

Perceptions are relative

An examination of the relationship between relative satisfaction metrics and share of wallet

Timothy Lee Keiningham

Ipsos Loyalty, Parsippany, New Jersey, USA

Bruce Cooil

Graduate School of Management, Vanderbilt University, Nashville, Tennessee, USA

Edward C. Malthouse

Department of Integrated Marketing Communication, Northwestern University, Evanston, Illinois, USA

Alexander Buoye and Lerzan Aksoy

Schools of Business, Fordham University, Bronx, New York, USA, and

Arne De Keyser and Bart Larivière

Department of Management, Innovation and Entrepreneurship, Ghent University, Ghent, Belgium

Abstract

Purpose – There is general agreement among researchers and practitioners that satisfaction is relative to competitive alternatives. Nonetheless, researchers and managers have not treated satisfaction as a relative construct. The result has been weak relationships between satisfaction and share of wallet in the literature, and challenges by managers as to whether satisfaction is a useful predictor of customer behavior and business outcomes. The purpose of this paper is to explore the best approach for linking satisfaction to share of wallet.

Design/methodology/approach – Using data from 79,543 consumers who provided 258,743 observations regarding the brands that they use (over 650 brands) covering 20 industries from 15 countries, various models such as the Wallet Allocation Rule (WAR), Zipf-AE, and Zipf-PM, truncated geometric model, generalization of the WAR and hierarchical regression models are compared to each other.

Findings – The results indicate that the relationship between satisfaction and share of wallet is primarily driven by the relative fulfillment customers perceive from the various brands that they use (as gauged by their relative ranked satisfaction level), and not the absolute level of satisfaction.

Practical implications – The findings provide practical insight into several easy-to-use approaches that researchers and managers can apply to improve the strength of the relationship between satisfaction and share of wallet.

Originality/value – This research provides support to the small number of studies that point to the superiority of using relative metrics, and encourages the adoption of relative satisfaction metrics by the academic community.

Keywords Customer behaviour, Consumer satisfaction

Paper type Research paper



Managers widely believe that customer satisfaction is a fundamental determinant of long-term consumer behavior (Oliver, 1980; Yi, 1990). This widespread acceptance has made customer satisfaction the most widely used metric in the measurement and management of consumer loyalty (Aksoy, 2013a; Zeithaml *et al.*, 2006). Companies spend substantial amounts of money to measure and manage customer satisfaction. For example, Inside Research (2012) found that for the 13 marketing research firms that responded to their survey, revenue from US-only customer satisfaction research would exceed \$750 million – this figure likely underestimates the total spent with marketing research firms given the small number of responding firms.

A review of the scientific literature on customer satisfaction supports management's focus on customer satisfaction. In particular, many studies have linked customer satisfaction to customers' purchasing behaviors (e.g. Bolton, 1998; Mittal and Kamakura, 2001; Rust and Zahorik, 1993). A close examination of the research regarding customer satisfaction and customers' share of category spending, however, reveals that while the relationship is positive, it tends to be very weak (Hofmeyr *et al.*, 2008; Mägi, 2003).

Because of the weak relationship, managers have difficulty connecting their efforts to improve satisfaction with tangible financial outcomes (Aksoy, 2013a; Keiningham *et al.*, 2014). For example, in an article entitled "Proof that it pays to be America's most-hated companies," *Bloomberg Businessweek* reported that their analysis of the relationship between the American Customer Satisfaction Index (ACSI) and stock performance found the relationship to be negative. Specifically, the magazine reported (Chemi, 2013):

[...] customer-service scores have no relevance to stock market returns [...] the most-hated companies perform better than their beloved peers [...] Your contempt really, truly doesn't matter [...] If anything, it might hurt company profits to spend money making customers happy.

Results such as these have led to calls by some managers and researchers to discontinue the measurement and management of satisfaction (Gupta and Zeithaml, 2006). Books like *Customer Satisfaction is Worthless*, *Customer Loyalty is Priceless*, by consultant Jeffery Gitomer (1998), and articles like "Customer satisfaction: it is dead, but it will not lie down," by researchers Williams and Visser (2002) are indicative of this general frustration.

Given customer satisfaction's weak relationship to business outcomes and customer behaviors, Mägi (2003, p. 104) argues "it might be informative to use relative measures of satisfaction when predicting customer share" (i.e. share of wallet). Researchers agree that perceptual metrics such as satisfaction need to be measured relative to competitive alternatives (e.g. Varki and Rust, 1997). Furthermore, there is a large body of research confirming the influence of competitive comparisons on both choice and post-purchase evaluations (e.g. Rust *et al.*, 2000).

The small number of studies that have used relative satisfaction in the scientific literature (e.g. Bolton *et al.*, 2000; Bowman and Narayandas, 2004; Hardie *et al.*, 1993; Wind, 1970) point to the superiority of relative metrics in linking to customer behavior. Nonetheless, the scientific community has been slow to use relative satisfaction in their research. None of the methods used by these researchers have been widely used in other scientific investigations. Rather, the overwhelming majority of scientific research investigating satisfaction relies on absolute metrics on a single firm. Furthermore, these methods are rarely used by managers.

The same reluctance to use relative metrics cannot be said for the practitioner community. Some of the world's largest survey research organizations specifically

advocate the use of relative metrics when linking customer satisfaction to a customer's share of wallet, and make them the foundation of their brand equity and customer experience measurement approaches, i.e., TNS: Conversion Model (Louw and Hofmeyr, 2012), Ipsos: Brand Value Creator (Hofmeyr *et al.*, 2008), and Ipsos: Wallet Allocation Optimizer (Keiningham *et al.*, 2011). These firms report strong correlations between their approaches and share of wallet.

The creators of these frameworks have made them widely available for managers to apply in their organizations by publishing their methodologies. Each of these approaches, however, uses a different technique to link relative metrics to share of wallet. Furthermore, despite their publication, these methodologies are not often used by managers outside of their application within a research firm's specific product offer. This, however, does not mean that they are not widely used. For example, the Conversion Model is used by "80% of the world's most valuable brands" (TNS, 2012).

The gap between the science and the practice of marketing in this regard has profound implications for both managers and researchers. There is no research in the peer reviewed literature that rigorously investigates various methodologies to determine their efficacy. As a result, researchers and managers are left with almost no guidance as to the usefulness of different approaches, or even to the validity of relative satisfaction metrics in general.

Additionally, if relative metrics more accurately reflect the relationship between satisfaction and customers' share of category spending, this would likely serve as impetus for new research in a number of areas. Clearly, this would necessitate new research into the relative nature of satisfaction and its corresponding impact on consumer behavior. It would also likely spur examinations into the potential relative impact of other perceptual and attitudinal metrics on consumer behavior (e.g. commitment, emotions, etc.).

As a result, there is a need for research regarding the efficacy of relative satisfaction metrics and best practices regarding the use relative satisfaction metrics. This research fills these gaps by investigating the relationship between relative satisfaction and customers' share of category spending (i.e. share of wallet) using data from 79,543 consumers who provided 258,743 observations regarding the brands that they use within a particular industry category. Data included ratings of over 650 brands in 20 industries from 15 countries.

The results of this investigation find that relative satisfaction significantly outperforms absolute satisfaction levels in linking to customers' share of category spending. Models based upon absolute satisfaction levels were consistently the worst performing models investigated. Moreover, we find that the most commonly used power laws in practice perform well compared to other models investigated in linking relative satisfaction to share of wallet. Finally, we note that there are significant differences in the complexity of the various approaches examined. Therefore managers need to consider the trade-off between relationship strength and complexity when selecting the best approach for use within their firms.

Structure of manuscript

This investigation relies upon a rigorous investigation of different power laws and hierarchical regression models. As a result, a thorough description of the investigation requires a detailed presentation of several models and analytic procedures. This has the potential to make the paper quite technical and fragmented, resulting in a paper that is

difficult for most managers to read. As a result, we believe that the core message of the paper can be lost in the technical descriptions of the models and analytics.

Therefore, in an effort to maximize the readability and insights gleaned from this investigation, this paper is divided into two main sections. The first section focuses on the theoretical foundation, core findings, and implications of the research. The second section is a Technical Appendix that provides a detailed overview of the models examined, and the various approaches used to investigate the properties of these models.

By using this approach, we hope that we are able to provide researchers and managers with clear and relevant insights while maintaining scientific rigor and transparency regarding our analyses and findings.

Theoretical background

Customer satisfaction

Satisfaction is the consumer's emotional response to the fulfillment of needs, expectations, wishes or desires. Specifically, Oliver (2010, p. 8) defines customer satisfaction as follows: "Satisfaction is the consumer's fulfillment response. It is a judgment that a product/service feature, or the product or service itself, provided (or is providing) a *pleasurable* level of consumption-related fulfillment, including levels of under- or overfulfillment."

Researchers have extensively examined the theoretical underpinnings of the satisfaction construct (e.g. Fornell *et al.*, 1996; Luo and Bhattacharya, 2006; Oliver, 1997). Researchers have also investigated the effects of customer satisfaction on future consumer behaviors (e.g. Crosby and Stephens, 1987; Keiningham *et al.*, 2003; Luo and Homburg, 2007).

Of particular importance to this investigation, there is general agreement among researchers and practitioners that satisfaction is relative to perceived competitive alternatives (e.g. Birtchnell, 1994; Holt and Huber, 1969; Varki and Rust, 1997; Semon, 1994). For example, Woodruff *et al.* (1983) argue that norms based on consumer experiences with brands within a product category and relative to competing alternatives in that category were a more natural comparison standard than focal brand expectations. Research by Cadotte *et al.* (1987) found that experience-based norms better explain variations in satisfaction than focal brand expectations. Additionally, Gardial *et al.* (1994) found that consumers tend to rely on competitive comparisons/norms when evaluating their consumption experiences.

This can in part be explained by expectancy-disconfirmation model of the appraisal sequence for satisfaction (Oliver, 2010, pp. 355-360). Oliver (2010, p. 22) defines expectancy-disconfirmation as "the psychological interpretation of an expectation-performance discrepancy. Consumers would describe this concept in terms of the performance of a product or service being better or worse than expected."

Although satisfaction and disconfirmation are not perfectly correlated, "satisfaction results primarily from disconfirmation" (Rust *et al.*, 1996, p. 234). As such, expectations tend to play a strong role in consumers' satisfaction judgments.

Consumers' expectations are strongly affected by their experiences. Experiences, however, are not limited to the focal/purchased brand, but frequently include broader experiences within a product or service category (Woodruff *et al.*, 1983). In addition, expectations may be affected by advertising and word of mouth (Boulding *et al.*, 1993; Miller, 1977). This, to a large degree, explains why satisfaction is influenced by competitive comparisons or norms.

Customer satisfaction and share of wallet

The relationship between satisfaction and consumer behavior is grounded in the theory of planned behavior (Ajzen and Madden, 1986), an offshoot of the theory of reasoned action (Ajzen, 2001; Ajzen and Fishbein, 1980). The theory argues that behaviors are influenced by three factors: attitudes, subjective norms, and perceived behavioral control. Specifically, favorable/unfavorable attitudes, in combination perceived societal “norms” are the primary determinants of a consumer’s intention to perform a behavior (provided the consumer believes he/she has the ability to perform the behavior). Although satisfaction is generally viewed as a perception (e.g. Oliver, 1980) this reflects the generally accepted view of how satisfaction ultimately influences consumer purchase decisions (Mittal and Frennea, 2010).

Share of wallet is widely believed to be driven in part by customers’ perceptions of the brands they use. The chain of effects can be thought of as product/service performance → satisfaction → share of wallet. In fact, this chain of effects is a logical adaptation of the core chain of effects proposed in some of the seminal models in marketing (Keiningham *et al.*, 2005); SERVQUAL (Parasuraman *et al.*, 1988; Zeithaml *et al.*, 1996), service profit chain (Heskett *et al.*, 1994), return on quality (Rust *et al.*, 1995), and the satisfaction profit chain (Anderson and Mittal, 2000).

The idea that customer satisfaction should link to share of category spending is intuitive (i.e. we tend to spend more with firms that better satisfy us). A large body of research does support this positive relationship (e.g. Baumann *et al.*, 2005; Bowman and Narayandas, 2004; Cooil *et al.*, 2007; Keiningham *et al.*, 2003, 2005; Larivière, 2008; Mägi, 2003; Perkins-Munn *et al.*, 2005; Silvestro and Cross, 2000).

The problem from a managerial perspective, however, is that while there tends to be a statistically significant positive relationship between satisfaction and share of wallet, the percentage of variance explained by this relationship is low (Hofmeyr *et al.*, 2008; Mägi, 2003). As a result, managers have openly challenged “whether the relationship between unobservable measures such as customer satisfaction and observable behavior such as purchasing was sufficiently strong to justify its use as the primary unobservable predictor” (Gupta and Zeithaml, 2006, p. 721).

Researchers have proposed two possible reasons to explain this weak relationship. First, customers appear to differ in their sensitivity to variations in satisfaction (Hofmeyr and Parton, 2010). For example, demographic differences have been shown to impact the satisfaction-share of wallet relationship (Cooil *et al.*, 2007). Second, researchers argue that satisfaction’s impact on customer behavior is nonlinear and asymmetric (e.g. Anderson and Mittal, 2000; Crofts *et al.*, 2008; Keiningham and Vavra, 2001). Accounting for the asymmetric, non-linear pattern of satisfaction has improved the relationship between satisfaction and share of wallet (e.g. Bowman and Narayandas, 2004; Keiningham *et al.*, 2003). Nonetheless, a large portion of the variance remains unexplained (Hofmeyr and Parton, 2010).

An alternative explanation for the weak relationship has been proposed by members of the practitioner community. Hofmeyr and Parton (2010) argue that the overriding reason for the asymmetric, non-linear relationship between satisfaction and share of wallet is not the absolute level of satisfaction per se. Rather at some point higher/lower levels of satisfaction correspond to a shift in a customer’s preference ranking for a brand *vis-à-vis* competitive brands that the customer also uses. As a result, Hofmeyr and colleagues (Hofmeyr *et al.*, 2008; Hofmeyr and Parton 2010) argue that the focus of satisfaction research should shift from absolute satisfaction levels to the relative preference rank that a brand’s satisfaction level represents among

competing brands used by customers to improve the strength of the relationship between satisfaction and share of wallet.

Relative measures

There is a large body of research confirming the influence of competitive comparisons on both choice and post-purchase evaluations (e.g. Gardial *et al.*, 1994; Rust *et al.*, 2000; Woodruff *et al.*, 1983). For a review of the psychology literature associated with relative thinking in the pre- and post-purchase consumption process, we refer the reader to Keiningham *et al.* (2014).

Relative thinking is central to the consumer decision process. For example, Jacoby and Chesnut (1978, p. 88) argue that “brand loyalty is a function of decision making, evaluative processes. It reflects a purchase decision in which the various brands have been psychologically (perhaps even physically) compared and evaluated on certain internalized criteria, the outcome of this evaluation being that one or more brands was (were) selected.”

Similarly, Dick and Basu (1994, pp. 100-101) observe, “Attitudes have been related to behaviors, although it is important to note that one may hold a favorable attitude toward a brand but not purchase it over multiple occasions because of comparable or greater attitudinal extremity toward other brands. For purposes of predictive validity, it is hence advantageous to compare brands that are viewed by consumers to be relevant in a given consumption context. The nature of relative attitudes is likely to provide a stronger indication of repeat patronage than the attitude toward a brand determined in isolation.”

