Marketing Science Institute Working Paper Series 2023

Report No. 23-111

The Net Promoter Score (NPS) Fails to Predict Revenue Growth

Sebastian Lundmark, Jon A. Krosnick, Ellen Konar, Matthew K. Berent, Yphtach Lelkes, Daniel Schneider, Randall K. Thomas, Alexander Tahk and David Yeager

The Net Promoter Score (NPS) Fails to Predict Revenue Growth

Sebastian Lundmark

Stanford University

Jon A. Krosnick

Stanford University

Ellen Konar

Stanford University

Matthew K. Berent

Matt Berent Consulting

Yphtach Lelkes

University of Pennsylvania

Daniel Schneider

Kry AB, Stockholm

Randall K. Thomas

Ipsos Public Affairs North America

Alexander Tahk

University of Wisconsin

David Yeager

University of Texas at Austin

Keywords: Net Promoter Score, Likelihood to Recommend, Total Revenue Growth, Survey Methodology

Jon Krosnick is University Fellow at Resources for the Future. Address correspondence to Sebastian Lundmark (email: sebastian.lundmark@som.gu.se), University of Gothenburg, Seminariegatan 1, Gothenburg, 413 14, Sweden, or Jon A. Krosnick, Stanford University, 432 McClatchy Hall, Stanford, CA 94305 (email: krosnick@stanford.edu). This paper was written with the financial support of the Marketing Science Institute (Grant 41914).

The Net Promoter Score (NPS) Fails to Predict Revenue Growth

Abstract

The Net-Promoter Score (NPS) may be the most popular measure of customer satisfaction, touted as the single best predictor of a company's revenue growth. However, little evidence exists in the literature to support this claim or the notion that it is the only measure of customer satisfaction needed. This paper reports results obtained with data from five surveys exploring 53 evaluations of satisfaction with companies in the U.S. and their revenue growth. The NPS was not a reliable predictor of revenue growth when represented in various different ways, when estimating revenue growth in various different ways, when allowing for various different functional forms of the relationship, and when measuring likelihood to recommend using various different question wordings. Asking people how many times they have recommended a company in the recent past did predict growth marginally significantly and positively. Thus, the value of the NPS is called into question by this evidence.

The Net Promoter Score (NPS) Fails to Predict Revenue Growth

In 2003, Fred Reichheld proposed that the Net-Promoter Score (NPS) should be the 'one number you need to grow' for any company looking to increase its revenue. Furthermore, he said, because it is so effective, this score should replace all other customer satisfaction metrics (Reichheld, 2003; 2006). Part of the appeal of the NPS is its simplicity, assessed with just one survey question: "How likely are you to recommend [company X] to a friend or a colleague?"[1] With this measurement tool, companies were told that they can assess their customers' loyalty, satisfaction, experiences, and—most importantly—predict future relative company earnings. Fifteen years ago, the NPS became a virtual standard across the Fortune 500 (Inc Magazine, 2006) and remains wildly popular among many companies today, large and small.

A few researchers have questioned the claim that the NPS is the best predictor of growth (Hayes, 2008; Keiningham, Cooil, Andreassen, and Aksoy, 2007; Morgan and Rego, 2006). Others have criticized the lack of transparency in the analytical approaches and the research designs used in Reichheld's original work (Grisaffe, 2007, p. 48; Keiningham et al., 2007). Indeed, Reichheld (2003; 2006) wrote revealed little about how he reached the conclusion that the NPS is optimal.

This paper reports the results of a new investigation of the predictive ability of the NPS, analyzing data from experiments embedded in five national surveys administered to samples of American adults. These datasets were used to analyze whether the NPS predicted revenue growth and whether its predictive power can be improved by using other methods of calculating NPS scores or other wordings of the survey question.

[1] The present manuscript will refer to this question as the 'likelihood to recommend' question.

We begin below by reviewing the NPS and the main critiques that have been lodged at it in the past. Next, we offer a new perspective, grounded in the survey methodology literature on optimal question design, of possible shortcomings in the NPS question and leading to possible improvements in its design. Next, we describe the results of the new empirical investigation and spell out its implications for best practices.

The NPS calculation

In 2003, Fred Reichheld proposed that customers needed only report their likelihood to recommend a company to a friend or a colleague for that company to estimate a reliable predictor of future revenue growth. The customers should simply answer "How likely is it that you would recommend [company X] to a friend or colleague?", reporting their likelihood by choosing one out eleven horizontally aligned response options numbered from 0 to 10, with verbal labels above 0 (Not at all likely), 5 (Neutral), and 10 (Extremely likely).

Through a data mining process of consumer surveys from around 50 companies (Satmetrix, 2004) and tests of optimal explanatory power, Reichheld (2003) found that a recode of responses of likelihood to recommend question produced the optimal growth prediction; respondents who reported 9 or 10 likelihood to recommend were the 'promoters' of a company, respondents who reported 7 or 8 were the "passive promoters," and respondents who reported 0 to 6 likelihood to recommend were the company's "detractors." By subtracting the percentage of detractors from the percentage of promoters, companies would obtain their Net Promoter Score (NPS) (see Eq. 1). [2]

$$NPS = \frac{Promoters\ (9-10)}{Total\ Respondents} - \frac{Detractors\ (0-6)}{Total\ Respondents} \qquad \text{(Eq. 1)}$$

[2] According to Reichheld (2003), the median NPS in his data was 16% and should serve as base line for what was a high or low NPS.

The analytical approaches Reichheld (2003) used for discovering that the NPS outperformed other measurements, nor what those other measurements were, remain unclear. It seems that a variety of coding schemes were applied to the likelihood to recommend responses, which were then compared to each other and to other customer satisfaction metrics. To the authors' knowledge, it is still unknown what 50 companies were included in the Reichheld's (2003; 2006) dataset, and the reader is simply left with only the graphs that illustrate the final coding of NPS and revenue growth for 14 out of the 50 companies (see Figure 1).

Despite the simplicity of the NPS and the lack of transparency of Reichheld's (2003; 2006) methods, there is some face validity in the claim that revenue growth is causally linked to customers' likelihood to recommend the company. If a customer claims to be likely to recommend a company, it would be reasonable that this increases the likelihood of him/her recommending that company, and sequentially, increases the likelihood that the friend or colleague ends up buying something of the company.

Despite that there might be some theoretical credence to a causal link between the 'likelihood to recommend' question and company revenue growth, studies have directed attention to the lack of rigid research designs and peer-reviewed evaluations of the NPS (Morgan and Rego, 2006; Grisaffe, 2007; Keiningham et al., 2007; Hayes, 2008).[3] Unfortunately, none of the studies attempting to assess the validity of the NPS did so by asking the 'likelihood to recommend' question that Reichheld suggested nor presented the question with the ordained eleven response options.

[3] The literature review in this manuscript will focus on peer-reviewed published research about the NPS, thus excluding such sources as the many blogs, opinion pieces, and unpublished working papers that have discussed the NPS.

