Intern. J. of Research in Marketing 32 (2015) 195–206

Contents lists available at ScienceDirect

Intern. J. of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

The predictive ability of different customer feedback metrics for retention☆

Evert de Haan [a,*], Peter C. Verhoef [a,1], Thorsten Wiesel [b,c,2]

[a] Department of Marketing, Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands
[b] Marketing Center Münster, Westfälische Wilhelms-Universität Münster, Am Stadtgraben 13-15, 48143 Münster, Germany
[c] Affiliate Researcher of Marketing, University of Groningen, Faculty of Business and Economics, 9700 AV Groningen, The Netherlands

ARTICLE INFO

Article history:
First received on February 14, 2014 and was under review for 5 months
Available online 5 March 2015

Area Editor: Koen H. Pauwels

Keywords:
Customer feedback metrics
Customer satisfaction
Net Promoter Score
Customer Effort Score
Customer retention
Firm performance

ABSTRACT

This study systematically compares different customer feedback metrics (CFMs) – namely customer satisfaction, the Net Promoter Score, and the Customer Effort Score – to test their ability to predict retention across a wide range of industries. We classify the CFMs according to a time focus (past, present, or future) and whether the full scale of the CFM is used or whether the focus is only on the extremes (e.g., top-2-box customer satisfaction). The data for this study represent customers of 93 firms across 18 industries. Multi-level probit regression models, which control for self-selection bias of respondents, investigate firm-, customer-, and industry-level effects simultaneously. Overall, we find that the top-2-box customer satisfaction performs best for predicting customer retention and that focusing on the extremes is preferable to using the full scale. However the best CFM does differ depending on industry and the unit of analysis (i.e., comparing customers or firms with one another). Furthermore, combining CFMs, along with simultaneously investigating multiple dimensions of the customer relationship, improves predictions even further.

1. Introduction

New customer feedback metrics (CFMs), as Morgan and Rego (2006) label them, including Reichheld's (2003) Net Promoter Score (NPS) and Dixon, Freeman, and Toman's (2010) Customer Effort Score (CES), are introduced frequently. These CFMs promise to be "the best" indicator of (future) firm performance, prompting leading companies in a wide range of industries to start using them (Bain and Co., 2013). Academic research challenges these promises (e.g., Keiningham, Cooil, Aksoy, Andreassen, & Weiner, 2007; Keiningham, Cooil, Andreassen, & Aksoy, 2007; Morgan & Rego, 2006); however, most studies investigate only a limited range of firms, industries, and settings, and they lack comparability, because they use different dependent variables, research settings, methodologies, units of analysis, and so on. Marketing managers thus lack guidance on which CFM to monitor and how to interpret changes in these CFMs, which can lead to uncertainty, frustration, and even abandonment of the CFMs in question. Such outcomes might lessen marketing departments' accountability and influence (Verhoef & Leeflang, 2009), hinder firms from becoming more customer centric (Shah, Rust, Parasuraman, Staelin, & Day, 2006), and negatively affect marketing-mix performance (Mintz & Currim, 2013).

This study aims to provide, for a wide range of industries, insights into the impact of different CFMs, including which (combinations of) CFM(s) a firm should monitor, how to interpret changes in CFMs, and how this differs across industries. We use actual customer retention data to compare the predictive power of various CFMs across a large number of firms and industries. Our focus on customer retention reflects three key considerations: (1) customers are among the most important marketing assets of the firm, (2) an almost one-to-one relationship exists between the value of the customer base and firm value, and (3) CFMs are frequently used as indicators of future loyalty (Gupta, Lehmann, & Stuart, 2004; Rust, Zeithaml, & Lemon, 2000). We simultaneously analyze customer-, firm-, and industry-level effects of CFMs on retention, using multi-level models to support comparisons of within-firm, between-firm, and between-industry effects. With this approach, we can provide generalizations and recommendations on which CFM(s) to monitor and how this differs across industries. We use surveys to collect CFM scores and customer background information from 6649 respondents, who in total filled out 8924 firm evaluations for 93 firms across 18 industries. In a follow-up survey two years later, filled out by 1308 respondents who provided 1375 firm evaluations (i.e., a 15.4% response rate), we measure our dependent variable, customer retention. We measure the usefulness of the CFMs in predicting

☆ Acknowledgement: This project is based on data collected through an annual survey on customer performance. The authors thank Metrixlab, MICompany, and the Customer Insights Center of the University of Groningen for data support. They also thank the participants of the 2013 GSBE Marketing-Finance Symposium in Maastricht, the participants of the 2013 INFORMS Marketing Science Conference in Istanbul, the two anonymous reviewers, the area editor, and the editor, Eitan Muller, for their valuable comments.
* Corresponding author. Tel.: +31 50 363 9621.
E-mail addresses: evert.de.haan@rug.nl (E. de Haan), p.c.verhoef@rug.nl (P.C. Verhoef), thorsten.wiesel@wiwi.uni-muenster.de (T. Wiesel).
[1] Tel.: +31 50 363 7320.
[2] Tel.: +49 251 83 28207.

http://dx.doi.org/10.1016/j.ijresmar.2015.02.004
0167-8116/

retention by analyzing both in-sample fit (with the Akaike information criterion [AIC] and Bayesian information criterion [BIC]) and out-of-sample fit (with the Gini coefficient, top-decile lift, and hit rate) for which we use a one-third holdout sample. In terms of the out-of-sample fit, the top-decile lift is an important criterion when the main goal is to identify customers most likely to churn, while the Gini coefficient and hit rate are important when the goal is to judge accuracy for all customers (i.e., make good predictions about both retainers and churners) (Blattberg, Kim, & Neslin, 2008).