Despite this recognition, academic research has overwhelmingly focused on absolute metrics. There are, however, some notable exceptions. Table I provides a brief summary of the research to date regarding the use of relative measures in the scientific literature.

An examination of the research in Table I supports the superiority of relative metrics in linking to customer intentions and behaviors. Interestingly, none of the methods used by these researchers have been widely employed in other scientific investigations. Furthermore, these methods are rarely used by managers.

Instead, the most prominent voices for the use of relative measures in the prediction of share (specifically market share and share of wallet) and the most widely used methodologies come from practitioners. The first widely adopted approach was customer value analysis (CVA), advocated by Bradley Gale (1994) in the book *Managing Customer Value*. One of the primary points of differentiation of the CVA approach was its incorporation of relative brand position in linking customer perceptions to business outcomes, most notably market share. At one time this metric was widely used in industry, although it has fallen out of favor because of underlying statistical issues with the ratios used in the process (Keiningham and Vavra, 2001, pp. 41-44) and the inability of many firms to validate the claimed link to market share (Keiningham *et al.*, 2008).

Hofmeyr *et al.* (2008) introduced a new brand “attitudinal equity” (AE) measure using the Zipf distribution (Zipf, 1935)[1]. The AE measure was calculated by transforming satisfaction (or other perceptual/attitudinal metrics) into relative ranks. Specifically, to transform a customer’s satisfaction ratings to ranks, the highest satisfaction rating a customer gave to a brand in his/her usage set would be assigned a “1,” the second highest a “2,” and so on; in the case of ties, the average is used for the ranks that would have been used had there been no ties. These ranks were then

Table I.
Summary of the
research to date
regarding the use of
relative measures

Study	Setting	Study type	Relative metric operationalization	Outcome	Most important findings/ propositions
Wind (1970)	Electronics industry	Research paper	Two relative metrics are used: 1. Relative attitude towards an ideal supplier 2. Relative attitude towards competitors (i.e. second favorite supplier)	Share-of-wallet	The relative attitude towards competitors is found to be one of the most important indicators of source loyalty
Hauser (1991)	Major consumer-product category	Research paper	Satisfaction rating relative to competition	Primary brand share	Relative scales are found to significantly outperform absolute scales in linking to the primary brand share
Hardie <i>et al.</i> (1993)	Retailing industry (Orange juice purchases)	Research paper	Econometric reference-dependent choice model (multinomial logit formulation)	Brand choice	Reference dependent models clearly outperform nonreference-based models, resulting in a better prediction of brand choice
Dick and Basu (1994)	na	Conceptual Paper	Relative attitude defined as the degree to which a customer's evaluation of one product/brand dominates that of other alternatives	Repeat patronage	The inclusion of relative attitudes is likely to result in higher predictive ability for loyalty compared to single-brand attitudes
Van den Putte <i>et al.</i> (1996)	1. Broadcasting industry 2. National/regional elections	Research paper	Two relative scales are used: 1. Indirect relative rank order scale 2. Direct relative rank order scale	1. Buying intention 2. Voting intention	Behavioral alternative models applying direct relative rank order scales have the best predictive power, significantly improving average explained variance of behavioral intentions compared to standard, non-relative scales

(continued)

Study	Setting	Study type	Relative metric operationalization	Outcome	Most important findings/propositions
Varki and Rust (1997)	Financial services industry	Research paper	Refinement of analysis of variance (ANOVA) for attribute satisfaction ratings	Customer satisfaction	The refined ANOVA-method allows firms to identify their relative performance to competitors at an attribute level, allowing for a better management practice
Bolton <i>et al.</i> (2000)	Financial services industry	Research paper	Gain/loss satisfaction scores by comparing focal brand and competitor ratings	Repeat patronage	Customers make re-patronage decisions on the basis of prior re-patronage intentions or behavior, updated by comparing their prior satisfaction level with the company vs that with the competitor(s)
Olsen (2002)	Retailing industry	Research paper	Comparative-attribute based survey format (i.e. quality/satisfaction questions for different alternatives are posed in sequence, making them salient for comparative evaluation)	Repurchase frequency	Using a comparative assessment, as opposed to an absolute measurement, results in higher predictive power and stronger relationships between quality, satisfaction and loyalty
Kumar (2002)	IT products and services industry	Research paper	Satisfaction gains and losses are computed using the proportional difference between the focal and competing firms	Repurchase intentions	Customers' repurchase intentions depend both on the satisfaction level with the supplier in question and the corresponding satisfaction level and costs of its referent competitor

(continued)

Perceptions
are relative

Table I.

Table I.

Study	Setting	Study type	Relative metric operationalization	Outcome	Most important findings/propositions
Bowman and Narayandas (2004)	Metal industry	Research paper	Satisfaction with the closest competitor. (0 if lower than focal vendor; 1 if equal or higher than focal vendor)	Share-of-wallet	Satisfaction with the closest competitor has a direct, negative impact on share-of-wallet
Rust <i>et al.</i> (2004)	Airline industry Electronic stores Facial tissues Grocery stores Rental cars industry	Research paper	Customer ratings on several customer-equity drivers are collected for four to five leading brands in each industry, and imputed in a multinomial logit regression model	Customer lifetime value	The developed CLV-model allows considering the impact of competitive responses on a firm's customer equity, and provides insight into competitive strengths and weaknesses
Ahearne <i>et al.</i> (2007)	Pharmaceutical industry	Research paper	Average ratings of competition are subtracted from the focal vendor's service quality and relationship quality measures	Share-of-wallet	Relative service quality evaluations are found to drive relationship quality, which in turn affects share-of-wallet

transformed to share of wallet estimates using the Zipf distribution. The parameters^[2] of the Zipf distribution were determined by fitting the relationship between the rank of a brand and the corresponding share of wallet that the customer allocated to that brand. (For the remainder of this paper, we will refer to this model as Zipf-AE.)

The results of the Zipf-AE approach showed a large improvement in model R^2 . In particular, Hofmeyr *et al.* report that the average R^2 between customer satisfaction and customers' share of wallet using absolute measures was 0.24, while using the rank-based Zipf-AE transformation resulted in a 0.44 R^2 .

Keiningham *et al.* (2011) introduced a power law for transforming relative "ranked" satisfaction into share of wallet predictions which they called the Wallet Allocation Rule (WAR). Satisfaction ranks were calculated using the same approach as Hofmeyr *et al.* (2008). WAR is a fixed parameter model; as such, no estimation (i.e. data fitting) is required to estimate the relationship between rank transformed satisfaction and share of wallet. Keiningham *et al.* (2011) report that changes in customers' WAR scores and changes in their share of wallet over time showed a correlation of approximately 0.8, which corresponds to an R^2 of approximately 0.6.

Recently Louw and Hofmeyr (2012) proposed what they described as "an improvement to the original measure of brand attitudinal equity proposed by Hofmeyr *et al.* (2008, p. 10)" which they refer to as a measure of "power of the mind" (PM). As with Hofmeyr *et al.* (2008), the calculation of PM is also based upon the Zipf distribution. (For the remainder of the paper, we will refer to this model as Zipf-PM.)

The primary distinguishing characteristic between the Zipf-AE and Zipf-PM approaches is that Zipf-PM uses "the share that a brand's rating achieves as a percent of the sum of a respondent's ratings of relevant brands" in the Zipf distribution equation (Louw and Hofmeyr, 2012, p. 11).

Louw and Hofmeyr (2012) report that the Zip-PM approach has a higher correlation to share of wallet "by a very small margin" (p. 14) than the Zip-AE and WAR approaches. It is important to note, however, that the comparison made in their investigation was not apples-to-apples; WAR and Zipf-AE were calculated using a single satisfaction question, whereas Zipf-PM was calculated using a combination of two questions in their comparison. Even with this difference, however, there was very little difference in terms of variance explained between the three approaches.

The Zipf and WAR approaches have received a great deal of attention by market researchers. Moreover, both the Zipf-AE (Hofmeyr *et al.*, 2008) and WAR (Keiningham *et al.*, 2011) approaches have received important industry awards for innovation (Gesulado, 2011; Humphrey, 2008).

The primary use of these approaches in practice is within specific products offered by two of the world's largest market research firms. Specifically, Ipsos and TNS use these power laws as core components of their brand equity and customer experience management approaches. As a result, it would be difficult to overstate their use by managers through the use of products offered by these firms. Even if we assume 100 percent overlap of clients, the research firms using these approaches work with over 5,000 different companies worldwide (Ipsos, 2012).

These approaches are not yet widely used by managers outside of the specific product offerings of these firms. As these approaches are not "black boxes" (i.e. these methods are published) and the creators actively promote these approaches (e.g. Hofmeyr, 2012; Keiningham, 2012), however, marketing managers are increasingly aware of the call for relative metrics to more strongly link satisfaction and share of wallet (e.g. Keiningham *et al.*, 2014).

Moreover, while the call for relative metrics has largely come from practitioners, there is early evidence that the academic community has taken notice. For example, Rust and Huang (2014, p. 4) argue that Keiningham (2014) “show convincingly that relative metrics (relative to competitors) are essential.”

Research objectives

The primary purpose of this study is to examine the relationship between relative satisfaction and share of wallet. As noted earlier, the research to date tends to support the superiority of relative perceptual and attitudinal metrics to monadic metrics in correlating to consumer buying behaviors such as share of wallet (e.g. Bowman and Narayandas, 2004; Hofmeyr *et al.*, 2008; Keiningham *et al.*, 2011). Therefore, we hypothesize:

- H1.* Ranked satisfaction levels are more strongly correlated to share of wallet than are absolute satisfaction levels.

Furthermore, although the empirical research appears to confirm the link between absolute satisfaction and share of wallet across various industries such as fleet trucking (Perkins-Munn *et al.*, 2005), pharmaceutical (Perkins-Munn *et al.*, 2005), institutional securities (Keiningham *et al.*, 2005), retail banking (Baumann *et al.*, 2005), processed metals (Bowman and Narayandas, 2004), and grocery retailing (Mägi, 2003; Silvestro and Cross, 2000), the majority of this research has relied on cross sectional data. Although longitudinal examinations of the effect of customer satisfaction on other performance measures have found a positive relationship to customer retention (Bolton, 1998), firm revenues and shareholder value (Anderson *et al.*, 2004), the impact on share of wallet is limited. One exception is the longitudinal share of wallet study by Cool *et al.* (2007) where results indicate a positive relationship between changes in satisfaction and changes in share of wallet over time. In line with these findings, we would expect longitudinal ranked satisfaction levels to link to changes in share or wallet over time. Therefore we hypothesize:

- H2.* Changes over time in ranked satisfaction levels are more strongly correlated to contemporaneous changes in share of wallet than are changes in absolute satisfaction levels.

In addition to testing the two hypotheses above, another important goal of this investigation is to provide insight into the most widely used approaches for linking satisfaction and SOW in practice, i.e., WAR (Keiningham *et al.*, 2011), Zipf-AE (Hofmeyr *et al.*, 2008), and Zipf-PM (Louw and Hofmeyr, 2012). In particular, we examine each of the proposed power laws to determine their efficacy in predicting SOW from ranked satisfaction. As noted earlier, to date there is no research in the peer-reviewed scientific literature that examines these various methods to determine their efficacy. Also, we seek to identify better approaches (if any) to link relative satisfaction levels to share of wallet.

Data and measures

Data collection

The data were collected by a large marketing research firm as part of its global norms database. In total, the data consisted of 79,543 customers providing 258,743 observations regarding the brands that they use within a particular industry category. Each respondent in the database used two or more brands in the category (i.e. single-brand

users were not included in our database for analysis since their SOW is, by definition, one).

Industries and brands. Data included ratings of over 650 brands in 20 industries. Airlines represented the largest industry in terms of number of respondents, although it should be noted that retail was broken out into more homogeneous subgroups. The complete industry breakdown is: airline (44.9 percent), asthma Rx OTC (0.4 percent), automobiles (0.3 percent), baby retail (1.8 percent), beauty (1.7 percent), clothing retail (2.4 percent), credit card (4.3 percent), DIY retail (0.7 percent), drugstores (1.0 percent), electronics retail (2.0 percent), furniture (2.9 percent), general retail (8.0 percent), grocery retail (13.9 percent), mass merchandise retail (0.5 percent), mobile phone carrier (0.03 percent), office supply (0.6 percent), personal computers (0.2 percent), pharmacy (1.6 percent), printer supplies (2.1 percent), and retail banking (10.7 percent).

Countries. Respondents were sampled from 15 countries, with the majority from the USA. The percentage of respondents from each country is: Australia (0.4 percent), Brazil (3.3 percent), China (0.8 percent), Denmark (0.6 percent), Finland (0.5 percent), Germany (0.6 percent), Italy (8.2 percent), the Netherlands (0.4 percent), Norway (0.6 percent), Peru (0.3 percent), South Africa (0.2 percent), Sweden (0.6 percent), Turkey (1.1 percent), the UK (10.8 percent) and the USA (71.7 percent).

Gender. In terms of total respondents, 51 percent of respondents are female, 49 percent male. The percentage of female respondents for each country is: Australia (30 percent), Brazil (43 percent), Denmark (31 percent), Finland (40 percent), Italy (29 percent), the Netherlands (48 percent), Norway (34 percent), Peru (31 percent), South Africa (20 percent), Sweden (35 percent), Turkey (21 percent), the UK (52 percent), and the USA (53 percent). Gender was not available in the database for Chinese and German respondents.

Age. The average age for all respondents is 49. The average age for respondents in each country is: Australia (48), Brazil (40), China (34), Denmark (49), Finland (45), Germany (38), Italy (48), the Netherlands (47), Norway (45), Peru (41), South Africa (47), Sweden (49), Turkey (34), the UK (48) and the USA (50).

Longitudinal data. A subset of these respondents (all from the USA) were contacted 6 months following the initial survey to provide longitudinal information regarding changes in satisfaction ratings and changes in share of wallet. The longitudinal data consisted of 1,138 customers providing 2,686 observations on the same brands in both periods 1 and 2. These customers provided a total of 3,228 rankings in period 1 and 3,365 rankings in period 2. These 1,138 customers were chosen because they ranked at least two brands in each period. We needed at least two brands from each customer in period 1 in order to be able to use their information to help estimate model parameters. Also, we needed at least two brands per customer in period 2 in order to estimate SOW < 100 percent (i.e. when number of brands equal one, SOW is by default 100 percent).

Gender distribution for the longitudinal sample is approximately even (51 percent male, 49 percent female) with an average age of 55.6. Breakdown of respondents by industry is as follows: grocery (13.4 percent), drugstore (13.4 percent), pharmacy (4.2 percent), mass merchandisers (10.1 percent), retail bank (0.5 percent), asthma Rx (7.9 percent), DIY (17.0 percent), office supply (13.5 percent), airline (12.2 percent), computers (3.1 percent), mobile phone carrier (0.4 percent), and automobiles (4.1 percent).

Constructs and measures

Customer satisfaction. Following Mittal *et al.* (1999) we measured overall satisfaction with the brand using a single item (1 = completely dissatisfied, 10 = completely

satisfied). Satisfaction levels were converted to ranks using the approach of Hofmeyr *et al.* (2008) discussed earlier.

It is important to note that relative “ranked” satisfaction is not a single-item construct in the commonly used sense. Rather ranks for customers when “number of brands ≥ 2 ” are based upon consumers’ perceptions of multiple brands. In example form, imagine that Brand A has a 7 in satisfaction on a ten-point scale. Its rank will depend on Brand B. If Brand B rates a 5, then Brand A is rank = 1. If Brand B rates a 9, then Brand A is rank = 2. In other words, the same satisfaction level can result in different ranks as information from all brands used by a respondent is used to determine rank. (Note: In this investigation, all respondents used two or more brands.)