Furthermore, due to the data-mining nature of Reichheld's (2003) analyses, the fixed cutoffs of promoters and detractors might be problematic. The lack of significant relationship between the NPS and growth in the published critiques against the NPS might have stemmed from the fact that the cutoffs merely fit the data he used for inventing the score (Grisaffe, 2007, p. 47). The present manuscript attempts to rectify this by investigating the exact question that Reichheld's (2003) proposed as well as comparing Reichheld's NPS to scores obtained from other 'likelihood to recommend' question formats.

Insights from survey methodology research

Question wording and its impact on validity Research have outlined how to measure attitudes best (Krosnick, 1999). First, the main obligation of each survey practitioner is to minimize task difficulty for all respondents (Krosnick 1999, p. 548). Every survey question should follow common conversational rules, use simple and unambiguous words, and ask questions with as few words as possible, without losing the intent of the question (Tourangeau, 1984; Krosnick, 1999; Shaeffer et al., 2005). In this regard, the 'likelihood to recommend' question uses few and relatively unambiguous words. However, ambiguousness might increase because the question does not state whether the customer should evaluate the company as a whole or a specific aspect of the company or the company (e.g., its service, quality of food, or brand). Hence, rewording the question to remove this ambiguity may increase the validity of the measurement. An alternative version of the 'likelihood to recommend' question is presented in Table 1 (question version 2).

Response options and their impact on validity Second, when developing a survey question, the practitioner should offer the respondents response alternatives that validly capture the strength of the attitude. The practitioner must also decide how to label the response options and how many response options to include. According to most survey research, both the number and the labeling of the scale points correlates with task difficulty (Krosnick, 1999; Lundmark et al., 2016). If the

response options are not labeled correctly, or a non-optimal number of response options are provided, respondents will have a harder time giving a valid answer to the survey question (Krosnick and Fabrigar, 1997).

Krosnick and Fabrigar (1997) argued that introducing more than two response options increased respondents' ability to translate their attitude to a survey response. Furthermore, instead of using numbers to label the response options, fully labeling them produced more accurate measurements (Krosnick and Fabrigar, 1997). Consequently, providing a greater number of response alternatives enabled respondents to give a more valid representation of the strength of their attitudes (Krosnick and Fabrigar, 1997). Moreover, using numbers instead of verbal labels was especially negative for validity among respondents with lower than average cognitive ability (Krosnick and Fabrigar, 1997). Although more response options increased the validity of the measurement, too many decreased the respondents' ability to distinguish between the different response options (Krosnick and Fabrigar, 1997).

The polarity of the survey question and its impact on validity Theoretically, a unipolar attitude exists on a continuum, ranging from a lack of the attitude to a strong attitude (e.g., not at all to extremely). A bipolar attitude ranges from a strong negative to a strong positive attitude (e.g., extremely negative to extremely positive). The 'likelihood to recommend' question is labeled as a unipolar construct with the labels "not at all likely" to "extremely likely." However, it is not obvious whether one's likelihood to recommend is a unipolar construct in the minds of the respondents. Perhaps the likelihood continuum ranges from 'extremely unlikely' to 'extremely likely' rather than 'not at all likely?'. An example of a bipolar scale can be found in Table 1 (question version 6).

The polarity of the attitude has also been found to affect the optimal number of response options (Krosnick and Fabrigar, 1997). Whereas unipolar attitudes were found to be most validly

measured with five verbally labeled response options, bipolar attitudes were most validly measured with seven options. Question version 7 in Table 1 represents a unipolar version, and question version 9 represents a bipolar version—both fully verbally labeled.

Thus, the 'likelihood to recommend' question that Reichheld (2003) proposed ignored the advice from survey methodology research by only partially verbally labeling and using as many as eleven response options. Hence, the validity of the measurement may be increased by decreasing the number of scale points and fully verbally labeling them. In addition, the original version has its midpoint labeled '*neutral*.' For a respondent, it is, at best, ambiguous what this '*neutral*' option means. The kind of behavior expected from a respondent who responds '*neutral*' to a recommend intention is difficult to predict. Thus, adding labels to all response options and making it bipolar should greatly improve the measurement of the attitude by making it easier for the respondents to understand the response options.

In the surveys analyzed in this manuscript, every respondent was randomly assigned to answer one out of several versions of the 'likelihood to recommend' question. The versions varied in both the number and labeling of the response options. Table 2 presents the versions of the 'likelihood to recommend' question that respondents answered.

Data

Five survey-embedded experiments were administered in online questionnaires to American adults in order to evaluate the NPS and customer satisfaction metrics. The five questionnaires used similar experimental designs, randomly assigning different wordings of the 'likelihood to recommend' question and varying the number and wording of the response options. One questionnaire was administered during the first quarter of 2006, two questionnaires during the first quarter of 2008, one questionnaire during the third quarter of 2009, and one questionnaire

during the third quarter of 2015. Two of the questionnaires were administered to a probability-based sample of U.S. adults, and three of the questionnaires were administered to non-probability samples of U.S. adults.

Method

Measurements

Dependent variable *Relative revenue growth.* Although the NPS is supposed to predict revenue growth, Reichheld's (2003; 2006) dependent variable ranged between a prediction and a post-prediction by estimating the average growth rate two years prior to the survey to the year after the survey was conducted (Grisaffe, 2006). Instead of only post-predicting and to give the NPS a fair chance of predicting growth, the predictive ability of the NPS was evaluated on a range of growth periods and growth equations.

Five equations of revenue growth were predicted: (1) the original Reichheld (2003) equation that estimated average revenue growth from two years prior to the survey was administered to one year after the survey was administered, (2) two-year growth (the year the survey was administered to the year after the survey was administered), (3) three-year growth (the year before the survey was administered to the year after the survey was administered), (4) four-year growth (the year before the survey was administered to two years after the survey was administered), and (5) four-year growth (two years before the survey was administered to one year after the survey was administered). The relative revenue growth was estimated as the percent growth within each company (see Eq. 2.0 to Eq. 2.5).