Our results show that the top-2-box customer satisfaction score offers the single-best predictor of retention across industries. In general, transforming scales of the CFMs to capture the proportion of most satisfied customers (as is done with the top-2-box customer satisfaction) or splitting customers up into groups (as is done with the promoters and detractors of the NPS) is preferable to using the full scale of the CFMs. In addition, our results show that the CES in itself has little to no predictive power and performs the worst of all CFMs studied. Which CFM performs best in predicting retention is however industry dependent, and it also depends on whether the CFM is meant to be used for customer management (i.e., compare customers of the same firm with one another) or to analyze the competitive position of a firm (i.e., compare different firms in the same industry with one another). Combining metrics, especially the CES with the customer satisfaction-related CFMs, results in improved out-of-sample retention predictions. A dashboard of CFMs that measure different dimensions, as indicated in our conceptualization, is preferable to monitoring a single CFM.

Table 1 illustrates the study's contribution with a selective literature overview. This study is the first to investigate the predictive power of CFMs over three levels (i.e., customer, firm, and industry) simultaneously. In doing so, we can distinguish between the heterogeneity of customers (i.e., which CFM is most appropriate for customer management) and the heterogeneity of firms (i.e., which CFM is most appropriate for competitive positioning). Furthermore, this study is one of the first to use the official NPS, as Reichheld (2003) intended it, and to investigate the CES in line with Dixon et al.'s (2010) approach. As such, we test the ability of two recently introduced metrics that have become famous as key CFMs. We also combine CFMs to determine whether using multiple CFMs improves the predictive power, as often done in firms' dashboards. Furthermore, we predict actual future performance, in contrast with other studies that investigate only same-period correlations (e.g., Anderson, Fornell, & Mazvancheryl, 2004; Keiningham, Cooil, Andreassen, & Aksoy, 2007). In doing so, we can test the usefulness of CFMs for predictive purposes. Finally, we judge the (combination of) CFMs on their out-of-sample predictions to determine whether they have real incremental predictive power and to overcome over-fitting problems. In doing so, we can show the validity and robustness of our results.

In summary, we investigate how valuable various CFMs are in predicting retention in different situations, both to increase the academic understanding of these CFMs and to help managers select the best CFM(s) according to their situation and demands. To do so, we examine (1) the overall usefulness of different CFMs in predicting retention, (2) the differences of this usefulness between industries, (3) the differences between different units of analysis (e.g., customer- or firm-level retention for customer management and competitive analysis purposes), and (4) the incremental power of monitoring multiple CFMs over using a single CFM. These insights can help practitioners decide which (combination of) CFM(s) to use in different situations and help academics understand how valuable CFMs are and what the determining factors (e.g., industry differences, unit of analysis) for this are.

2. Conceptual background

As Gupta and Zeithaml (2006) note, it is critical to understand the relationships among CFMs, customer behavior, and firm performance. Although the positive relationship between customer satisfaction and firm performance is well established (Gupta & Zeithaml, 2006; Hanssens, 2009), a similar state does not exist for other CFMs. In this section, we classify the CFMs under study and highlight the importance of the unit (or level) of analysis.

2.1. Conceptual classification of metrics

Research in marketing has discussed a large number of metrics. Farris, Bendle, Pfeifer, and Reibstein (2006) classify these metrics as share-of-mind metrics and consider customer satisfaction and willingness to recommend a specific sub-group within these metrics. In marketing research practice, these metrics are known as customer feedback metrics (CFMs) (Morgan & Rego, 2006). These CFMs have specifically gained attention in the service and relationship marketing and customer (relationship) management literature. Given the broad number of CFMs, we distinguish between these metrics on two dimensions. The first dimension is introduced by Bolton, Lemon, and Verhoef (2004) and more recently by Zeithaml et al. (2006), who focus on the time span of measures and distinguish between more backward-looking (including the present) and more forward-looking metrics. Forward-looking CFMs focus on what customers plan to do in the future and may signal something about the future performance of the relationship. Reichheld's (2003) NPS is an example of a forward-looking CFM because it considers the willingness to recommend a firm in the future, which may also signal one's future relationship with the firm (e.g., Zeithaml et al., 2006). Backward-looking metrics focus on the past and current performance of a company toward customers. The CES is a typical backward-looking CFM because it measures the

Table 1
Literature overview on (predictive) performance of CFMs (selection).

Study	Level of analysis			CFM			Combine multiple CFMs	Predictive power	Out-of-sample prediction
	Customer	Firm	Industry	Satisfaction	NPS	CES			
Hallowell (1996)		√		√					
Ittner and Larcker (1998) (ch. 3)	√			√				√	
Ittner and Larcker (1998) (ch. 5)		√	√	√				√	
Mittal and Kamakura (2001)	√			√				√	
Anderson et al. (2004)		√	√	√					
Gruca and Rego (2005)		√	√	√				√	
Morgan and Rego (2006)		√		√	[1]			√	
Keiningham, Cooil, Aksoy, Andreassen, and Weiner (2007)	√		√	√	[1]		√	√	
Keiningham, Cooil, Andreassen, and Aksoy (2007)		√	√	√	[1]				
Rego et al. (2013)		√		√				√	
Van Doorn et al. (2013)		√		√	√			[2]	
Current paper	√	√	√	√	√	√	√	√	√

[1] A proxy question used, instead of the official NPS question developed by Reichheld (2003).
[2] Van Doorn et al. (2013) test predictive power but find no significant effects of the CFMs.