With regard to the use of single-item measures in general, although marketing academics typically prefer multi-item measures, single-item measures of overall satisfaction have been used in many prior studies and shown to perform adequately (e.g. Bolton, 1998; Bolton and Lemon, 1999; Cooil *et al.*, 2007; Crosby and Stephens, 1987; Drolet and Morrison, 2001; Mittal and Kamakura, 2001; Mittal *et al.*, 1998, 1999).

Bergkvist and Rossiter (2007) have demonstrated that single-item measures achieve the same predictive ability as multi-item measures, provided that the focal construct is concrete and singular in nature. Satisfaction would appear to meet this standard. Zeithaml *et al.* (2006, p. 170) observe, “Customer satisfaction is the most widely used perceptual metric because it is generic and can be universally gauged for all products and services (including nonprofit and public services). Even without a precise definition of the term, customer satisfaction is clearly understood by respondents, and its meaning is easy to communicate to managers.”

Moreover, psychometric analyses conducted by Drolet and Morrison (2001) finds that the incremental information from even the second or third item in a multi-item scale contributes very little to the information obtained from the first item in a multi-item scale. They also find that “added items actually aggravate respondent behavior, inflating across-item error term correlation and undermining respondent reliability” (p. 196).

Of particular relevance to this investigation, Hofmeyr *et al.* (2008) and Keiningham *et al.* (2011) specifically create ranks based upon responses to a single-item measure. This is not surprising given that in practice most firms use single-item measures of satisfaction (Morgan *et al.*, 2005), and these approaches were developed in large part by industry practitioners. Therefore, it is appropriate to apply this same approach in our investigation of these methods.

It is important to note, however, that the longitudinal data examined in this analysis also contained the survey measures used in the ACSI to measure overall customer satisfaction, specifically: overall satisfaction (as used in the single item measure), performance relative to expectations, and performance relative to the customer’s ideal (Fornell *et al.*, 1996). Therefore, to be certain that our findings were robust we compared the overall satisfaction measure with two reliable composites of these three questions: both the average of all three and the first principal component of the three items. The average and first principal component are essentially the same (the correlation between the two summaries is 1.000 across both periods) and overall satisfaction has a correlation of 0.95 with each. Given this equivalence, the single-item measure is preferred as the most direct estimate of overall satisfaction.

Share of wallet. Following the approach of Cooil *et al.* (2007), share of wallet was measured as the percent of spending in the category that respondents allocate to the various brands that they use.

Analysis

Description of the relationship between SOW and rank

As noted earlier, research consistently finds that correlation between satisfaction and SOW (at the individual customer level) is very weak. A core argument of the Zipf-AE, Zipf-PM, and WAR approaches under investigation is that relative “ranked” satisfaction is more strongly correlated to SOW. Therefore, the first step was to test the accuracy of this claim.

Table II summarizes the correlations and partial correlations between SOW and both rank and satisfaction (absolute). It also includes correlations with logarithmic transformations of each variable and the logit transformation of SOW[3].

The correlations of SOW, and transformations of SOW, with *Rank* and $\log(Rank)$ are invariably stronger than the correlations of SOW, and the transformations of SOW, with the two versions of *Satisfaction*. Nevertheless all correlations are highly significant ($p < 0.001$; which is not surprising given the sample size, $n = 258,743$). The strongest relationships are for $\log(Rank)$ with SOW and $\text{logit}(SOW)$; $\log(Rank)$ explains 30 percent (or $r^2 \times 100\%$, with $r = -0.545$) of the variance in SOW and 29 percent of the variance of $\text{logit}(SOW)$ ($r = -0.536$). The largest nominal correlation with *Satisfaction* are for *Satisfaction* with the $\text{logit}(SOW)$ ($r = 0.239$), which indicates it accounts for 5.7 percent of the variation in $\text{logit}(SOW)$.

Remarkably, the correlations of *Rank* and $\log(Rank)$ with SOW and its transformations still remain strong and quite significant when we condition on *Satisfaction* levels, as seen from the partial correlations (the percent variance explained ranges from 19 to 26 percent in each case). In contrast, the partial correlations of *Satisfaction* and $\log(Satisfaction)$ with SOW and its transformations are actually negative, and correspond to R^2 values that are below 1 percent in absolute value in every case.

Our results provide strong evidence of the superiority of relative ranked satisfaction to absolute satisfaction in linking to SOW.

Investigating the models[4]

The next step in our analysis was to investigate the efficacy of the three most widely used power laws (i.e. Zipf-AE, Zip-PM, and WAR) in predicting SOW. A fair assessment, however, requires that we compare these power laws to other models that would be reasonably expected to perform similarly based upon the properties of these models.

	Log		Log		Partial correlations after removing <i>Satisfaction</i>		Partial correlations after removing rank (as $\log(Rank)$)	
	<i>Rank</i>	<i>Satisfaction</i>	<i>Rank</i>	<i>Satisfaction</i>	<i>Rank</i>	<i>Satisfaction</i>	<i>Rank</i>	<i>Satisfaction</i>
SOW	-0.484	-0.545	0.237	0.192	-0.437	-0.505	-0.010	-0.027
Log(SOW)	-0.492	-0.521	0.231	0.191	-0.448	-0.479	-0.003*	-0.016
Logit(SOW)	-0.491	-0.536	0.239	0.195	-0.443	-0.494	-0.002**	-0.019

Notes: $n = 258,743$. Except as indicated, all correlations are significant at the $p < 0.001$ level; * $p = 0.124$; ** $p = 0.209$

Table II.
Correlations between transforms of SOW and transforms of *Rank* and *Satisfaction*

The Zipf functions imply a Pareto decay in SOW as rank increases, which is distinct from a geometric decay and more rapid than the linear decay of the WAR model (when there are more than two brands). Therefore, to provide an additional reasonable point of comparison for the Zipf-AE and Zipf-PM power laws, we examine the effectiveness of the truncated geometric model in using ranked satisfaction to predict SOW.

Whenever possible, we examine three versions of these discrete distributions: a fixed-parameter version, a one-parameter version (i.e. the parameter does not vary by the total number of brands), and what we label as a nine-parameter version (i.e. the parameter varies by the total number of brands used; we consider customers who use from two to ten brands). It is important to note that there is no one-parameter version of WAR, and no established fixed-parameter version of the truncated geometric. In total, we explore ten versions of the discrete distribution models by including fixed-parameter, one-parameter and nine-parameter versions of the various models.

Additionally, because hierarchical regression models are commonly used in research and practice to assess the relationship between satisfaction and SOW (e.g. Keiningham *et al.*, 2003) we investigated these models as a point of comparison. In each of these models, a random effect at the customer level is used to accommodate the dependence among observations from the same customer within a product category. Specifically, we consider four hierarchical regression models (where for each of set of predictors, we estimate one version with common parameters across all m -categories, where m represents the total number of brands, and another with separate parameters within each m -category).

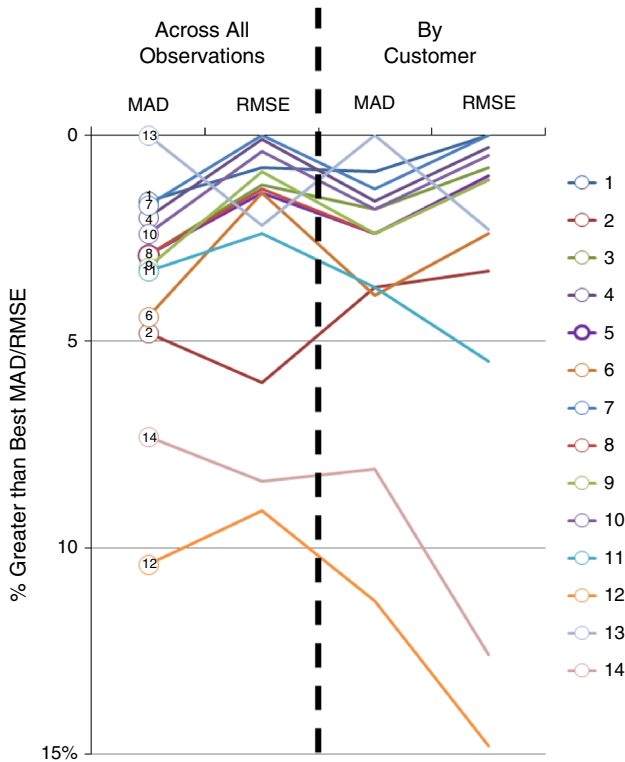
Overall model performance (cross-sectional)

To evaluate the overall performance, we first examined each model's ability to link customer satisfaction (absolute or relative ranked satisfaction) with SOW for the same time period. We assess each model's performance in four ways: mean absolute deviation (MAD), and root mean squared error (RMSE) across all observations and also by customer. Figure 1 shows the performance of each of the models relative to the best performing models[5].

The fixed-parameter versions of the discrete distribution models do remarkably well overall. Among these distributions, the fixed-parameter Zipf-AE model is best in terms of MAD, both overall and per customer, and it actually outperforms all models (including the regression models) in terms of average customer RMSE. The nine-parameter version of Zipf-AE is the best performer in terms of overall RMSE. Nevertheless, the discrete distributions generally do quite well: eight of the other ten discrete distributions have RMSE values that are within 1.5 percent of the best fit. The one exception is the fixed parameter Zipf-PM which has an RMSE that is 6 percent larger overall, relative to the best performing nine-parameter Zipf-AE model.

The nine-class regression with $\log(\text{Rank})$ is actually the best performing model in terms of MAD, and it is just ahead of the fixed parameter Zipf-AE with MAD values that are 1.6 percent and 0.9 percent larger overall, and per customer, respectively. This regression model is also uniformly the best among the four regression alternatives, but paradoxically it does not fit as well in terms of RMSE, where it actually achieves the 10th and 9th highest overall and per customer RMSE, respectively. Still, even in these cases its error rates are only larger than the lowest RMSE values by 2.2v overall, and 2.3 percent per customer.

By contrast, the regression models based on *Satisfaction* are uniformly the worst models in every case, and here the error rates are substantially larger than the best



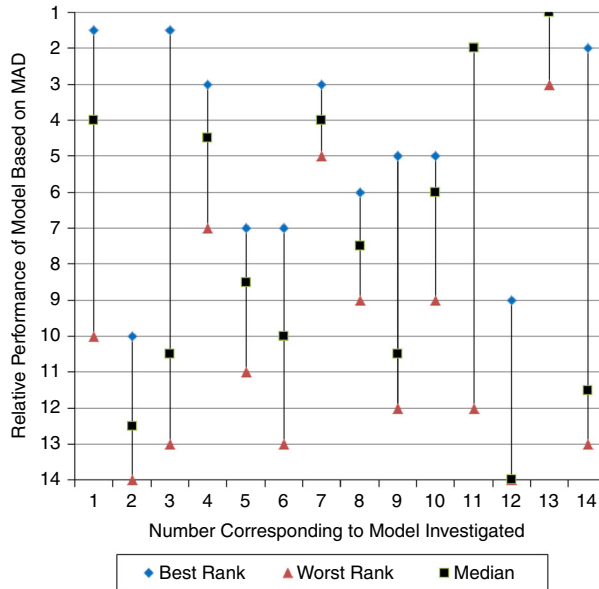
1. Zipf-AE, Fixed Parameter
2. Zipf-PM, Fixed Parameter
3. WAR, Fixed Parameter
4. Zipf-AE, One-Parameter
5. Zipf-PM, One-Parameter
6. Truncated Geometric, One-Parameter
7. Zipf-AE, Nine-Parameter
8. Zipf-PM, Nine-Parameter
9. WAR, Nine-Parameter
10. Truncated Geometric, Nine-Parameter
11. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Log}(\text{Rank}) + \beta_2 \text{Log}(\text{Total Brands} + 1 - \text{Rank})$, 1 Class
12. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Satisfaction} + \beta_2 (\text{Total Brands})$, 1 Class
13. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Log}(\text{Rank}) + \beta_2 \text{Log}(\text{Total Brands} + 1 - \text{Rank})$, 9 Classes
14. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Satisfaction}$, 9 Classes

Figure 1. Model performance overall and at the customer level in terms of mean absolute deviation (MAD) and root mean squared error (RMSE)

model in every instance. Although the nine-class version of this model is the better performer, even its error rates range from being higher by 7.3 percent (MAD overall) to 12.6 percent (RMSE per customer).

Overall model performance by number of brands used. In addition to examining overall performance, we investigated whether the number of brands used by the customer affect which model performs best. Figure 2 provides a comparison of model performance by the number of total brands that are used by the customer. An examination of Figure 2 shows that the relative performance of most models varies widely depending upon the number of brands used by the customer[6].

The fixed-parameter versions of Zipf-AE and WAR are the best in the two-product category with MAD values of 20.5 percent. WAR and Zipf-AE are equivalent in this case.



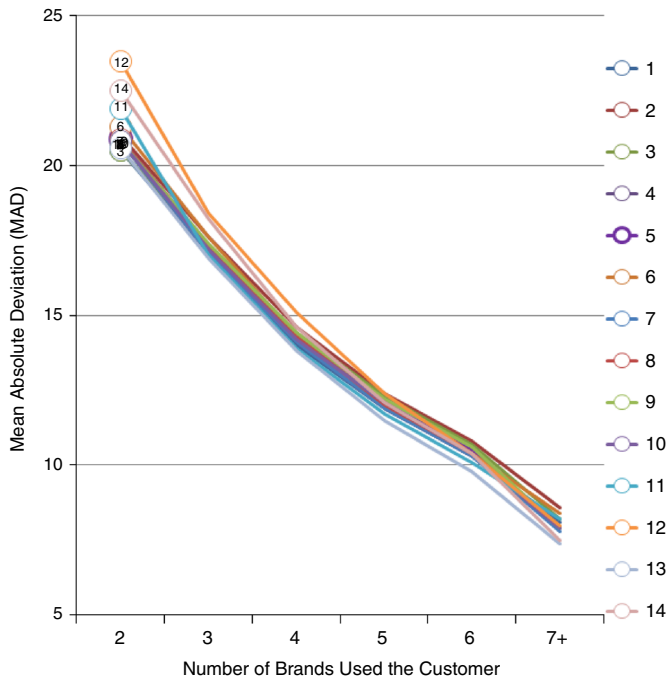
1. Zipf-AE, Fixed Parameter
2. Zipf-PM, Fixed Parameter
3. WAR, Fixed Parameter
4. Zipf-AE, One-Parameter
5. Zipf-PM, One-Parameter
6. Truncated Geometric, One-Parameter
7. Zipf-AE, Nine-Parameter
8. Zipf-PM, Nine-Parameter
9. WAR, Nine-Parameter
10. Truncated Geometric, Nine-Parameter
11. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Log}(\text{Rank}) + \beta_2 \text{Log}(\text{Total Brands} + 1 - \text{Rank})$, 1 Class
12. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Satisfaction} + \beta_2 (\text{Total Brands})$, 1 Class
13. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Log}(\text{Rank}) + \beta_2 \text{Log}(\text{Total Brands} + 1 - \text{Rank})$, 9 Classes
14. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Satisfaction}$, 9 Classes

Figure 2.
Relative model performance in terms of total brands used based upon mean absolute deviation (MAD)

This is the only category where the nine-class regression with $\log(\text{Rank})$ is not the best model, and even in the two-category case this regression model is nearly the best with a MAD that is 20.6 percent (relative to the best MAD of 20.5 percent).