Average (post-predicting) relative revenue growth % = (Eq. 2.0)

$$\left(\left(\frac{\frac{\text{Total revenue one year before the survey was conducted}}{\text{Total revenue two years before the survey was conducted}}+\frac{\text{Total revenue the year the survey was conducted}}{\text{Total revenue one year before the survey was conducted}}+\frac{\text{Total revenue one year after the survey was conducted}}{\text{Total revenue the year the survey was conducted}}}{3}\right)-1\right)*100$$

One year (predicting) relative revenue growth % = (Eq. 2.1)

$$\left(\left(\frac{\text{Total revenue the year after the survey was conducted}}{\text{Total revenue the year that the survey was conducted}}\right)-1\right)*100$$

Two years (post-predicting) relative revenue growth % = (Eq. 2.2)

$$\left(\left(\frac{\text{Total revenue the year after the survey was conducted}}{\text{Total revenue the year before the survey was conducted}}\right)-1\right)*100$$

Three years (post-predicting) relative revenue growth % = (Eq. 2.3)

$$\left(\left(\frac{\text{Total revenue two years after the survey was conducted}}{\text{Total revenue the year before the survey was conducted}}\right)-1\right)*100$$

Three years (post-predicting alternative) relative revenue growth % = (Eq. 2.4)

$$\left(\left(\frac{\text{Total revenue one year after the survey was conducted}}{\text{Total revenue two years before the survey was conducted}}\right)-1\right)*100$$

Three years (predicting) relative revenue growth % = (Eq. 2.5)

$$\left(\left(\frac{\text{Total revenue two years after the survey was conducted}}{\text{Total revenue the year the survey was conducted}}\right)-1\right)*100$$

The relative revenue for each company's last logged income statement for each year was extracted from the financial information database Capital IQ (2018). The companies in the five studies were linked to the public companies owning the brand names during 2004-2017. If a company's ownership changed during 2004-2017, the revenue of the company that owned the brand during the year that the survey was administered was selected. If ownership changed after the year that the survey was administered, the subsequent years were coded as missing. Revenue was reported in thousands of U.S. dollars using the value of one USD on April 28, 2018. Private companies that did not provide public income statements were excluded from the analyses (for a full list of the brands and the coding of their owners, see Table A1 in the appendix).

Companies with relative revenue growth or loss of more than three standard deviations larger than the average revenue growth of the companies in the datasets were excluded from all regressions (Figure A1, A2, and A3 in the appendix label the excluded observations). After

removing outliers, 53 observations from 30 companies in eight industries were included in the analyses.

Company evaluations. *NPS.* Four calculations of the NPS were compared: (1) The original NPS (see Eq. 1). The score ranged from 0 (everyone detracted) to 1 (everyone promoted). (2) A more restrictive NPS that subtracted the percentage of respondents who answered 0-6 on the original 11-point version from the percentage of respondents who answered 10. (3) An NPS that kept more of the variation in the original 11-point version by subtracting the percentage of respondents who answered 0-6 from the percentage of respondents who answered 7-10. (4) The full variation of the original 11-point 'likelihood to recommend' unipolar partially labeled question. All calculations were coded to range from 0 to 1 instead of -1 to 1 in order to allow comparisons between them.

Customer satisfaction measures. Furthermore, the NPS was compared to eight customer satisfaction measures: satisfaction with the company, liking of the company, the 'likelihood to recommend' question with eleven unipolar partially labeled response options, with five unipolar fully labeled response options, with seven bipolar fully labeled response options, or with seven unipolar partially labeled response options, the percent likelihood to recommend, and the number of times the respondent had recommended the company in the past.

Predictive validity

The customer satisfaction metrics were compared in terms of their predictive validity. Predictive validity refers to the rate that a specific measure or survey question predicts factors that it theoretically and empirically should predict (Ciuk and Jacoby, 2014; Krosnick and Fabrigar, 1997, p. 143; Lundmark et al., 2016; Shaeffer et al., 2005). Survey questions with non-optimal response options or a non-optimal NPS calculation would show weaker predictive validity compared to better measurements.

Parameters of OLS regressions were estimated, predicting the revenue growth with the company evaluation, the dummies for the experimental groups, and their interaction (see Shaeffer et al., 2005; Ciuk and Jacoby, 2014; or Lundmark et al., 2016). The estimated parameters for the interaction between the NPS and the dummies for the experimental groups gauged whether any of the versions were statistically significantly stronger predictors of revenue growth (see Eq. 3). Furthermore, in line with Reichheld's (2006) methodology, all analyses controlled for the industry of the company (see Keiningham et al., 2007), the year that the surveys were administered, and only respondents who answered that they were customers of the evaluated company were included.[4]

$$\begin{aligned} \textit{Relative revenue growth}_i = \textit{Company evaluation}_i + \textit{Experimental group}_i + \\ (\textit{Company evaluation}_i * \textit{Experimental group}_i) + \textit{Industry}_i + \textit{Year}_i + \varepsilon_i \end{aligned} \quad \text{(Eq. 3)}$$

A positive interaction parameter meant that the NPS was outperformed by the alternative version, whereas a negative interaction parameter meant that the NPS outperformed the alternative version (Dinesen, 2011, p. 172; Krosnick and Fabrigar, 1997, p. 143; Shaeffer et al., 2005).

Non-parametric prediction of revenue growth.

Bayesian Additive Regression Trees (BART) introduced by Chipman et al. (2010) were used to investigate NPS's prediction of revenue growth without assuming the functional form of the relationship. The BART framework is built around decision trees. Decision trees are simple non-parametric models that successively partition the data, producing predictions based on the partition in which an observation lies. Typically, each partitioning occurs along the values of a

[4] The regressions used the NPS as the reference category. Furthermore, the dependent variables were coded to expect a positive relationship (i.e., greater NPS should predict greater revenue growth). All regressions were estimated using cluster-robust standard errors with observations nested within companies.

single variable, but subsequent sub-partitioning of that partition can occur along the values of other variables or at other values along the same variable. As a result, decision trees can capture arbitrary relationships in the data, including non-linear relationships and interactions between variables.

BART is an ensemble method that improves predictive accuracy by averaging over numerous decision trees rather than relying on a single tree. If each tree is thought of as a possible relationship between the predictors and the outcome, averaging over these possible relationships according to their likelihood rather than choosing a single relationship. BART performs as well as or better than competing methods in a variety of contexts (e.g., Alves and Artes, 2018).

Recent work on BART has developed a formal approach to variable selection (Bleich et al., 2014). This offers a principled way to determine which predictors are likely to have a real effect on the outcome. In this manuscript, the "bartMachine" package for the statistical software R (Kapelner and Bleich, 2018) was used to implement the BART technique.

BART is particularly well-suited to investigate the NPS's prediction of revenue growth because BART minimizes the risk that the importance of a variable is missed due to misspecification of the functional form of the relationship. BART can fit arbitrary relationships between predictors and outcome variables. This includes relationships which are highly non-linear as well as interactive relationships in which the effect of one variable on the outcome is conditional on any number of other variables.[5]

5 This can be contrasted with generalized additive models (GAMs), which can accommodate arbitrary relationships between each individual relationship and the outcome but assume that these effects are additive.

Results

Replicating Reichheld: Did the NPS predict average relative revenue growth?

Contrary to Reichheld's (2003) findings, the original NPS equation did not statistically significantly positively predict the average relative revenue growth over a three-year period ($\beta_{NPS\ predicting\ Eq.\ 2.0.} = 0.15$, $p = 0.18$) (see Table 4, column 1). Hence, despite that the relative revenue growth was estimated in the same way as Reichheld (2003) did and several of the companies were identical to the companies he investigated or, at the very least, were from the same industries, the data could not support the hypotheses that the NPS predicts increases in relative revenue growth.[6]

However, it is plausible that the NPS can positively predict other relative revenue growth equations. At the very least, the NPS should be able to predict increases in growth during the periods temporally closest to when the NPS was recorded.

Did the NPS predict other relative revenue growth equations?

But, the NPS did not statistically significantly positively predict the growth for years temporally closest to the NPS measurement. That is, the relative revenue growth from the year of the survey to the year after the survey was not positively predicted by the NPS. The relationship between the NPS and relative revenue growth was actually directionally negative ($\beta_{NPS\ predicting\ Eq.\ 2.1.} = -0.15$, $p = 0.60$) (see Table 4, column 2). In fact, the NPS failed to statistically significantly predict all of the relative revenue growth predictions ($\beta_{NPS\ predicting\ Eq.\ 2.2.} = 0.14$, $p = 0.37$; β_{NPS}

[6] When not correcting for that revenues were confounded in company industries and years, the NPS marginally significantly predicted three-year revenue growth ($\beta_{NPS\ predicting\ Eq.\ 2.2.} = 0.26$, $p = 0.08$) (see Table A4, column 2). However, not correcting for the known confounds in the data might bias the statistical significance test of the prediction. Hence, the predictive validity of the NPS could not be confirmed when estimating the likely less biased parameters. The unweighted data returned smaller estimated parameters between the NPS and the percent revenue growths.

$_{predicting\ Eq.\ 2.3.}$ = -0.03, p = 0.90; $\beta_{NPS\ predicting\ Eq.\ 2.3.}$ = -0.11, p = 0.50) (see Table 4, column 3, 4, and 6).

The slight exception to the failed predictions was when estimating equation 2.4.—the relative revenue growth from two years before the survey to one year after the survey—greater NPS marginally significantly predicted increases in relative revenue growth ($\beta_{NPS\ predicting\ Eq.\ 2.4.}$ = 0.16, p = 0.10). Thus, according to these results, a company that had increased its NPS from -100 to +100 should expect a 16% increase in relative revenue growth, but only when comparing revenues between two years before the NPS was recorded to the year after the NPS was recorded. In none of the other possible growth periods should the company expect a positive relationship between their NPS and their relative revenue growth.

However, before concluding that the NPS failed to predict growth positively, the fact may be that the NPS shared a non-linear relationship with relative revenue growth, and estimating the relationship with an OLS regression equation unjustly forced the relationship to have a linear functional form.

DId the NPS have a non-linear relationship with revenue growth?

Two groups of BART models predicting the relative revenue growth predictions with the NPS were estimated. The first group used the NPS as a predictor along with the industry and year. The BART analyses revealed that whether the company was a car manufacturer, and what year the questionnaire was administered were important for predicting revenue growth, whereas NPS was not selected (see Table 5, rows 1 to 5).

In the second group of models, the NPS was replaced with eleven cumulative response frequencies for each of the eleven possible responses to the ‘likelihood to recommend’ question on which the NPS is based. That is, the fraction who responded a zero, the fraction who responded with one or below, the fraction who responded with two or below, and so on were

included in the BART regression. These models allow for the possibility that the responses to the question might be predicted by the 'likelihood to recommend' question but in a different form than the NPS (see Table 5, rows 6 to 10). In similarity to the models with the NPS, car manufacturing and year were selected as important variables in most of the models. However, none of the eleven cumulative response frequencies from the 'likelihood to recommend' question were selected.

The next logical step would thus be to test if the relationship between customers' likelihood to recommend and relative revenue growth can be restored by changing the boundaries (or cutoffs) for the NPS equation.

Did changing the cutoff points help?

However, the NPS's inability to predict relative growth could not be restored by changing the cutoffs of promoters and detractors. For almost all of the relative revenue growth predictions estimated in this manuscript, varying the cutoff points did not statistically significantly improve the predictions of the revenue growth time-periods (see Table 6, column 1 to 6). One exception to this was the prediction of the average revenue growth (Eq. 2.0.), where the 10 as the promoters' cutoff point was marginally significantly superior to the original NPS ($\beta_{\textit{NPS 10 as promoters * company evaluation – predicting Eq. 2.4.}} = 0.05$, $p = 0.10$) (see Table 6, column 1). However, the superior predictive power of the 10 as cutoff compared to the original NPS still did not reveal a statistically significant main effect of predicting increases in average revenue growth ($\beta_{\textit{NPS 10 as promoters – predicting Eq. 2.0.}} = 0.20$, $p = 0.12$) (see Table A3, column 1).

The original NPS calculation was, however, statistically significantly superior in predicting the relative growth from two years before to one year after the survey (Eq. 2.4.) when comparing it to an NPS where customers who answered between 7 and 10 were considered to be promoters and 0 to 6 were considered detractors ($\beta_{\textit{NPS 10-7 as promoters * company evaluation – predicting Eq. 2.4.}}$

= -0.05, p = 0.04) (see Table 6, column 5). But, when comparing the original NPS to cutoffs where the customers who answered 10 on the 'likelihood to recommend' question, the two NPS equations predictions of revenue growth (Eq. 2.4) were not statistically indistinguishable from each other (β *NPS 10 as promoters * company evaluation – predicting Eq. 2.4.* = 0.02, p = 0.32) (see Table 6, column 5).

Did other customer satisfaction predict relative growth?

Striving to uncover evidence for the superiority of the NPS, attempts were made to compare it to other customer satisfaction metrics. Nor here did the analyses reveal any support that the NPS were superior. The NPS did not predict any of the relative revenue growth equations more strongly than what customers' satisfaction with the companies (β *Satisfaction * company* evaluation – *predicting Eq. 2.0.* = 0.11, p = 0.28; β *Satisfaction * company* evaluation – *predicting Eq. 2.1.* = -0.22, p = 0.19; β *Satisfaction * company* evaluation – *predicting Eq. 2.2.* = 0.02, p = 0.87; β *Satisfaction * company* evaluation – *predicting Eq. 2.3.* = -0.08, p = 0.59; β *Satisfaction * company* evaluation – *predicting Eq. 2.4.* = 0.05, p = 0.49; β *Satisfaction * company* evaluation – *predicting Eq. 2.5.* = 0.09, p = 0.41), nor the customers' liking of the companies did (β *Liking * company* evaluation – *predicting Eq. 2.0* = 0.08, p = 0.29; β *Liking * company* evaluation – *predicting Eq. 2.1.* = -0.02, p = 0.86; β *Liking * company* evaluation – *predicting Eq. 2.2.* = 0.03, p = 0.72; β *Liking * company* evaluation – *predicting Eq. 2.3.* = 0.02, p = 0.85; β *Liking * company* evaluation – *predicting Eq. 2.4.* = 0.02, p = 0.70; β *Liking * company* evaluation – *predicting Eq. 2.5.* = 0.03, p = 0.67) (see Table 7). Thus, in similarity to the NPS, neither customers' satisfaction nor their liking of the companies statistically significantly predicted increases in relative revenue growth (see Table A4, columns 1 to 12).

Did improving the likelihood to recommend question help the prediction of relative revenue growth?