The nine-parameter Zipf-AE model and the nine-class regression with $\log(\text{Rank})$ are the best overall performers across categories, and the Zipf-AE models are always among the top 5 models when total brands is less than seven ($m \leq 6$). Finally the regression models based on *Satisfaction* are the worst models overall, in terms of median rank across categories, although the nine-class regression on *Satisfaction* is the second best model in the last category ($7 \leq m \leq 10$). The regression models based on *Satisfaction* are uniformly the poorest performers when there are four or fewer total brands ($m \leq 4$).

It is important to note that while the relative performance of most models varies by the number of brands used, MAD values decrease as the total number of brands used increases (see Figure 3), which is to be expected, given that one is predicting smaller SOW values as the total number of brands increases. Across models, the lowest MAD values decrease by 64 percent as total brands increase across the six categories, and it ranges from 20.5 percent (when $m = 2$) to 7.4 percent (when $7 \leq m \leq 10$).



1. Zipf-AE, Fixed Parameter
2. Zipf-PM, Fixed Parameter
3. WAR, Fixed Parameter
4. Zipf-AE, One-Parameter
5. Zipf-PM, One-Parameter
6. Truncated Geometric, One-Parameter
7. Zipf-AE, Nine-Parameter
8. Zipf-PM, Nine-Parameter
9. WAR, Nine-Parameter
10. Truncated Geometric, Nine-Parameter
11. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Log}(\text{Rank}) + \beta_2 \text{Log}(\text{Total Brands} + 1 - \text{Rank})$, 1 Class
12. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Satisfaction} + \beta_2 (\text{Total Brands})$, 1 Class
13. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Log}(\text{Rank}) + \beta_2 \text{Log}(\text{Total Brands} + 1 - \text{Rank})$, 9 Classes
14. $\text{Logit}(SOW_i) = \beta_0 + \beta_1 \text{Satisfaction}$, 9 Classes

Figure 3.
Mean absolute
deviation (MAD) of
model in terms of
total brands used

Overall model performance (longitudinal)

For managers, the most important criterion for determining the success of any model of customer satisfaction is the strength of its relationship to changes in customer behavior (Oliver *et al.*, 1997, p. 312). Researchers similarly maintain that “marketers should examine changes in customer satisfaction over time due to customer “touches” (i.e. customer or firm-initiated encounters) as well as perceptions of competitors (e.g. Bowman and Narayandas, 2004)” (Bolton *et al.*, 2004, p. 277).

Figure 4 provides and analysis of the two-period data. The figure summarizes the correlations of change in SOW with contemporaneous changes in model estimates[7].

Figure 4 shows that the correlation between the two-period change in the WAR estimates of SOW (fixed-parameter version) and change in SOW are nominally the largest overall ($r = 0.407$, $p < 0.001$), but nearly all of the discrete distributions perform at the same level in terms of predicting change in SOW.

The weakest performing models are the nine-parameter truncated geometric model (for change in SOW: $r = 0.371$, $p < 0.001$) and the regression models (for change in

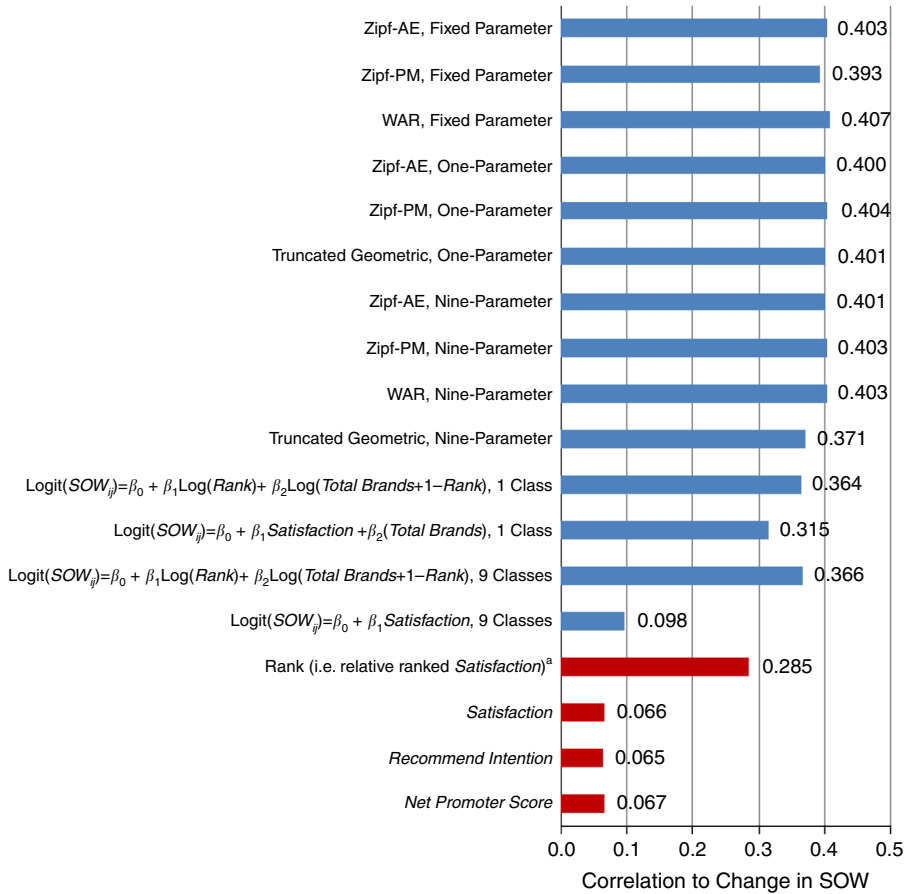


Figure 4.
Correlations between changes in SOW and changes in model estimates

Notes: $n \geq 2,686$ for each correlation. ^aAbsolute value used as the correlation between rank and SOW would be expected to be negative (i.e. the lower the number associated with rank, the higher the expected SOW). All correlations are significant at the level $p < 0.001$

SOW: the largest $r = 0.366$, $p < 0.001$). Surprisingly, the one-class regression model based on *Satisfaction* performs better than the nine-class version. The nine-class regression based on *Satisfaction* is by far the worst performer overall ($R^2 < 12$ percent in each case).

Finally, Figure 4 shows the inadequacy of changes in absolute satisfaction levels in correlating to changes in share of wallet. Changes in *Satisfaction* explain < 1 percent of the variation in changes in share of wallet ($r = 0.066$).

The disaffection with customer satisfaction has caused many managers to shift to a measure of recommend intention, specifically the Net Promoter Score (NPS), to gauge customer loyalty (Reichheld, 2003). Our results, however, clearly indicate that changes in a customer's Net Promoter classification similarly has almost no correlation to changes in share of wallet ($r = 0.067$). These results are comparable when using change in recommend intention levels ($r = 0.065$).

Results summary. *H1* postulated that “ranked satisfaction levels are more strongly correlated to share of wallet than are absolute satisfaction levels.” This study conclusively showed this to be true.

First, Table II shows that *Rank* and $\log(\text{Rank})$ account for 23 and 30 percent of the variance in *SOW* (the corresponding correlations are -0.484 and -0.545, respectively). In contrast, the percentage of variance explained by satisfaction is 5.7 percent ($r = 0.239$). Even more striking, the conditional correlations between *Rank* and $\log(\text{Rank})$ with *SOW* and $\text{logit}(\text{SOW})$ remain strong, when conditioning on *Satisfaction* levels, while the conditional (partial) correlations with *Satisfaction* and $\log(\text{Satisfaction})$ are not significant, when we condition on $\log(\text{Rank})$.

Further, the regressions based on *Satisfaction* in Figure 2 demonstrate that these models do not fit as well as the corresponding regression based on *Rank*, and almost without exception, the *Satisfaction* models are the worst performing models. The one exception occurs in the largest total brand category (see Table AI in the Appendix 1), where the nine-class regression with *Satisfaction* is second only to the nine-class regression model based on *Rank*. Finally, among the discrete distribution models, only the three Zipf-PM models use information on absolute satisfaction, and these are generally among the worst performing discrete distribution models. The findings therefore clearly indicate the superiority of using a relative ranked approach to customer satisfaction measurement compared to absolute satisfaction when attempting to link to a customers’ share of wallet. *H1* is therefore supported.

H2 postulated that “changes over time in ranked satisfaction levels are more strongly correlated to contemporaneous changes in share of wallet than are changes in absolute satisfaction levels.”

The two-period analysis summarized in Figure 4 shows how two-period changes in ranked satisfaction levels are more strongly correlated to contemporaneous changes in share of wallet, than are changes in absolute satisfaction levels. Change in *Rank* and $\log(\text{Rank})$ have substantially larger absolute correlations with change in *SOW* ($r = -0.285$, and $r = -0.332$, respectively) and with $\text{Logit}(\text{SOW})$ ($r = -0.278$, and $r = -0.328$, respectively), than with *Satisfaction* and $\log(\text{Satisfaction})$ (here the largest correlation is $r = 0.111$ between satisfaction and $\text{Logit}(\text{SOW})$). Among the 14 models considered, the regression models based on *Satisfaction* provide estimates of change in *SOW* and $\text{logit}(\text{SOW})$ that have the smallest correlations with actual change in *SOW* and $\text{Logit}(\text{SOW})$.

Using a longitudinal data set, the findings therefore clearly demonstrate that when linking *changes* in customers’ satisfaction levels to *changes* in corresponding share of wallet over time, compared to absolute satisfaction, relative ranked satisfaction remains the more closely linked measure to share of wallet. *H2* is therefore supported.

Discussion and conclusion

The analysis reported here advances the empirical research regarding the relationship between customer satisfaction and share of wallet in two overarching ways. First, our findings clearly demonstrate that relative ranked satisfaction is superior to absolute satisfaction in linking to the share of category spending that customers allocate to the brands that they use.

Specifically, our research finds that absolute satisfaction explains only 5.6 percent of the variation in share of wallet when examined cross-sectionally, and changes in absolute satisfaction explain only a very small 0.4 percent of the variation in contemporaneous changes in share of wallet. By contrast, relative ranked satisfaction

explains 23.4 percent in the variation in share of wallet, and changes in relative ranked satisfaction explains 8.1 percent of the variation in changes in share of wallet. Furthermore, almost without exception, models based on absolute satisfaction are the worst performing models examined in our investigation.

Second, our findings indicate that there are multiple methodologies available to researchers and managers to transform ranked satisfaction into relatively good approximations of customers' share of wallet allocations.

Specifically, we find that all of the most commonly used discrete distributions (i.e. Zipf-AE, Zipf-PM, and WAR) perform remarkably well. For example, the percentage of variance explained from changes in the share of wallet estimates from these models and changes in customers' share of category spending ranged from a high of 16.6 percent (for the WAR-fixed parameter model) to a low of 15.4 percent (for the Zip-PM fixed parameter model), with the rest of the models explaining 16 percent or more of the variance.

Additionally, when examined cross-sectionally, the percentage of variance explained by these models ranges from a low of 34.6 percent (for the Zip-PM fixed parameter model) to a high of 37.6 percent (for the Zipf-AE nine-parameter model).

Similar cross-sectional results were obtained for hierarchical regression models based on rank (36.0 and 37.0 percent for the two models examined). Longitudinally, however, these models explained approximately 13 percent of the variation in changes in share of wallet.

Taken together, these findings have wide reaching implications for both the practice and the science of marketing.

Implications for researchers

These results also have several important implications for scientific researchers, and point to the need for new research in several areas. The most obvious implication of this research is that the traditional view of the satisfaction and share of wallet relationship (i.e. a non-linear, s-shaped relationship) based upon absolute satisfaction levels is at best incomplete. Our findings indicate that the relationship is instead primarily driven by the relative fulfillment customers perceive from the various brands that they use (as gauged by their relative ranked satisfaction level), and not the absolute level of satisfaction. Therefore, while consumer satisfaction represents a widely studied area of research (for a review, see Oliver, 2010), our findings indicate a need for additional research into the nature of satisfaction and its corresponding impact on consumer behavior which better takes competitive effects into account.

Choice modelers have known for years that you need to consider all brands in the usage set (Luce, 1959, 1977), yet this simple fact has not been applied by most satisfaction researchers. Satisfaction researchers must recognize that consumers are making a choice, and that the choice is relative.

The relative nature of satisfaction also indicates that we need new, more comprehensive models linking satisfaction to business results. As noted earlier, the seminal satisfaction-based chain of effects models in the literature focus on absolute, focal-firm only metrics.

Additionally, given the relative nature of consumer satisfaction, this raises the likelihood that other perceptual and attitudinal metrics display similar properties. For example, since most researchers presume that satisfaction is an antecedent to commitment (e.g. Bansal *et al.*, 2004; Garbarino and Johnson, 1999; Hennig-Thurau *et al.*, 2002), this begs the question, "Is commitment also relative?" If yes, how do consumers

trade off different types of commitment (e.g. affective, calculative, and normative) with the various brands that they use in a category?

Furthermore, previous satisfaction literature has devoted attention to the moderating impact of customer and situational characteristics on the relationship between satisfaction and share of wallet (e.g. Cooil *et al.*, 2007). Hence, given our new insights, additional research is warranted, investigating these moderating influences in a relative context. For example, length of relationship could be of particular importance to this research context, as this has been found to lower the relationship between absolute satisfaction and loyalty (e.g. Homburg *et al.*, 2003).

Finally, this research relied on using ranks to capture relative satisfaction. While ranks have been used in other marketing applications to capture relative performance (e.g. Kohli and Sah, 2006; Shugan and Mitra, 2013), and Shugan and Mitra (2013) offer a compelling argument regarding the benefits of using ranks as a unit of analysis, more research is needed to determine the best means of capturing relative satisfaction (and other perceptual metrics). To date, there are several approaches proposed. For example, rank transformation (e.g. Hofmeyr *et al.*, 2008; Keiningham *et al.*, 2011) and mean-centering (Wind, 1970) are two common approaches for deriving relative position. Van den Putte *et al.* (1996) use direct ranking scales (i.e. respondents assign a rank). Still other researchers have proposed relative scales (e.g. Hauser, 1991). Therefore, there is a need to examine different relative measurement approaches to determine which methods work best and under what conditions.

Implications for managers

One of the most important implications is that firms need to shift from focusing on their satisfaction score (i.e. rating level) to focusing on their rank to which the satisfaction level corresponds. This need not be complicated, particularly since ranks are used in multiple aspects of our lives (sports, education, etc.). If the firm already has a customer satisfaction tracking program in place, managers can simply add questions about competitors used and ask respondents to provide satisfaction ratings for these competitors in addition to the focal firm. For firms that do not have a tracking system in place, managers can institute one with new questionnaires that measure satisfaction perceptions for the firm and its competitors which could then be transformed into ranks. The information collected would provide valuable input for calculating metrics to be tracked and/or included in dashboards and also provide opportunities to benchmark over time.

For example, Keiningham *et al.* (2014) argue that managers should focus on the percentage of their customers who would be classified as ranking the firm first among all the competitors that they used; they refer to this metric as the percentage “First Choice.” There is an obvious appeal to managers for such a metric. Regardless of the level of the employee within the organization, all have a visceral sense of the importance of being first-choice *vis-à-vis* competition. While there are limitations with a focus on being “first,” it does offer managers a measure that is easy to communicate and easy to rally support around that keeps the focus of the organization on relative rank.