The full variation of the eleven unipolar partially labeled version did not outperform the NPS in terms of predictive validity (β *Eleven unipolar partially labeled * company* evaluation – *predicting Eq. 2.0.* = 0.04,

p = 0.45; β *Eleven unipolar partially labeled * company* evaluation – *predicting Eq. 2.1.* = -0.04, p = 0.73; β *Eleven unipolar partially labeled * company* evaluation – *predicting Eq. 2.2.* = 0.04, p = 0.66; β *Eleven unipolar partially labeled * company* evaluation – *predicting Eq. 2.3.* = -0.06, p = 0.55; β *Eleven unipolar partially labeled * company* evaluation – *predicting Eq. 2.4.* = 0.03, p = 0.48; β *Eleven unipolar partially labeled * company* evaluation – *predicting Eq. 2.5.* = -0.05, p = 0.33) (see Table 8, row 11).

Furthermore, improvements in the 'likelihood to recommend' survey question did not increase the positive prediction of relative revenue growth. Decreasing the number of response options to seven and changing the construct to bipolar (unlikely to likely to recommend) did not statistically significantly strengthen the prediction compared to the NPS for most of the growth equations (β *Seven bipolar partially labeled * company* evaluation – *predicting Eq. 2.1.* = -0.02, p = 0.94; β *Seven bipolar partially labeled * company* evaluation – *predicting Eq. 2.2.* = 0.10, p = 0.43; β *Seven bipolar partially labeled * company* evaluation – *predicting Eq. 2.3.* = 0.01, p = 0.98; β *Seven bipolar partially labeled * company* evaluation – *predicting Eq. 2.4.* = 0.11, p = 0.11) (see Table 8, row 12). However, for the average relative growth (Eq. 2.0), the seven bipolar partially labeled response options offered some positive findings in that it did outperform the NPS (β *Seven bipolar partially labeled * company* evaluation – *predicting Eq. 2.0.* = -0.02). But despite it outperforming the NPS, the seven bipolar partially labeled response options still did not statistically significantly predict average relative growth (β *Seven bipolar partially labeled.* = 0.42, p = 0.10) (see Table A5.1, column 7). This finding contradicts Keiningham and his colleagues (2007) finding where they found that a bipolar 'likelihood to recommend' survey question statistically significantly predicted relative revenue growth.

None of the other attempts to improve the survey question showed any indication of helping the prediction. Fully labeling the response options and reducing them to five did not statistically significantly increase the predictive validity (β *Five unipolar fully labeled * company* evaluation – *predicting Eq. 2.0.* = -0.01, p = 0.89; β *Five unipolar fully labeled * company* evaluation – *predicting Eq. 2.1.* = 0.02, p = 0.88;

β *Five unipolar fully labeled * company* evaluation – *predicting Eq. 2.2.* = -0.02, p = 0.87; β *Five unipolar fully labeled * company* evaluation – *predicting Eq. 2.3.* = -0.05, p = 0.73; β *Five unipolar fully labeled * company* evaluation – *predicting Eq. 2.4.* = -0.02, p = 0.78; β *Five unipolar fully labeled * company* evaluation – *predicting Eq. 2.5.* = -0.06, p = 0.36), nor did fully verbally labeling seven bipolar response options (β *Seven bipolar fully labeled * company* evaluation – *predicting Eq. 2.0.* = -0.10, p = 0.16; β *Seven bipolar fully labeled * company* evaluation – *predicting Eq. 2.1.* = -0.04, p = 0.87; β *Seven bipolar fully labeled * company* evaluation – *predicting Eq. 2.2.* = -0.10, p = 0.26; β *Seven bipolar fully labeled * company* evaluation – *predicting Eq. 2.3.* = -0.05, p = 0.84; β *Seven bipolar fully labeled * company* evaluation – *predicting Eq. 2.4.* = -0.07, p = 0.20; β *Seven bipolar fully labeled * company* evaluation – *predicting Eq. 2.5.* = 0.10, p = 0.35) (see Table 8, rows 13 and 14).

Were there any alternative 'likelihood to recommend' version that predicted relative revenue growth?

The last analyses of this manuscript investigated exploratory versions of measuring the likelihood to recommend. First, across every growth equation, asking the customers the percent likelihood to recommend a company did statistically significantly worse than the NPS in predicting relative revenue growth (β *Percent likelihood * company* evaluation – *predicting Eq. 2.0.* = -0.24, p = 0.03; β *Percent likelihood * company* evaluation – *predicting Eq. 2.1.* = -0.56, p = 0.01; β *Percent likelihood * company* evaluation – *predicting Eq. 2.2.* = -0.29, p = 0.01; β *Percent likelihood * company* evaluation – *predicting Eq. 2.3.* = -0.51, p = 0.01; β *Percent likelihood * company* evaluation – *predicting Eq. 2.4.* = -0.16, p = 0.03; β *Percent likelihood * company* evaluation – *predicting Eq. 2.5.* = -0.22, p = 0.06) (see Table 8, row 15). In fact, as the percent likelihood to recommend increased, the closer the company moved toward bankruptcy (see Table A5.3, column 1 to 6).

There is little, if any, theoretical reason to expect that this relationship is accurate. It is more likely that the relationship between the percent likelihood to recommend and revenue growth is a statistical artifact due to that the percent question being asked only in 2008, 2009, and 2015, and contributed with only 17 observations (see Table A5.3, columns 1 to 6).

Lastly, asking for the number of times a customer recommended a company in the past showed promising predictive power. Although the number of times did not statistically significantly outperform the NPS in predicting relative company revenue growth (β *Number of recommendations * company* evaluation – *predicting Eq. 2.0.* = -0.12, p = 0.21; β *Number of recommendations * company* evaluation – *predicting Eq. 2.1.* = 0.14, p = 0.44; β *Number of recommendations * company* evaluation – *predicting Eq. 2.2.* = -0.08, p = 0.46; β *Number of recommendations * company* evaluation – *predicting Eq. 2.3.* = 0.06, p = 0.74; β *Number of recommendations * company* evaluation – *predicting Eq. 2.4.* = -0.10, p = 0.13; β *Number of recommendations * company* evaluation – *predicting Eq. 2.5.* = 0.10, p = 0.35) (see Table 8, row 16), the number of times the customer recommended the company marginally significantly predicted three (β *Number of recommendations – predicting Eq. 2.1.* = 0.16, p < 0.01; β *Number of recommendations – predicting Eq. 2.3.* = 0.12, p = 0.09; β *Number of recommendations – predicting Eq. 2.5.* = 0.05, p = 0.06) out of the six relative growth equations (see Table A5.3, column 7 to 12).

Did including non-customers help?

Adding non-customers to the estimation of the NPS, the NPS using different cutoffs, other customer satisfaction metrics, and the improved likelihood to recommend questions did not affect the outcome of the analyses presented above. All analyses were rerun when also including evaluations of the companies made by non-customers in the five questionnaires. But, for each growth equation, the predictions exhibited similar flat predictions or even weaker predictions than when excluding non-customers.

Summary

In this manuscript, attempts were made to predict the relative revenue growth of thirty companies with their respective NPS's. Replicating the analyses in Reichheld (2003) and attempting to predict the relative revenue growth of the same companies and from the same

industries as Reichheld originally used, the NPS was not found to predict the average relative revenue growth.

Other relative growth equations were then tested to try to establish a link between the NPS and relative revenue growth. However, in five out the six growth equations investigated, no relationship with the NPS could be found. Only once did a significant relationship appear, and then only marginally so, and only when the relative revenue growth was not temporally close to the measurement of the NPS.

In attempts to rectify the non-predictive ability of the NPS, different cutoffs of promoters and detractors were estimated and used for predicting relative revenue growth. However, none of these alternative NPS estimations provided us with clearer evidence of a link between the score and relative revenue growth. Furthermore, allowing the full variation across the original eleven unipolar partially labeled 'likelihood to recommend' question neither increased nor decreased the predictive validity compared to the NPS.

Perhaps most discouraging for proponents of the NPS's ability to predict revenue growth was that improvements to the survey question (i.e., optimizing the number and the labeling of the response options) did not yield statistically significant improvements to the predictions. If a clear link between relative revenue growth and customers' likelihood to recommend existed, then one should expect that such a link would be strengthened by asking better-formatted questions.

Hence, the most likely conclusion is that, regardless of how well one measures the survey question, the customers' likelihood to recommend a company is not as strongly related to the company's revenue growth as Reichheld (2003) suggested.

The manuscript's contribution compared to previous research

The lack of relationship between the NPS and revenue growth in this manuscript adds to the literature contradicting the value of score's (Morgan and Rego, 2006; Keiningham et al., 2007;

Hayes, 2008). In similarity of Neil A. Morgan and Lopo Leotte Rego's (2006) analyses of 80 companies, the present manuscript found that the NPS was not predicting several types of growth metrics. However, in contrast to the present manuscript's contribution, Morgan and Rego (2006) did not use the correct 'likelihood to recommend' question to estimate the number of net promoters of the company.[7] Furthermore, Bob E. Hayes (2008) critique against the score was not based on the NPS, but instead of a composite measure of various likelihood to recommend items.

Furthermore, Timothy Keiningham and his colleagues (2007) studied 21 Norwegian firms and concluded that the NPS was positively related to revenue growth, albeit not statistically significantly so. However, the lack of relationship in their paper could have been due to the few companies they studied, or the fact that their respondents reported their likelihood to recommend on a 10-point scale instead of an 11-point scale. However, as the present manuscript has shown, using the correct survey question to form improving the measurement of the 'likelihood to recommend' question nor other customer satisfaction metrics predicted revenue growth.

Perhaps most damning to the NPS, Keiningham and his colleagues (2007) argued that the NPS was not even significantly outperforming other customer satisfaction metrics in Reichheld's (2003) and Satmetrix's (2004) own datasets. However, given that Reichheld (2003) and Satmetrix (2004) never made their full dataset publicly available, Keiningham and his colleagues (2007) could only conclude that, in the bits and pieces of data available through graphs, the NPS did not produce a statistically significant superior prediction of revenue growth compared to other metrics. It may well be that the full dataset in Reichheld (2003), in fact, did show the strong positive relationship between the NPS and revenue growth.

[7] Respondents in Morgan and Rego (2006) were asked "Have you discussed your experiences with [brand or company x] with anyone?" and "Have you formally or informally complained about your experiences with [brand or company x]?" (p. 429).

Thus, the overall picture that starts to appear through the peer-reviewed literature and through this manuscript is that there is no, or at the very best, a weak relationship between the NPS and relative revenue growth. In contrast to Reichheld's (2003; 2006) findings, something as multi-faceted as a company's revenue growth does not seem to be predicted by surveying a valid sample of one's customers and asking them their likelihood to recommend the company (Reichheld 2003, p. 1).

Limitations

The present manuscript presented 53 observation of companies. A potential reason for the lack of statistical significance presented in this manuscript might thus be due to its weak statistical power. However, although Reichheld (2003) claimed to have investigated over 400 companies, it seems that only about 50 of them were included in the actual analyses (Satmetrix, 2004, p. 6). Hence, the 53 observations analyzed in this manuscript should yield enough statistical power to expect statistically significant predictions of revenue growth.

Furthermore, some of the companies in the dataset had few self-reported customers. However, given the claim that the NPS should be considered the 'ultimate' customer satisfaction measure, we expected a strong relationship between the NPS and revenue growth with small residuals and large regression coefficients, none of which were found here. Additionally, Reichheld (2003; 2006) did not disclose the number of customers he had for each company to be considered a valid sample for predicting relative revenue growth.

Another limitation might be that three out of the five questionnaires in this manuscript were administered to non-probability samples of the population. Reichheld (2003) made no claims that probability samples of customers were required to predict revenue growth, and the data he used for his analyses were acquired through *"email invitations sent to addresses purchased from publicly available, opt-in email lists*" (Satmetrix, 2004, p. 3). However, from

previous research, we know that probability samples render vastly superior representation of populations' beliefs, attitudes, and behaviors (Yeager et al., 2011; MacInnis et al., 2018). Hence, compared to Reichheld's (2003; 2006) studies, a strength of the present manuscript was that two of the datasets were generated using a probability sample of respondents. Future studies should investigate whether a probability sample of customers can conjure a significant relationship between the NPS and relative revenue growth.

Market volatility and short-term trends in the economy might also limit our findings if the relationship between the NPS is strongly influenced by such trends. Three of the questionnaires were administered during one of the greatest economic recessions in U.S history. Hence, many of the companies experienced great losses. However, since we had several companies from the same industry, and controlled for the type of industries, we would expect the companies with a higher NPS would show less revenue loss compared to companies with a lower NPS. Such a relationship could not be found in our data. Therefore, it is unlikely that the recession of 2008 and 2009 caused the lack of predictive validity of the NPS.

Future research

A glimmer of hope in the failed replications presented in this manuscript was that a relationship appeared between the customers' number of past recommendations of the company and relative revenue growth. Although not statistically significantly outperforming the NPS, the number of past recommendations marginally significantly predicted three out of the six growth equations. Due to the exploratory nature of the analyses of the past recommendations and given that the question measuring past recommendations were always asked after the 'likelihood to recommend' question, it is too early to tell whether the number of past recommendations is a true predictor of relative revenue growth.

References

Alves de Brito Filho, D., & Artes, R. (2018). Application of Bayesian additive regression trees in the development of credit scoring models in Brazil. *Production*, 28: e20170110.

Bleich, J., Kapelner, A., George, E.I.. & Jensen., S.T. (2014). Variable selection for BART: An application to gene regulation. *Annals of Applied Statistics*, 8 (3): 1750–1781.