Another important finding for managers is that managers have several viable options when deciding on how they wish to link satisfaction to SOW. The discrete distributions examined perform remarkably well. Nearly all of the discrete distributions perform at the same level when predicting change in SOW.

It is important to note, however, that with the exception of “automatic decision models such as those involved in search engine optimization, revenue management systems and so forth,” simple models tend to perform better when users are involved (Lilien, 2013). Little (1970, 2004) observes that for models to be both useful and used in practice they must be “(1) simple, (2) robust, (3) easy to control, (4) adaptive, (5) complete on important issues, (6) easy to communicate with” (2004, p. 1855). While marketing academics likely view most (if not all) of the models investigated – particularly the most commonly used discrete models – as being relatively straightforward, the reality is that most managers do not. In fact, when explaining the WAR – the simplest model investigated – the *Harvard Business Review* first implored managers with “Don’t let the math scare you” (Keiningham *et al.*, 2011, p. 30).

The danger is that managers tend to reject models that they don’t understand and “revert to models of great simplicity” (Little, 2004, p. 1855). For example, the simplicity of the calculation and the ease of communicating the underlying philosophy would appear to explain in large part the continued popularity of the Net Promoter metric (Owen and Brooks, 2009, p. 10) despite a wide body of scientific evidence (including this investigation) which casts doubt on its reported claims to link to business outcomes (e.g. Keiningham *et al.*, 2007; Morgan and Rego, 2006; Sharp, 2008).

Therefore, managers need to balance precision with the ability to easily understand and communicate the fundamentals of the model selected. In the case of this examination, several of the models tested require no data fitting to arrive at share of wallet estimates, specifically the WAR (both fixed parameter and nine-parameter versions) and all of the fixed parameter discrete distributions). As a result, managers have relatively simple models to use which can significantly increase the strength of the relationship between satisfaction and share of wallet.

These findings have another important implication for managers. Because rank-based models are substantially superior to absolute satisfaction based models in linking to SOW, the drivers of satisfaction and the drivers of share of wallet are likely to be different. By “drivers” we mean the underlying attributes that influence overall satisfaction levels (Morgan *et al.*, 2005; Anderson and Mittal, 2000).

Most managers identify drivers of satisfaction based upon consumer ratings regarding the performance of their firm only. Relative ranked satisfaction models, however, by their nature take competition into account. Early research into the differences between drivers of satisfaction and drivers of rank (based on relative satisfaction levels) indicates that consumers who use more than one brand in a category at the same time do so to fulfill different needs (Aksoy, 2013b). Therefore, improving rank would imply not only increasing satisfaction with a firm’s offering, but also reducing consumers’ perceived needs to use competitors.

Conclusion

There is general agreement among researchers and practitioners that satisfaction is relative to competitive alternatives (e.g. Birtchnell, 1994; Holt and Huber, 1969; Varki and Rust, 1997; Semon, 1994). Nonetheless, researchers and managers have not treated satisfaction as a relative construct. The result has been weak relationships between satisfaction and SOW in the literature, and challenges by managers as to whether satisfaction is a useful predictor of customer behavior and business growth (Chemi, 2013; Gupta and Zeithaml, 2006; Reichheld, 2003).

This research similarly challenges the usefulness of using absolute satisfaction levels, and absolute levels of other commonly used metrics such recommend intention

and the NPS, in linking to customers' share of category spending. Our findings indicate that these commonly used metrics explain less than one half of one percent of the variance in share of wallet. While this may be statistically significant, it is almost certainly not managerially relevant.

This investigation provides compelling evidence of the superiority of relative ranked satisfaction to absolute satisfaction in linking to share of wallet. Moreover, it provides practical insight into several easy-to-use approaches that researchers and managers can apply to improve the strength of the relationship between satisfaction and share of wallet. For example, our research found that almost all versions of the three most commonly used power laws explained 35 percent or more of the variance in share of wallet when examined cross-sectionally, and 16 percent or more of the variance in changes in share of wallet when examined longitudinally.

Finally, this research points to the critical need for new research into the relative nature of satisfaction, as well as other perceptual and attitudinal constructs, to better understand their influence on consumer behavior.

Limitations

Although this investigation used a large data set comprised of multiple brands, industries, and countries, there are limitations that should be noted. Inclusion of additional brands, industries and countries would more clearly establish the generalizability of our findings.

Additionally, our investigation analyzed only multi-brand usage markets and customers. Therefore, research needs to be conducted in single brand usage categories to better understand the relationship between satisfaction and consumer behavior to determine if and how relative satisfaction levels impact this relationship.

Finally, our analysis identified the presence of a statistically significant relationship between current share of wallet levels and relative ranked satisfaction, and changes in share of wallet and concomitant changes in relative ranked satisfaction levels. We did not, however, prove causation. Therefore, additional longitudinal research should be conducted to examine the robustness of these findings.

Nonetheless, we believe these results provide compelling evidence of the superiority of relative satisfaction metrics in linking to customers' share of wallet allocations. Moreover, this investigation provides insight into several viable approaches that researchers and managers can apply to more strongly link satisfaction to customers' spending behaviors.

Notes

1. At its core, Zipf's Law states that the frequency an event is inversely proportional to its rank. Many types of data studied in the physical and social sciences have been shown to be inversely proportional to rank. Of importance to this investigation, Zipf's Law has been shown to apply to market share (Kohli and Sah 2006), corporation sizes (Ramsden and Kiss-Haypál, 2000), and the income distribution of companies (Okuyama *et al.*, 1999).
2. Note, Hofmeyr *et al.* (2008) modeled separate Zipf Distributions for each brand usage size.
3. Although the logit transformation is the standard link function used when general linear models are applied to binomial data, it is used here, and in the models introduced later, to provide as an unbounded and relatively familiar dependent variable for linear regression.
4. A detailed description of all models investigated and all analytics conducted is provided in the Technical Appendix.

5. Table AI in the Appendix 1 provides a detailed comparison of model performance overall and at the customer level in terms of mean absolute deviation (MAD) and root mean squared error (RMSE) as percent of total SOW.
6. Table AII in the Appendix 1 provides a detailed comparison of model performance by the number of total brands (m) that are used by the customer.
7. Table AIII in the Appendix 1 provides a detailed summary of the correlations of change in SOW and logit(SOW) with contemporaneous changes in model estimates and changes in other variables.
8. The nine-class models are fit separately to nine groups defined by the number of *Total Brands*, and they include one additional hierarchical parameter per group).
9. In the regression models that use satisfaction, total brands does not need to be used as a predictor, because a separate intercept is fit within each total brand category.

References

- Ahearne, M., Jelinek, R. and Jones, E. (2007), "Examining the effect of salesperson service behavior in a competitive context", *Journal of the Academy of Marketing Science*, Vol. 35 No. 4, pp. 603-616.
- Ajzen, I. (2001), "Nature and operation of attitudes", *Annual Review of Psychology*, Vol. 52 No. 1, pp. 27-58.
- Ajzen, I. and Fishbein, M. (1980), *Understanding Attitudes and Predicting Social Behavior*, Prentice Hall, Englewood Cliffs, NJ.
- Ajzen, I. and Madden, T.J. (1986), "Prediction of goal-directed behavior: attitudes, intentions, and perceived behavioral control", *Journal of Experimental Social Psychology*, Vol. 22 No. 5, pp. 453-474.
- Aksoy, L. (2013a), "How do you measure what you can't define? The current state of loyalty measurement and management", *Journal of Service Management*, Vol. 24 No. 4, pp. 356-381.
- Aksoy, L. (2013b), "Linking member satisfaction to share of deposits: applying the Wallet Allocation Rule in credit unions", research report, Filene Research Institute Publication Number No. 290, Filene Research Institute, Madison, WI, February.
- Anderson, E.W. and Mittal, V. (2000), "Strengthening the satisfaction-profit chain", *Journal of Service Research*, Vol. 3 No. 2, pp. 107-120.
- Anderson, E.W., Fornell, C. and Mazvancheryl, S.K. (2004), "Customer satisfaction and shareholder value", *Journal of Marketing*, Vol. 68 No. 4, pp. 172-185.
- Bansal, H.S., Irving, P.G. and Taylor, S.F. (2004), "A three-component model of customer commitment to service providers", *Journal of the Academy of Marketing Science*, Vol. 32 No. 3, pp. 234-250.
- Baumann, C., Burton, S. and Elliot, G. (2005), "Determinants of customer loyalty and share of wallet in retail banking", *Journal of Financial Services Marketing*, Vol. 9 No. 3, pp. 231-248.
- Bergkvist, L. and Rossiter, J.R. (2007), "The predictive validity of multiple-item versus single-item measures of same constructs", *Journal of Marketing Research*, Vol. 44 No. 2, pp. 175-184.
- Birtchnell, J. (1994), "The interpersonal octagon: an alternative to the interpersonal circle", *Human Relations*, Vol. 47 No. 5, pp. 511-529.
- Bolton, R.N. (1998), "A dynamic model of the duration of the customer's relationship with a continuous service provider: the role of satisfaction", *Marketing Science*, Vol. 17 No. 1, pp. 45-65.

- Bolton, R.N., Kannan, P.K. and Bramlett, M.D. (2000), "Implications of loyalty program membership and service experiences for customer retention and value", *Journal of the Academy of Marketing Science*, Vol. 28 No. 1, pp. 95-108.
- Bolton, R.N. and Lemon, K.N. (1999), "A dynamic model of customers' usage of services: usage as an antecedent and consequence of satisfaction", *Journal of Marketing Research*, Vol. 36 No. 2, pp. 171-186.
- Bolton, R.N., Lemon, K.N. and Verhoef, P.C. (2004), "The theoretical underpinnings of customer asset management: a framework and propositions for future research", *Journal of the Academy of Marketing Science*, Vol. 32 No. 3, pp. 271-292.
- Boulding, W., Kalra, A., Staelin, R. and Zeithaml, V.A. (1993), "A dynamic process model of service quality", *Journal of Marketing Research*, Vol. 30 No. 1, pp. 7-27.
- Bowman, D. and Narayandas, D. (2004), "Linking customer management effort to customer profitability in business markets", *Journal of Marketing Research*, Vol. 41 No. 4, pp. 433-447.
- Cadotte, E.R., Woodruff, R.B. and Jenkins, R.L. (1987), "Expectations and norms in models of consumer satisfaction", *Journal of Marketing Research*, Vol. 24 No. 3, pp. 305-314.
- Chemi, E. (2013), "Proof that it pays to be America's most-hated companies," *Bloomberg Businessweek*, December 17, available at: www.businessweek.com/articles/2013-12-17/proof-that-it-pays-to-be-americas-most-hated-companies (accessed April 18, 2014).
- Cooil, B., Keiningham, T.L., Aksoy, L. and Hsu, M. (2007), "A longitudinal analysis of customer satisfaction and share of wallet: investigating the moderating effect of customer characteristics", *Journal of Marketing*, Vol. 71 No. 1, pp. 67-83.
- Crosby, L.A. and Stephens, N. (1987), "Effects of relationship marketing on satisfaction, retention, and prices in the life insurance industry", *Journal of Marketing Research*, Vol. 24 No. 4, pp. 404-411.
- Crotts, J.C., Pan, B. and Raschid, A.E. (2008), "A survey method for identifying key drivers of guest delight", *International Journal of Contemporary Hospitality Management*, Vol. 20 No. 4, pp. 462-470.
- Dick, A.S. and Basu, K. (1994), "Customer loyalty: toward an integrated conceptual framework", *Journal of the Academy of Marketing Science*, Vol. 22 No. 2, pp. 99-113.
- Drolet, A.L. and Morrison, D.G. (2001), "Do we really need multiple-item measures in service research?", *Journal of Service Research*, Vol. 3 No. 3, pp. 196-204.
- Fornell, C., Johnson, M.D., Anderson, E.W., Cha, J. and Bryant, B.E. (1996), "The American customer satisfaction index: nature, purpose, and findings", *Journal of Marketing*, Vol. 60 No. 4, pp. 7-18.
- Gale, B.T. (1994), *Managing Customer Value*, The Free Press, New York, NY.
- Garbarino, E. and Johnson, M.S. (1999), "The different roles of satisfaction, trust, and commitment in customer relationships", *Journal of Marketing*, Vol. 63 No. 2, pp. 70-87.
- Gardial, S.F., Clemons, D.S., Woodruff, R.B., Schumann, D.W. and Burns, M.J. (1994), "Comparing consumers recall of prepurchase and postpurchase product evaluation experiences", *Journal of Consumer Research*, Vol. 20 No. 4, pp. 548-560.
- Gasbarra, D., Kulathinal, D., Pirinen, M. and Sillanpää, M.J. (2011), "Estimating haplotype frequencies by combining data from large DNA pools with database information", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 8 No. 1, pp. 36-44.
- Gesulado, N. (2011), "Team including Prof. Aksoy wins innovation award", Gabelli Connect, November 9, available at: www.gabelliconnect.com/disruptiveinnovation (accessed December 18, 2012).

- Gitomer, J. (1998), *Customer Satisfaction is Worthless, Customer Loyalty Is Priceless*, Bard Press, Austin, TX.
- Gupta, S. and Zeithaml, V. (2006), "Customer metrics and their impact on financial performance", *Marketing Science*, Vol. 25 No. 6, pp. 718-739.
- Hardie, B., Johnson, E.J. and Fader, P.S. (1993), "Modeling loss aversion and reference dependence effects on brand choice", *Marketing Science*, Vol. 12 No. 4, pp. 378-394.
- Hauser, J.R. (1991), *Comparison of Importance Measurement Methodologies and their Relationship to Consumer Satisfaction*, MIT Sloan School of Management, Cambridge, MA.
- Hennig-Thurau, T., Gwinner, K.P. and Gremler, D.D. (2002), "Understanding relationship marketing outcomes: an integration of relational benefits and relationship quality", *Journal of Service Research*, Vol. 4 No. 3, pp. 230-247.
- Heskett, J.L., Jones, T.O., Loveman, G.W., Sasser, W.E. Jr and Schlesinger, L.A. (1994), "Putting the service-profit chain to work", *Harvard Business Review*, Vol. 72 No. 2, pp. 164-174.
- Hofmeyr, J. (2012), "Bringing marketing research into the 21st Century", TNS Conversion Model Seminar, TNS Sifo, September 11, available at: www.slideshare.net/tnssifo/tns-conversionmodel-seminar-911-hur-du-vinner-slaget (accessed December 6, 2013).
- Hofmeyr, J. and Parton, G. (2010), "Rank matters", *Marketing Research*, Vol. 22 No. 3, pp. 6-12.
- Hofmeyr, J., Goodall, V., Bongers, M. and Holtzman, P. (2008), "A new measure of brand attitudinal equity based on the Zipf distribution", *International Journal of Market Research*, Vol. 50 No. 2, pp. 181-202.
- Holt, C.C. and Huber, G.P. (1969), "A computer aided approach to employment service placement and counseling", *Management Science*, Vol. 15 No. 11, pp. 573-594.
- Homburg, C., Giering, A. and Menon, A. (2003), "Relationship characteristics as moderators of the relationship between customer satisfaction and loyalty: an empirical analysis", *Psychology and Marketing*, Vol. 18 No. 1, pp. 35-62.
- Humphrey, J. (2008), "Synovate wins MRS David Winton Award for Innovation", press release, December 11, available at: www.pressreleasepoint.com/synovate-wins-mrs-david-winton-award-innovation (accessed December 18, 2012).
- Inside Research (2012), "Special report: US CSM spend back", Vol. 23, Issue 289, Number 3 (March), p. 16.
- Ipsos (2012), "Fast facts", available at: www.ipsos.com/Fast_facts (accessed December 27, 2012).
- Jacoby, J. and Chesnut, R.W. (1978), *Brand Loyalty Measurement and Management*, John Wiley, New York, NY.
- Johnson, N.L. (1960), "An approximation to the multinomial distribution: some properties and applications", *Biometrika*, Vol. 47 Nos 1/2, pp. 93-102.
- Keiningham, T.L. (2012), "Customer loyalty measurement is broken – let's fix it!", presented at the 2012 Conference Board Customer Experience Leadership Conference, New York, NY, March 22-23.
- Keiningham, T.L. and Vavra, T.G. (2001), *The Customer Delight Principle*, McGraw Hill, New York, NY.
- Keiningham, T.L., Perkins-Munn and Evans, H. (2003), "The impact of customer satisfaction on share-of-wallet in a business-to-business environment", *Journal of Service Research*, Vol. 6 No. 1, pp. 37-50.
- Keiningham, T.L., Perkins-Munn, T., Aksoy, L. and Estrin, D. (2005), "Does customer satisfaction lead to profitability?: The mediating role of share-of-wallet", *Managing Service Quality*, Vol. 15 No. 2, pp. 172-181.