Capital IQ. (2018). *S&P Capital IQ Platform.* Retrieved May 2, 2018, from https://www.capitaliq.com/CIQDotNet/Login.aspx.

Ciuk, D. J., & Jacoby, W. G. (2015). Checking for Systematic Value Preferences Using the Method of Triads. *Political Psychology, 36*(6), 709-728.

Chipman, H.A., George, E.I., & McCulloch, R.E. (2010). BART: Bayesian additive regression trees. *Annals of Applied Statistics*, 4 (1): 266–298.

Dinesen, P. T. (2011). A Note on the Measurement of Generalized Trust of Immigrants and Natives. *Social Indicators Research, 103*, 169-177.

Hayes, B. E. (2008). The True Test of Loyalty. *Quality Progress*, *41*(6), 20–26.

Inc Magazine. (2006). *Would you recommend us? Perfect your service by asking the only question that matters.* Retrieved May 13, 2019, from http://www.inc.com/magazine/20060901/handson-customer-service.html.

Kapelner, A., & Bleich, J. (2016). bartMachine: Machine Learning with Bayesian Additive Regression Trees. *Journal of Statistical Software*, 70 (4), 1–40.

Keiningham, T. L., Cooil, B., Andreassen, T. W., & Aksoz, L. (2007). A Longitudinal Examination of Net Promoter and Firm Revenue Growth. *Journal of Marketing*, *71*, 39-51.

Krosnick, J. A. (1999). Survey Research. *Annual Review of Psychology*, *50*, 537-67.

Krosnick, J. A., & Leandre, F. R. (1997). Designing Rating Scales for Effective Measurement in Surveys. In *Survey Measurement and Process Quality*. New York: John Wiley & Sons.

Lundmark, S., Gilljam, M., & Dahlberg, S. (2016). How to Measure Generalized Trust: Wording and Number Of Scale Points. *Public Opinion Quarterly*. *80*(1), 26-43.

MacInnis, B., Krosnick, J. A., Ho, A. S., & Cho, M.-J. (2018). The accuracy of measurements with probability and nonprobability survey samples. *Public Opinion Quarterly*, *82*(4), 707-744. doi:10.1093/poq/nfy038.

Morgan & Rego (2006). The Value of Different Customer Satisfaction and Loyalty Metrics in Predicting Business Performance. *Marketing Science*. *25*(5), 426–439.

Reichheld, F. F. (2003). The One Number You Need. *Harvard Business Review*, *81*(12), 46-54.

Reichheld, F. F. (2006). *The Ultimate Question: Driving Good Profits and True Growth.* Boston: Harvard Business School Press.

Satmetrix. (2004). The Power Behind a Single Number: Growing Your Business with Net Promoter. *Satmetrix Systems white paper*. Retrieved April 15, 2019, from http://www.marketingritson.com/wp-content/uploads/2014/11/week1netpromoterscores.pdf].

Shaeffer, E. M., Krosnick J. A., Langer G. E., & Merkle, D. M. (2005). Comparing the Quality of Data Obtained By Minimally Balanced and Fully Balanced Attitude Questions. *Public Opinion Quarterly*, *69*, 417-28.

Tourangeau, R. (1984). Cognitive Sciences and Survey Methods. In *Cognitive Aspects* of S*urvey Methodology: Building a Bridge between Disciplines*, (pp. 73-100). Washington, DC: National Academy Press.

Yeager, D., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. *Public Opinion Quarterly, 75*(4), 709-747. https://doi.org/10.1093/poq/nfr020

Table 1. Examples of likelihood to recommend questions

Question Wording	Response options										
(1) – Standard version "How likely are you to recommend [company X] to a friend or a colleague?"	Not at all likely 0	1	2	3	4	Neutral 5	6	7	8	9	Extremely likely 10
(2) – Product version "How likely are you to recommend [flying/buying/eating] [with/a/at] [company X] to a friend or a colleague?"	Not at all likely 0	1	2	3	4	Neutral 5	6	7	8	9	Extremely likely 10
(3) – Recommend against "How likely are you to [not] recommend [company X] to a friend or a colleague?"	Not at all likely 0	1	2	3	4	Neutral 5	6	7	8	9	Extremely likely 10
(4) – Natural metric: suppose/imagine "[Suppose/Imagine] you are talking about cell phone companies with a friend or colleague. Out of one hundred such conversations that you might have with friends or colleagues, how many times would you recommend Verizon?"	[text-box] (type a number between zero and 100)										

Continued

Table 1. *Continued*

Question Wording	Response options
(5) – Natural metric: Percent "What is the percent chance out of 100 that you will recommend [company] to a friend or colleague?"	Please type a number between 0 and 100, with the higher the number meaning it's more likely.
(6) – Standard version 11-point bipolar	Extremely unlikely 0 1 2 3 4 5 6 7 8 9 Extremely likely 10
(7) – Standard version unipolar 5-point	○ Not at all likely ○ Slightly likely ○ Moderately likely ○ Very likely ○ Extremely likely
(8) – Standard version unipolar 7-point	○ Not at all likely ○ Slightly likely ○ Moderately likely ○ Fairly likely ○ Pretty likely ○ Very likely ○ Extremely likely

Continued

Table 1. *Continued*

Question Wording	Response options
(9) – Standard version bipolar fully labeled 7-point	○ Extremely likely ○ Moderately likely ○ Slightly likely ○ Neither likely nor unlikely ○ Slightly unlikely ○ Moderately unlikely ○ Extremely unlikely
(10) – Recommend against bipolar fully labeled 7-point	○ Extremely likely to recommend against ○ Moderately likely to recommend against ○ Slightly likely to recommend against ○ Neither likely to recommend nor recommend against ○ Slightly likely to recommend ○ Moderately likely to recommend ○ Extremely likely to recommend
(11) – Standard version unipolar partially labeled 7-point	Not at all likely — Neutral — Extremely likely 1 2 3 4 5 6 7

Table 2. Question versions in each Study.

Study 1. 32 Companies.

How likely is it that you would recommend each of the following companies to a friend or colleague?

Group 1 Original	Not at all likely					Neutral					Extremely likely
	0	1	2	3	4	5	6	7	8	9	10
Group 2 7 unipolar partially labeled	Not at all likely					Neutral					Extremely likely
	1		2		3	4		5		6	7
Group 3 7 unipolar fully labeled	Not at all likely		Slightly likely		Somewhat likely	Likely		Very likely		Remarkably likely	Extremely likely
Group 4 5 unipolar fully labeled	Not at all likely			Slightly likely		Moderately likely			Very likely		Extremely likely

Continued

Table 2. Continued.

Study 2. 16 Companies.
How likely is it that you would recommend [company product] to a friend or colleague?