-
- Keiningham, T.L., Cooil, B., Andreassen, T.W. and Aksoy, L. (2007), "A longitudinal examination of net promoter and firm revenue growth", *Journal of Marketing*, Vol. 71 No. 3, pp. 39-51.
- Keiningham, T.L., Aksoy, L., Cooil, B. and Andreassen, T.W. (2008), "Linking customer loyalty to growth", *MIT Sloan Management Review*, Vol. 49 No. 4, pp. 50-57.
- Keiningham, T.L., Aksoy, L., Buoye, A. and Cooil, B. (2011), "Customer loyalty isn't enough. Grow your share of wallet", *Harvard Business Review*, Vol. 89 No. 10, pp. 29-31.
- Keiningham, T.L., Gupta, S., Aksoy, L. and Buoye, A. (2014), "The high price of customer satisfaction", *MIT Sloan Management Review*, Vol. 55 No. 3, pp. 37-46.
- Keiningham, T.L., Aksoy, L., De Keyser, A., Larivière, B., Buoye, A. and Williams, L. (2014), "It's not your score that matters: the importance of relative metrics", in Roland, R. and Huang, M.-H. (Eds), *Handbook of Service Marketing Research*, Edward Elgar Publishing, Cheltenham, pp. 166-181.
- Kohli, R. and Sah, R. (2006), "Some empirical regularities in market shares", *Management Science*, Vol. 52 No. 11, pp. 1792-1798.
- Kumar, P. (2002), "The impact of performance, cost, and competitive considerations on the relationship between satisfaction and repurchase intent in business markets", *Journal of Service Research*, Vol. 5 No. 1, pp. 55-68.
- Larivière, B. (2008), "Linking perceptual and behavioral customer metrics to multiperiod customer profitability – a comprehensive service-profit chain application", *Journal of Service Research*, Vol. 11 No. 1, pp. 3-21.
- Lilien, G. (2013), "Personal communication", e-mail, July 17.
- Little, J.D.C. (1970), "Models and managers: the concept of a decision calculus", *Management Science*, Vol. 16 No. 8, pp. B466-B485.
- Little, J.D.C. (2004), "Comments on 'models and managers: the concept of a decision calculus': managerial models for practice", *Management Science*, Vol. 50 No. Suppl 12, pp. 1854-1860.
- Louw, A. and Hofmeyr, J. (2012), "Reality check in the digital age: the relationship between what we ask and what people actually do", in Fellows, D.S. (Ed.), *3D Digital Dimensions 2012 (ESOMAR Publication Series Volume S355 (ISBN 920831-0261-4))*, ESOMAR, Amsterdam.
- Luce, R.D. (1959), *Individual Choice Behavior: A Theoretical Analysis*, John Wiley & Sons, New York, NY.
- Luce, R.D. (1977), "The choice axiom after twenty years", *Journal of Mathematical Psychology*, Vol. 15 No. 3, pp. 215-233.
- Luo, X. and Bhattacharya, C.B. (2006), "Corporate social responsibility, customer satisfaction, and market value", *Journal of Marketing*, Vol. 70 No. 4, pp. 1-18.
- Luo, X. and Homburg, C. (2007), "Neglected outcomes of customer satisfaction", *Journal of Marketing*, Vol. 71 No. 2, pp. 133-149.
- Mägi, A.W. (2003), "Share of wallet in retailing: the effects of customer satisfaction, loyalty cards and shopper characteristics", *Journal of Retailing*, Vol. 79 No. 2, pp. 97-106.
- Miller, J.A. (1977), "Studying satisfaction, modifying models, eliciting expectation, posing problems and making meaningful measurement", in Hunt, H.K. (Ed.), *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, Marketing Science Institute, Cambridge, MA, pp. 72-91.
- Mittal, V. and Frennea, C.M. (2010), "Customer satisfaction: a strategic review and guidelines for managers", Fast Forward MSI Report No. 10-701, Cambridge, MA.

- Mittal, V. and Kamakura, W.A. (2001), "Satisfaction, repurchase intent, and repurchase behavior: investigating the moderating effect of customer characteristics", *Journal of Marketing Research*, Vol. 38 No. 1, pp. 131-142.
- Mittal, V., Kumar, P. and Tsiros, M. (1999), "Attribute-level performance, satisfaction, and behavioral intentions over time: a consumption system approach", *Journal of Marketing*, Vol. 63 No. 2, pp. 88-101.
- Mittal, V., Ross, W.T. Jr and Baldasare, P.M. (1998), "The asymmetric impact of negative and positive attribute-level performance on overall satisfaction and repurchase intentions", *Journal of Marketing*, Vol. 62, January, pp. 33-47.
- Morgan, N.A., Anderson, E.W. and Mittal, V. (2005), "Understanding firms' customer satisfaction information usage", *Journal of Marketing*, Vol. 69 No. 3, pp. 131-151.
- Morgan, N.A. and Rego, L.L. (2006), "The value of different customer satisfaction and loyalty metrics in predicting business performance", *Marketing Science*, Vol. 25 No. 5, pp. 426-439.
- Okuyamaa, K., Takayasu, M. and Takayasu, H. (1999), "Zipf's law in income distribution of companies", *Physica A: Statistical Mechanics and its Applications*, Vol. 269 No. 1, pp. 125-131.
- Oliver, R.L. (1980), "A cognitive model of the antecedents and consequences of satisfaction decisions", *Journal of Marketing Research*, Vol. 17 No. 4, pp. 460-469.
- Oliver, R.L. (1997), *Satisfaction: A Behavioral Perspective on the Consumer*, McGraw Hill, New York, NY.
- Oliver, R.L. (2010), *Satisfaction: A Behavioral Perspective on the Consumer*, 2nd ed., M.E. Sharpe, Armonk, NY.
- Oliver, R.L., Rust, R.T.R. and Varki, S. (1997), "Customer delight: foundations, findings, and managerial insight", *Journal of Retailing*, Vol. 73 No. 3, pp. 311-336.
- Olsen, S.O. (2002), "Comparative evaluation and the relationship between quality, satisfaction, and repurchase loyalty", *Journal of the Academy of Marketing Science*, Vol. 30 No. 3, pp. 240-249.
- Owen, R. and Brooks, L. (2009), *Answering the Ultimate Question*, Jossey-Bass, San Francisco, CA.
- Parasuraman, A., Zeithaml, V.A. and Berry, L.L. (1988), "SERVQUAL: s multiple-item scale for measuring consumer perceptions of service quality", *Journal of Retailing*, Vol. 64 No. 1, pp. 12-40.
- Perkins-Munn, T., Aksoy, L., Keiningham, T.L. and Estrin, D. (2005), "Actual purchase as a proxy for share of wallet", *Journal of Service Research*, Vol. 7 No. 3, pp. 245-256.
- Ramsden, J.J. and Kiss-Haypál, G. (2000), "Company size distribution in different countries", *Physica A: Statistical Mechanics and its Applications*, Vol. 277 Nos 1-2, pp. 220-227.
- Reichheld, F.F. (2003), "The one number you need to grow", *Harvard Business Review*, Vol. 81 No. 12, pp. 46-54.
- Rust, R.T., Danaher, P.J. and Varki, S. (2000), "Using service quality data for competitive marketing decisions", *International Journal of Service Industry Management*, Vol. 11 No. 5, pp. 438-469.
- Rust, R.T. and Huang, M. (2014), "Service marketing research: emerging directions", in Roland, R. and Huang, M.-H. (Eds), *Handbook of Service Marketing Research*, Edward Elgar Publishing, Cheltenham, pp. 3-8.
- Rust, R.T., Lemon, K.N. and Zeithaml, V.A. (2004), "Return on marketing: using customer equity to focus marketing strategy", *Journal of Marketing*, Vol. 68 No. 1, pp. 109-127.
- Rust, R.T. and Zahorik, A.J. (1993), "Customer satisfaction, customer retention, and market share", *Journal of Retailing*, Vol. 69 No. 2, pp. 193-215.

-
- Rust, R.T., Zahorik, A.J. and Keiningham, T.L. (1995), "Return on quality (ROQ), making service quality financially accountable", *Journal of Marketing*, Vol. 59 No. 2, pp. 58-70.
- Rust, R.T., Zahorik, A.J. and Keiningham, T.L. (1996), *Service Marketing*. Harper Collins, HarperCollins, New York, NY.
- Semon, T.T. (1994), "Customer satisfaction is relative", *Marketing News*, Vol. 28 No. 10, pp. 15.
- Sharp, B. (2008), "Net promoter score fails the test", *Marketing Research*, Vol. 20 No. 4, pp. 28-30.
- Shugan, S.M. and Mitra, D. (2013), "A theory for market growth or decline", *Marketing Science*, Vol. 33 No. 1, pp. 47-65.
- Silvestro, R. and Cross, S. (2000), "Applying the service profit chain in a retail environment – challenging the 'satisfaction mirror'", *International Journal of Service Industry Management*, Vol. 11 No. 3, pp. 244-268.
- TNS (2012), "Conversion Model™ – commitment measurement", available at: www.tns-global.be/solutions/commitment-research/ (accessed December 5, 2012).
- Van den Putte, B., Hoogstraten, J. and Meertens, R. (1996), "A comparison of behavioural alternative models in the context of the theory of reasoned action", *British Journal of Social Psychology*, Vol. 35 No. 2, pp. 257-266.
- Varki, S. and Rust, R.T. (1997), "Satisfaction is relative", *Marketing Research*, Vol. 9 No. 2, pp. 14-19.
- Williams, R. and Visser, R. (2002), "Customer satisfaction: it is dead but it will not lie down", *Managing Service Quality*, Vol. 12 No. 3, pp. 194-200.
- Wind, Y. (1970), "Industrial source loyalty", *Journal of Marketing Research*, Vol. 7 No. 4, pp. 450-457.
- Woodruff, R.B., Cadotte, E.R. and Jenkins, R.L. (1983), "Modeling consumer satisfaction processes using experience-based norms", *Journal of Marketing Research*, Vol. 20 No. 3, pp. 296-304.
- Yi, Y. (1990), "A critical review of consumer satisfaction", in Zeithaml, V.A. (Ed.), *Review of Marketing*, American Marketing Association, Chicago, IL, pp. 68-123.
- Zeithaml, V.A., Berry, L.L. and Parasuraman, A. (1996), "The behavioral consequences of service quality", *Journal of Marketing*, Vol. 60 No. 2, pp. 31-46.
- Zeithaml, V.A., Bolton, R., Deighton, J., Keiningham, T.L., Lemon, K. and Petersen, A. (2006), "Forward-looking customer focus: can firms have adaptive foresight?", *Journal of Service Research*, Vol. 9 No. 2, pp. 168-183.
- Zipf, G.K. (1935), *The Psycho-Biology of Language: an Introduction to Dynamic Philology*, Houghton Mifflin, Boston, MA.

Further reading

- Anderson, E.W. and Sullivan, M.W. (1993), "The antecedents and consequences of customer satisfaction for firms", *Marketing Science*, Vol. 12 No. 2, pp. 125-143.

(The Appendix follows overleaf.)

Table AI.
Model performance overall and at the customer level in terms of mean absolute deviation (MAD) and root mean squared error (RMSE) as percent of total SOW

	R^2 -ADJ (%)	MAD	Rank (% greater than best)	RMSE	Rank (% greater than best)	MAD	Rank (% greater than best)	RMSE	Rank (% greater than best)
<i>Fixed parameter models</i>									
Zipf-AE, $s = 1$	37.0	15.2	2 (1.6)	20.4	4 (0.8)	16.9	2 (0.9)	18.2	1 (-)
Zipf-PM, $s = 1$	34.6	15.7	12 (4.8)	21.5	12 (6.0)	17.4	10 (3.7)	18.8	11 (3.3)
WAR, $p = 2/[(Total Brands)+1]$	36.3	15.4	6 (2.9)	20.5	6 (1.2)	17.1	5 (1.8)	18.4	5 (0.8)
<i>One-parameter models</i>									
Zipf-AE	37.5	15.3	4 (2.0)	20.3	2 (0.1)	17.0	4 (1.6)	18.3	3 (0.3)
Zipf-PM	36.0	15.4	8 (2.9)	20.5	8 (1.4)	17.2	7 (2.4)	18.4	6 (1.0)
Truncated Geometric	35.9	15.7	11 (4.4)	20.5	9 (1.4)	17.4	12 (3.9)	18.7	10 (2.4)
<i>Nine-parameter models</i>									
(1 Parameter per Total Brands Class)									
Zipf-AE	37.6	15.3	3 (1.7)	20.3	1 (-)	17.0	3 (1.3)	18.2	2 (0.0)
Zipf-PM	36.0	15.4	7 (2.9)	20.5	7 (1.3)	17.2	9 (2.4)	18.4	7 (1.1)
WAR	36.5	15.5	9 (3.2)	20.4	5 (0.9)	17.2	8 (2.4)	18.4	8 (1.1)
Truncated Geometric	37.2	15.4	5 (2.4)	20.3	3 (0.4)	17.1	6 (1.8)	18.3	4 (0.5)
<i>Hierarchical regressions</i>									
(customer i , product category j)									
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \log(Total Brands + 1 - Rank)$ 1 Class									
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \cdot Satisfaction + \beta_2 (Total Brands)$									
1 Class									
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \log(Total Brands + 1 - Rank)$									
9 Classes									
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \cdot Satisfaction$, 9 Classes									
9 Classes									
29.1 16.1 13 (7.3) 22.0 13 (8.4) 18.1 13 (8.1) 20.5 13 (12.6)									
Notes: There are 258,743 observations and 79,543 sets of customer rankings. The regression models use an additional hierarchical parameter for the additional customer level standard error. The nine-class models are fit separately to nine groups defined by the number of <i>Total Brands</i> , and they include one additional hierarchical parameter per group									

Total brands Group size (% of total brands > 1) Model	2		3		4		5		6		7-10		Best rank	Worst rank	Median rank	
	Parameters	MAD	Rank	MAD	Rank	MAD	Rank	MAD	Rank	MAD	Rank	MAD				Rank
Fixed parameter models																
Zipf-AE, $s = 1$	Fixed	20.5	1.5	17.1	3	14.0	3	11.9	5	10.3	5	8.1	10	1.5	10	4
Zipf-PM, $s = 1$	Fixed	21.0	10	17.6	12	14.6	12	12.4	13	10.8	14	8.6	14	10	14	12.5
WAR, $p = 2/[(Total Brands)+1]$	Fixed	20.5	1.5	17.3	7	14.4	10	12.3	12	10.7	13	8.2	11	1.5	13	10.5
One-parameter models																
Zipf-AE	1	20.8	7	17.1	5	14.1	5	11.9	3	10.3	3	7.9	4	3	7	4.5
Zipf-PM	1	20.9	8	17.3	9	14.3	9	12.1	8	10.5	11	7.9	7	7	11	8.5
Truncated geometric	1	21.3	11	17.6	11	14.3	7	12.0	7	10.4	9	8.4	13	7	13	10
Nine-parameter models																
Zipf-AE	9	20.8	5	17.1	4	14.1	4	11.9	4	10.3	4	7.8	3	3	5	4
Zipf-PM	9	20.9	9	17.3	8	14.3	8	12.0	6	10.4	7	7.9	6	6	9	7.5
WAR	9	20.8	5	17.4	10	14.4	11	12.2	11	10.6	12	7.9	8	5	12	10.5
Truncated Geometric	9	20.8	5	17.2	6	14.2	6	12.1	9	10.4	8	7.9	5	5	9	6
Hierarchical regression																
Logit(SOW _{<i>i</i>}) = $\beta_0 + \beta_1 \log(Rank_i) + \beta_2 \log(Total Brands) + 1 - Rank_i$																
1 Class	4	21.9	12	17.0	2	13.9	2	11.7	2	10.1	2	8.2	12	2	12	2
Logit(SOW _{<i>i</i>}) = $\beta_0 + \beta_1 Satisfaction_i + \beta_2 (Total Brands)$																
1 Class	4	23.5	14	18.4	14	15.1	14	12.4	14	10.4	10	8.0	9	9	14	14
Logit(SOW _{<i>i</i>}) = $\beta_0 + \beta_1 \log(Rank_i) + \beta_2 \log(Total Brands) + 1 - Rank_i$																
9 Classes	36	20.6	3	16.9	1	13.8	1	11.5	1	9.8	1	7.4	1	1	3	1
Logit(SOW _{<i>i</i>}) = $\beta_0 + \beta_1 Satisfaction_i$																
9 Classes	27	22.5	13	18.2	13	14.6	13	12.1	10	10.4	6	7.5	2	2	13	11.5

Notes: There are 258,743 observations and 79,543 customers. The regression models use an additional hierarchical parameter for the additional customer level standard error. The nine-class models are fit separately to nine groups defined by the number of *Total Brands*, and they include one additional hierarchical parameter per group. MAD is measured in percent of total SOW

Table AII.
Model performance
in terms of total
brands ranked

Table AIII.
Correlations between
change in SOW and
changes in rank,
satisfaction and
model estimates
of SOW

Change in SOW	Change in Logit(SOW)
<i>With changes in logit of</i>	
Fixed parameter	
Attitudinal equity(AE), $s = 1$	0.379
Conversion, $s = 1$	0.374
WAR, $p = 2/[(Total Brands)+1]$	0.381
One-parameter	
Attitudinal equity	0.376
Conversion model	0.383
Truncated geometric	0.365
Nine-parameter	
Attitudinal equity	0.381
Conversion	0.384
WAR	0.380
Truncated Geometric	0.344
Logit(SOW) predicted as	
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \text{LN}(Total Brands + 1 - Rank)$, 1 Class	0.341
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \text{Satisfaction} + \beta_2 (Total Brands)$, 1 Class	0.293
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \log(Total Brands + 1 - Rank)$, 9 Classes	0.358
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \text{Satisfaction}$, 9 Classes	0.108
<i>(continued)</i>	
Change in SOW	
<i>With change in</i>	
Fixed parameter	
Zipf-AE, $s = 1$	0.403
Zipf-PM, $s = 1$	0.393
WAR, $p = 2/[(Total Brands)+1]$	0.407
One-parameter	
Zipf-AE	0.400
Zipf-PM	0.404
Truncated Geometric	0.401
Nine-parameter	
Zipf-AE	0.401
Zipf-PM	0.403
WAR	0.403
Truncated Geometric	0.371
SOW as predicted from	
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \text{LN}(Total Brands + 1 - Rank)$, 1 Class	0.364
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \text{Satisfaction} + \beta_2 (Total Brands)$, 1 Class	0.315
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \log(Total Brands + 1 - Rank)$, 9 Classes	0.366
$Logit(SOW_{ij}) = \beta_0 + \beta_1 \text{Satisfaction}$, 9 Classes	0.098

Change in SOW	Change in Logit(SOW)
Rank	-0.285
Log(Rank)	-0.332
Satisfaction	0.066
Log(Satisfaction)	0.078
Log((Total Brands)+1-Rank)	-0.038 ^a
Logit(Rank/(Total Brands+1))	-0.133
Logit(Satisfaction/(Maximum(Satisfaction)+1))	0.069
Recommend Intention	0.065
Log(Recommend Intention)	0.056 ^a
Net Promoter	0.067
Log(Net Promoter)	0.068
Rank	-0.278
Log(Rank)	-0.328
Satisfaction	0.111
Log(Satisfaction)	0.098
Log((Total Brands)+1-Rank)	-0.010 ^a
Logit(Rank/(Total Brands+1))	-0.180
Logit(Satisfaction/(Maximum(Satisfaction)+1))	0.112
Recommend Intention	0.070
Log(Recommend Intention)	0.061 ^a
Net Promoter	0.073
Log(Net Promoter)	0.073

Notes: $N \geq 2,686$ for each correlation. ^aAll correlations are significant at the level $p < 0.001$, except for the correlations of change in $\log((Total\ Brands)+1-Rank)$ with change in SOW and Change in Logit(SOW) (where $p = 0.050$, and $p = 0.608$, respectively), and the correlations of change in $\log(Recommend\ Intention)$ with change in SOW and change in $\logit(SOW)$ (where $p = 0.004$, and $p = 0.001$, respectively)

Table AIII.

Appendix 2. Technical Appendix

Perceptions are relative: an examination of the relationship between relative satisfaction metrics and share of wallet

Most managers would consider our empirical analysis to be overly complex and therefore difficult to read and interpret. As a result, we believe that the findings and implications of our investigation would be lost in the technical descriptions of the models and analytics if included in the main document.

The goal of our research is to spur a change in current practice (as well as academic research) with regard to the measurement and management of customer satisfaction. Therefore, in an effort to maximize the impact on both the practice and the science of service management, we have chosen to present the details of our analysis in a Technical Appendix.

Models investigated

Currently, the three most widely used models for linking relative ranked satisfaction with SOW are:

- (1) The “attitudinal equity” model of Hofmeyr *et al.* (2008). Because this is based upon the Zipf distribution, we refer to this throughout the manuscript as Zipf-AE.
- (2) The “power of the mind” model of Louw and Hofmeyr (2012). Because this is also based upon the Zipf distribution, we refer to this throughout the manuscript as Zipf-PM.
- (3) The “Wallet Allocation Rule” model of Keiningham *et al.* (2011). Throughout manuscript we refer to this as WAR.

Each of these models is described below. It is important to note that the notation we use to describe the models differs slightly from their original presentation in the respective articles that introduced them (i.e. Hofmeyr *et al.*, 2008; Keiningham *et al.*, 2011; Louw and Hofmeyr, 2012). This is done so that common variable labels can be used across all models investigated. The models are in fact unchanged from their original presentation.

Zipf-AE

The Zipf-AE model (Hofmeyr *et al.*, 2008) posits that customer *i*’s share of wallet for brand *j* (with *Rank_{ij}*) in a usage set of size *m_i* is:

$$SOW_{ij}^{(ZIPf-AE)} = \frac{1}{Rank_{ij}^{s(m_i)} \left[\sum_{k=1}^{m_i} \left(\frac{1}{Rank_{ik}^{s(m_i)}} \right) \right]}$$

where *s(m_i)* is a constant that depends on the number of brands (*m_i*) taken from Table AII of Hofmeyr *et al.* (2008), and these constants were found with the “solver” function in Excel” (p. 190). (In the summation, *k* is the index that runs through all possible brands in the product category.) We have added *m_i* to the notation to clarify that there are different exponents depending on the number of brands in customer *i*’s usage set. Hofmeyr *et al.* (2008) recommend using the values *s(m_i)* published in their article (p. 191, step 2), but one could estimate the exponent using the data at hand as discussed below. AE is a Zipf probability distribution, and its values must therefore sum to 1 (for a given number of brands); the constant in brackets in the denominator guarantees that the sum is 1.

Zipf-PM

The Zipf-PM model (Louw and Hofmeyr, 2012; Hofmeyr, 2012) posits that customer *i*’s share of wallet for brand *j* (with *Rank_{ij}*) in a usage set of size *m_i* is:

$$SOW_{ij}^{(ZIPf-PM)} = \frac{Share_{ij}}{Rank_{ij}^s \left[\sum_{k=1}^{m_i} \left(\frac{Share_{ik}}{Rank_{ik}^s} \right) \right]}$$

where j is the brand being scored, and m is the number of brands (Hofmeyr, 2012, p. 18 states: “There is still an ‘s’, but it’s set to ‘1’. No exponential transform needed”). Also, following Hofmeyr (2012), we define $Share_{ij}$ as the share of “total satisfaction” that customer i assigns to brand j :

$$Share_{ij} \equiv \frac{Satisfaction_{ij}}{\sum_{k=1}^m Satisfaction_{ik}}.$$

The distinguishing characteristic of the Zipf-PM approach is that they propose using “the share that a brand’s rating achieves as a percent of the sum of a respondent’s ratings of relevant brands” in the Zipf distribution equation (Louw and Hofmeyr, 2012, p. 11).

WAR

The WAR (Keiningham *et al.*, 2011) posits that customer i ’s SOW for brand j is:

$$SOW_{ij}^{(WAR)} = \left(1 - \frac{Rank_{ij}}{m_i + 1}\right) \times \left(\frac{2}{m_i}\right).$$

WAR is a fixed parameter model; as such, no estimation (i.e. data fitting) is required to estimate the relationship between rank transformed satisfaction and share of wallet.

General WAR

The WAR is actually a special case of the family of discrete probability distributions that assign an arithmetic sequence of probabilities (representing SOW values) to successive ranks. Consequently, all of these distributions imply the SOW is a linear function of the brand’s rank. If $p(m_i)$ represents the probability (or SOW value) assigned to rank 1 when there are total of m_i brands, then the generalization of WAR would allow this probability to vary by m_i -category, so that the generalization becomes:

$$SOW_{ij}^{(GWAR)} = p(m_i) - \frac{2(Rank_{ij} - 1)[m_i p(m_i) - 1]}{m_i(m_i - 1)},$$

where $p(m_i)$ is the SOW assigned to the brand with rank 1, $1/m_i \leq p(m_i) < 2/m_i$ (note that non-positive SOW values would be assigned to ranks if $p(m_i) \geq 2/m_i$, and if $p(m_i) < 1/m_i$, we would not have a non-increasing sequence of SOW values that adds to 1). Consequently, the arithmetic sequence of SOW assignments begins with $p(m_i)$ at rank 1 and decreases by:

$$\frac{2[m_i p(m_i) - 1]}{m_i(m_i - 1)}$$

for each successive rank. Since these probabilities must add to 1, this is the only arithmetic sequence possible when the SOW value $p(m_i)$ is assigned to rank 1. Note that the WAR is the special case where $p(m_i) = 2/(m_i + 1)$, and the discrete uniform is the case where $p(m_i) = 1/m_i$ (i.e. in this case all ranks would be assigned the same SOW value).

Other models

In addition to the three most widely used models, we investigated whether better approaches existed for linking relative satisfaction levels to share of wallet. Based upon the properties of the Zipf-based models, we examined another discrete distribution that seemed plausible: the truncated geometric model. In addition, because hierarchical regression models are commonly used in satisfaction and SOW research (e.g. Keiningham *et al.*, 2003) we investigate hierarchical regression models.

These models are described below.

Truncated geometric

The truncated geometric model provides an alternative way of accommodating the decay in SOW with increasing rank. According to this model, if customer i assigns $Rank_{ij}$ to brand j (in a usage

set of m_i brands), then customer i 's share of wallet for brand j is:

$$SOW_{ij}^{(G)} = \frac{p(m_i)[1-p(m_i)]^{Rank_{ij}-1}}{[p(m_i) \sum_{k=1}^{m_i} ([1-p(m_i)]^{k-1})]}$$

(The denominator would be $[1-(1-p(m_i))^{m_i}]$ if there were no rank ties.) Here $p(m_i)$ represents the share of wallet corresponding to the brand ranked 1 (in the untruncated case). In the one-parameter model, it is estimated across all customers, and in the nine-parameter model it is estimated separately across all customers in one of the nine usage sets (m_i), as is the case with the Zipf models. Following the truncated geometric paradigm, the SOW for the brand with $Rank_{ij}$ is proportional to the probability of not finding the $(Rank_{ij}-1)$ preferred brands, where we assume an equal failure probability $(1-p(m_i))$ of not finding each one of the preferred brands.

Hierarchical regression

Finally, we also consider two-level regression models for SOW_{ij} , which represents customer i 's SOW for brand j when it is assigned a rank of $Rank_{ij}$ among the m_i total brands in that category:

$$\text{Logit}(SOW_{ij}^{(HR)}) = f(Rank_{ij}, m_i) + \varepsilon_i + \varepsilon_{ij},$$

and:

$$\text{Logit}(SOW_{ij}^{(HR)}) = f(Satisfaction_{ij}, m_i) + \varepsilon_i + \varepsilon_{ij},$$

where ε_i represents the customer random effect and ε_{ij} represents overall model error, both of which are normally distributed with mean zero, and distinct variances. The random effect at the customer level provides a flexible way to accommodate the natural dependence among the observations from one customer across brands within a product category. As models in this category, we considered the best two-predictor regressions based on the total number of brands (m_i) and either $Rank_{ij}$ or $Satisfaction_{ij}$. (Models with more than two predictors based on these variables generally did not explain more than an additional 0.6 percent of the variance in SOW .) $Rank_{ij}$ and $Satisfaction_{ij}$ were considered directly as candidate predictors along with the log transforms of each variable, and the logit transforms of each when it is expressed as a proportion, i.e. the logits of the proportions $P_R = Rank_{ij}/(m_i+1)$, and $P_s = Satisfaction_{ij}/(Maximum\ Rating+1)$ (here the maximum possible satisfaction rating is 10). As candidate predictors, we also considered the components of these logit transforms: $\log(m_i+1-Rank_{ij})$, and $\log(Maximum\ Rating+1-Satisfaction_{ij})$.

Parameter estimation

With the exception of the hierarchical regression models, the other models are discrete probability distributions that automatically provide SOW estimates that sum to 1 for each customer. If we group the customer observations by the number of total brands ranked (m) we can view each set of customer ranks within m -category as an estimate of the same continuous multinomial distribution (Johnson, 1960). This is complicated by the fact that we allow tied ranks, but it still provides a straightforward method of obtaining maximum likelihood estimates for the parameters in each of the proposed models. This approach is consistent with the hierarchical structure of repeated customer rankings within each brand category.

For each of the discrete-distribution models, we consider, whenever possible, three versions: a popular fixed-parameter version, a one-parameter version (where the parameter does not vary by the total number of brands, m), and the M -parameter version where the parameter is allowed to vary by m -category. All three versions are possible for Zipf-AE and Zipf-PM. We refer to the WAR as a fixed parameter model (since in this case $p(m) \equiv 2/(m+1)$, so that no estimation is required). There is no one-parameter version of WAR (since $p(m)$ must vary with the total number of brands, m), and no popular (or established) fixed-parameter version of the truncated geometric.

Consequently, we explore ten versions of the discrete distribution models, and consider four hierarchical regression models (where for each of set of predictors, we estimate one version with common parameters across all m -categories, and another with separate parameters within each m -category). The 14 models in total investigated are as follows.

Fixed parameter models:

- (1) Zipf-AE, $s = 1$
- (2) Zipf-PM, $s = 1$
- (3) WAR, $p = 2/[(Total\ Brands)+1]$

One-parameter models

- (4) Zipf-AE
- (5) Zipf-PM
- (6) Truncated Geometric

Nine-parameter models (one parameter per Total Brands class) [8]

- (7) Zipf-AE
- (8) Zipf-PM
- (9) WAR
- (10) Truncated Geometric

Hierarchical regression (customer i , product category j) [9]

- (11) $\text{Logit}(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \log(Total\ Brands + 1 - Rank)$, 1 Class
- (12) $\text{Logit}(SOW_{ij}) = \beta_0 + \beta_1 Satisfaction + \beta_2 (Total\ Brands)$, 1 Class
- (13) $\text{Logit}(SOW_{ij}) = \beta_0 + \beta_1 \log(Rank) + \beta_2 \log(Total\ Brands + 1 - Rank)$, 9 Classes
- (14) $\text{Logit}(SOW_{ij}) = \beta_0 + \beta_1 Satisfaction$, 9 Classes

We chose maximum likelihood estimation throughout as relatively non-controversial way of finding good representative estimates of each model. For the discrete distributions we used the continuous multinomial likelihood (next section); for the hierarchical regressions we used standard multivariate normal distributions.

Maximum likelihood estimation using the continuous multinomial

Let SOW_{ij} represent the customer i 's SOW for brand j , when there are a total of m_i brands, then the vector of SOW for customer j , $[SOW_{i1}, \dots, SOW_{im}]$ would occur with the continuous multinomial probability (Johnson, 1960):

$$P[SOW_{i1}, \dots, SOW_{im}] = C(m(\underline{r})) \prod_{j=1}^m \frac{1}{\Gamma(SOW_{ij} + 1)} [p(r_j)]^{SOW_{ij}} \quad (A1)$$

with $\sum_{j=1}^m SOW_{ij} = 1$, $\sum_{j=1}^m p(r_j) = 1$, and where $C(m(\underline{r}))$ would be the normalizing constant necessary across the m -category with a given vector of ranks $\underline{r} = (r_1, \dots, r_m)$ (to compensate for the continuous nature of the share of wallet values; see for example: Gasbarra *et al.*, 2011, p. 37, Equation (3)). $C(m(\underline{r}))$ is not part of the kernel likelihood and consequently does not affect the maximum likelihood estimates. Each of the different discrete distributions proposed in this study (the WAR, Zipf and Truncated Geometric Models) provides alternative

models for how the rank-category probabilities $\{p(r_k)\}$ are determined as functions of the ranks $\{\underline{r}\} = \{r_1, \dots, r_m\}$.

Since we must allow for tied ranks, there are more than one possible set of ranks per m -category. For example, sets of ranks $\{2,2,2\}$ and $\{1,2,3\}$ would be two of the four possible sets of ranks possible there are $m = 3$ brands. Let $N(m(\underline{r}))$ represent the total number of customers that use a particular set of ranks $\{\underline{r}\}$ when there are $m(\underline{r})$ total brands. For example, $N(m(2,2,2))$ would represent the total number of customers that use the ranks $\{2,2,2\}$ to rank 3 brands, i.e., $m(2,2,2) = 3$. Let $S(r_k, m(\underline{r}))$ represent the sum of the customer share of wallet values for all brands with rank r_k across the total number of customers, $N(m(\underline{r}))$, who use the specific set of ranks $\{\underline{r}\}$ (which includes r_k), i.e.:

$$S(r_k, m(\underline{r})) = \sum_{i=1}^{N(m(\underline{r}))} SOW_{ir_k}$$

Here we are summing overall customer share of wallet values SOW_{ir_k} that correspond to the same rank r_k . Thus, given a specific set of ranks $\{\underline{r}\}$, with L distinct ranks, the corresponding vector of the total share of wallets for each of those ranks:

$$[S(r_1, m(\underline{r})), \dots, S(r_L, m(\underline{r}))],$$

would also have a continuous multinomial distribution, that is:

$$P[S(r_1, m(\underline{r})), \dots, S(r_L, m(\underline{r}))] = D(\underline{r}, N(m(\underline{r}))) \prod_{k=1}^L \frac{1}{\Gamma(S(r_k, m(\underline{r})) + 1)} [p(r_k)]^{S(r_k, m(\underline{r}))}, \quad (A2)$$

where $D(\underline{r}, m(\underline{r}))$ is the appropriate normalizing constant. Note that in those cases where customers are assigning only one rank across the full set of brands (e.g. when $\{2,2,2\}$ is assigned to each of 3 brands ($m = 3$)), then there is only one distinct category, $L = 1$, and the distribution in (2) becomes an example of the degenerate multinomial case where there is only one category.

For a fixed number of total brands m , there can be many possible sets of ranks $\{\underline{r}\}$ that are used by customers (because of the different ways there can be ties), and a different multinomial distribution for each set. The full likelihood would then be a product of all the independent likelihoods (of the form given in (2)) across all the distinct sets of ranks $\{\underline{r}\}$ that are used by customers when m is the total number of brands ranked. For example, when $m = 3$, the full likelihood would be the product of four different versions of the likelihood in (2), that correspond to the four different ways of assigning ranks ($\{1, 2, 3\}$, $\{2, 2, 2\}$, $\{1.5, 1.5, 3\}$, $\{1, 2.5, 2.5\}$, and the number of distinct ranks for each set are $L = 3, 1, 2$, and 3, respectively). As the sample size increases, the non-integral nature of the $\{S(r_k, m(\underline{r}))\}$ makes very little difference, and estimates based on the rounded sufficient statistics (and the standard multinomial distribution) are virtually the same as they would be using the continuous multinomial.

Application of the models

Overall model performance (cross-sectional). Table AI shows four comparisons for each model: MAD, and RMSE across all observations and by customer.

In Table AI, the fixed-parameter versions of the discrete distribution models do remarkably well overall. Among these distributions, the fixed-parameter Zipf-AE model is best in terms of MAD, both overall and per customer, and it actually outperforms all models (including the regression models) in terms of average customer RMSE. The nine-parameter version of Zipf-AE is the best performer in terms of overall RMSE. Nevertheless, the discrete distributions generally do quite well: eight of the other ten discrete distributions have RMSE values that are

within 1.5 percent of the best fit. The one exception is the fixed parameter Zipf-PM which has an RMSE that is 6 percent larger overall, relative to the best performing nine-parameter Zipf-AE model.

The nine-class regression with $\log(\text{Rank})$ is actually the best performing model in terms of MAD, and it is just ahead of the fixed parameter Zipf-AE with MAD values that are 1.6 and 0.9 percent larger overall, and per customer, respectively. This regression model is also uniformly the best among the four regression alternatives, but paradoxically it does not fit as well in terms of RMSE, where it actually achieves the 10th and 9th highest overall and per customer RMSE, respectively. Still, even in these cases its error rates are only larger than the lowest RMSE values by 2.2 percent overall, and 2.3 percent per customer. In contrast, the regression models based on *Satisfaction* are uniformly the worst models in every case, and here the error rates are substantially larger than the best model in every instance. Although the 9-class version of this model is the better performer, even its error rates range from being higher by 7.3 percent (MAD overall) to 12.6 percent (RMSE per customer).

Table AII provides a comparison of model performance by the number of total brands (m) that are considered by the customer. The fixed-parameter versions of Zipf-AE and WAR are the best in the two-product category with MAD values of 20.5 percent. WAR and Zipf-AE are equivalent in this case, where each predicts *SOW* values of (2/3,1/3) when the ranks are (1,2) and values of (1/2,1/2) when ranks are tied (1.5,1.5). This is the only category where the nine-class regression with $\log(\text{Rank})$ is not the best model, and even in the two-category case this regression model is nearly the best with a MAD that is 20.6 percent (relative to the best MAD of 20.5 percent). Across models, the lowest MAD values decrease by 64 percent as total brands increase across the six categories, and it ranges from 20.5 percent (when $m = 2$) to 7.4 percent (when $7 \leq m \leq 10$). The 9-parameter Zipf-AE model and the nine-class regression with $\log(\text{Rank})$ are the best overall performers across categories, and the Zipf-AE models are always among the top 5 models when total brands is less than seven ($m \leq 6$). Finally the regression models based on *Satisfaction* are the worst models overall, in terms of median rank across categories, although the nine-class regression on *Satisfaction* is the second best model in the last category ($7 \leq m \leq 10$). The regression models based on *Satisfaction* are uniformly the poorest performers when there are four or fewer total brands ($m \leq 4$).

Overall model performance (longitudinal)

Table AIII provides an analysis of the two-period data. This table summarizes the correlations of change in *SOW* and $\text{logit}(\text{SOW})$ with contemporaneous changes in model estimates and changes in other variables. For the eleven models that require parameter estimates (i.e. all models except for three fixed-parameter models), the parameters are estimated in period 1 and those estimates are then used to predict *SOW* in period 2 (using period 2 information on *Rank*, *Total Brands*, and *Satisfaction*).

Table AIII shows that the correlation between the two-period change in the WAR estimates of *SOW* (fixed-parameter version) and actual change in *SOW* are nominally the largest overall ($r = 0.407$, $p < 0.001$), but nearly all of the discrete distributions perform at the same level in terms of predicting change in *SOW* and change in $\text{logit}(\text{SOW})$. The weakest performing models are the nine-parameter truncated geometric model (for change in *SOW*: $r = 0.371$, $p < 0.001$) and the regression models (for change in *SOW*: the largest $r = 0.366$, $p < 0.001$). Surprisingly, the one-class regression model based on *Satisfaction* performs better than the nine-class version. The nine-class regression based on *Satisfaction* is by far the worst performer overall ($R^2 < 12$ percent in each case).

The comparison of regression models indicates that the one-class regression on *Satisfaction* and *Total Brands* provides an alternative way of “calibrating” satisfaction relative to total brands, so that it is comparable (in the two-period case) to a regression on a relative measure of satisfaction (like rank). This “calibration” is not achieved by the nine-class version of the

same model because, although this category-specific estimation of the coefficient for *Satisfaction* provides a better fit within category, it does not provide a single “calibration” of *Satisfaction* relative to the total brands, across the nine categories.

Finally, Table AIII shows the inadequacy of changes in absolute satisfaction, recommend intention, and NPS levels in correlating to changes in share of wallet. Changes in these variables explain <1 percent of the variation in changes in share of wallet.

About the authors

Dr Timothy Lee Keiningham is the Global Chief Strategy Officer and the Executive Vice President at Ipsos Loyalty. He is the Author and/or Editor of eight books, his most recent being *Why Loyalty Matters*. He has received best paper awards from the *Journal of Marketing* (twice), the *Journal of Service Research*, and *Managing Service Quality* (twice), and has received the Citations of Excellence “Top 50” award (top 50 management papers of approximately 20,000 papers reviewed) from Emerald Management Reviews. Tim also received the best reviewer award from the *Journal of Service Research*. Dr Timothy Lee Keiningham is the corresponding author and can be contacted at: tlkeiningham@yahoo.com

Dr Bruce Cooil is The Dean Samuel B. and Evelyn R. Richmond Professor of Management at the Owen Graduate School of Management, Vanderbilt University. His research interests include the adaptation of grade-of-membership and latent class models for marketing and medical research, estimation of qualitative data reliability, large sample estimation theory and extreme value theory. He has also written and consulted on models for mortality, medical complications, medical malpractice, and automobile insurance claims. His publications have appeared in business, statistics and medical journals, including the *Journal of Marketing Research*, *Journal of Marketing*, *Psychometrika*, *Journal of the American Statistical Association*, *Annals of Probability*, *Circulation*, and the *New England Journal of Medicine*.

Dr Edward C. Malthouse is the Theodore R and Annie Laurie Sills Professor of Integrated Marketing Communications at the Northwestern University and was the Co-Editor of the *Journal of Interactive Marketing* between 2005 and 2011. His research interests center on media marketing, database marketing, advertising, new media, and integrated marketing communications. He is the Co-Editor of *Medill on Media Engagement*, and has extensively published on these topics, with his work appearing in *Journal of Service Research*, *Journal of Consumer Psychology*, *Journal of Interactive Marketing*, *Expert Systems with Applications*, *Data Mining and Knowledge Discovery*, *IEEE Transactions on Neural Networks*, *Computers in Chemical Engineering*, *Journal of Advertising*, *Journal of Advertising Research*, *Journal of Broadcasting and Electronic Media*, *International Journal of Market Research*, *Journal of Media Business Studies*, *International Journal of Media Management*, and others.

Alexander Buoye is an Assistant Professor of Marketing at the Fordham University Schools of Business in New York. Prior to joining Fordham, he was the Head of Loyalty Analytics and a Senior Vice President at Ipsos Loyalty. He is the co-recipient of the 2011 Next Gen Market Research Disruptive Innovator Award for his role in the discovery and development of the Wallet Allocation Rule. His work has been accepted for publication in such journals as *Harvard Business Review*, *MIT Sloan Management Review*, *Journal of Service Research*, *Journal of Service Management*, and *Journal of Business Research*. He received his MA and PhD in Sociology and a Bachelor of Business Administration (Marketing) from the University of Notre Dame.

Dr Lerzan Aksoy is a Professor of Marketing at the Fordham University Schools of Business in New York. She is author and/or editor of four books, including *Loyalty Myths* and *Why Loyalty Matters*. She has received best paper awards from the *Journal of Marketing*, *Journal of Service Management*, and *Managing Service Quality* (twice), and has received the Citations of Excellence “Top 50” award (top 50 management papers of approximately 20,000 papers reviewed) from Emerald Management Reviews. She was awarded finalist for best paper in the *Journal of Service Research*. Lerzan also received best reviewer awards from the *Journal of Service Management* and the *Journal of Service Research*.

Arne De Keyser is a PhD-Candidate at the Ghent University, where he is also a Member of the Center for Service Intelligence. His interests focus on multichannel marketing management and customer experience management. His research is forthcoming in *Journal of Service Management* and *Managing Service Quality*. He won the Best Doctoral Service Paper Award at the second Positive Marketing Conference (New York, Fordham University, 2013) and received the Best Practitioner Presentation Award at the 2013 Frontiers in Service Conference (National Taiwan University).

Dr Bart Larivière is an Assistant Professor of Service Management at the Ghent University, where he is the Founder and the Executive Director of the Center for Service Intelligence. His research focuses on customer loyalty, servitization, servicescape, consumer well-being, and multichannel management. He intensively collaborates with Belgian companies bridging the gap between practice and academia. His research has been published in *Journal of Service Research*, *Journal of Service Management*, *Journal of Interactive Marketing*, the *European Journal of Operational Research*, and *Expert Systems with Applications*. He was finalist for best paper in *Journal of Service Research*, and his research won the Best Practitioner Presentation Award (twice) at the Frontiers in Service Conference.