Group 1 Original	Not at all likely 0	1	2	3	4	Neutral 5	6	7	8	9	Extremely likely 10
Group 2 7 unipolar partially labeled	Not at all likely 1		2	3		Neutral 4		5	6		Extremely likely 7
Group 3 7 unipolar fully labeled	Not at all likely		Slightly likely	Moderately likely		Fairly likely		Pretty likely	Very likely		Extremely likely
Group 4 5 unipolar fully labeled	Not at all likely		Slightly likely			Moderately likely		Very likely			Extremely likely

Comment: *Not included in the analysis due to too few companies evaluated. *Continued*

Table 2. Continued.

Study 3: Wave 5. Three Companies.
How likely is it that you would recommend [company product] to a friend or colleague?

Group 1 Original	Not at all likely				Neutral					Extremely likely
	0	1	2	3	4	5	6	7	8	9	10
Group 2 7 bipolar fully labeled	Extremely unlikely		Moderately unlikely		Slightly unlikely	Neither likely nor unlikely		Slightly likely	Moderately likely		Extremely likely
Group 3 Percentage	"What is the percent chance out of 100 that you will recommend [PRODUCT] by [COMPANY] a friend or colleague? Please type a number between 0 and 100, with the higher the number meaning it's more likely."										

Study 3: Wave 12. Fifteen Companies
How likely is it that you would recommend [company product] to a friend or colleague?

Group 1 Original (reversed order)	Extremely likely				Neutral					Not at all likely
	0	1	2	3	4	5	6	7	8	9	10
Group 2 7 bipolar fully labeled	Extremely likely		Moderately likely		Slightly likely	Neither likely nor unlikely		Slightly unlikely	Moderately unlikely		Extremely unlikely
Group 3 Percentage	"Please type a number between 0 and 100, with the higher the number meaning it's more likely."										

Continued

Table 2. Continued.

Study 4. Three Companies.
How likely are you to recommend [company] to a friend or colleague?

Group 1 Original	Not at all likely 0 1 2 3 4 Neutral 5 6 7 8 9 Extremely likely 10
Group 2 5 fully labeled	Not at all likely; Slightly likely; Moderately likely; Very likely; Extremely likely
Group 3 7 bipolar fully labeled	Extremely likely; Moderately likely; Slightly likely; Neither likely nor unlikely; Slightly unlikely; Moderately unlikely; Extremely unlikely
Group 4 Percentage	"What is the percent chance out of 100 that you will recommend [company] to a friend or colleague? Please type a number between 0 and 100, with the higher the number meaning it's more likely."

Source: Study 1, 2, 3, and 4.

Table 3. Companies and their number of customers in Study 1, 2, 3, and 4.

	Number of Customers				
Company	Study 1	Study 2	Study 3: Wave 5	Study 3: Wave 12	Study 4
Alaska Airlines	126	-	-	-	-
Albertsons Supermarkets	606	-	-	-	-
America West Airlines	217	-	-	-	-
American Airlines	459	387	609	258	403
An-Jan Feed and Pet Supply	51	-	-	-	-
Avis Rent a Car	266	-	-	-	-
Best Buy	1,316	-	-	-	-
BMW	-	20*	-	-	-
Budget Rent a Car	398	-	-	-	-
Chevrolet	-	252	-	318	-
Circuit City	1,014	-	-	-	-
CompUSA	540	-	-	-	-
Continental Airlines	320	280	-	171	-
Delta Airlines	488	376	-	215	-
Dodge	-	158	-	204	-
Drug Emporium	96	-	-	-	-
Enterprise Rent-A-Car	652	-	-	-	-
Ford	-	255	806	339	444
Hertz Rent a Car	345	-	-	-	-
Home Depot	1,669	-	-	-	-
Honda	-	90	-	174	-
Jet Blue	-	74	-	72	-
Long's Pharmacy	229	-	-	-	-
Lowes Home Improvement	1,490	-	-	-	-
McDonalds	-	-	960	-	-
National Car Rental	232	-	-	-	-
Northwest Airlines	288	239	-	139	-
Orchard Supply Hardware	141	-	-	-	-
P W Supermarkets	59*	-	-	-	-
Petco Pet Store	691	-	-	-	-
PetSmart Pet Store	848	-	-	-	-
Rite Aid	813	-	-	-	-
Safeway Food & Drug	458	-	-	-	-
Save Mart Supermarkets	105	-	-	-	-
Southwest Airlines	492	306	-	247	-
Toyota	-	138	-	253	-
United Airlines	405	345	-	217	-
United Drugs	63	-	-	-	-
U.S. Airways	324	209	-	169	-
Walgreen pharmacies	1,448	-	-	-	-
Verizon	-	-	-	-	360
Volkswagen	-	100	-	145	-
Volvo	-	22*	-	51	-

Comment: *Excluded from analyses due to too fewer than 10 customers on average per likelihood to recommend question version.

Table 4. Predicting the percent revenue growth with the NPS.

	Relative revenue growth					
	Eq. 2.0. Average growth two years before to one year after the survey	Eq. 2.1. The year of the survey to one year after the survey	Eq. 2.2. One year before the survey to one year after the survey	Eq. 2.3. The year before the survey to two years after the survey	Eq. 2.4. Two years before the survey to one year after the survey	Eq. 2.5. The year of the survey to two years after the survey
Company evaluation						
NPS	0.17	-0.15	0.14	-0.03	0.16+	-0.11
	(0.18)	(0.60)	(0.37)	(0.90)	(0.10)	(0.50)
Company industries						
Automobile manufacturer	-0.09+	-0.06	-0.10+	-0.17*	-0.06	-0.06
	(0.09)	(0.36)	(0.07)	(0.02)	(0.10)	(0.13)
Computer retail	-0.07	0.09	-0.02	0.01	-0.03	0.04
	(0.15)	(0.21)	(0.72)	(0.93)	(0.31)	(0.26)
Pharmacy	-0.05	0.04	-0.04	0.00	-0.03	0.05
	(0.20)	(0.41)	(0.41)	(0.96)	(0.29)	(0.18)
Food retail	-0.06**	0.00	-0.07*	-0.09*	-0.05*	-0.02
	(0.01)	(0.99)	(0.03)	(0.01)	(0.02)	(0.36)
Home improvement	-0.02	0.03	-0.04	-0.06	-0.01	-0.02
	(0.67)	(0.66)	(0.56)	(0.45)	(0.74)	(0.62)
Restaurants	0.02	0.07	0.02	-0.08	0.02	-0.03
	(0.61)	(0.40)	(0.73)	(0.36)	(0.51)	(0.39)
Pet stores	-0.01	0.06	-0.01	0.05	-0.00	0.06**
	(0.80)	(0.37)	(0.83)	(0.22)	(0.98)	(0.00)
Telecommunications	-0.03	0.02	-0.04	-0.11**	-0.02	-0.04*
	(0.27)	(0.53)	(0.18)	(0.01)	(0.40)	(0.05)
Year						
2008	-0.07	-0.25**	-0.11*	-0.08	-0.05	-0.05
	(0.11)	(0.00)	(0.04)	(0.15)	(0.13)	(0.16)
2009	-0.04	0.09	-0.07	0.00	-0.03	0.07*
	(0.39)	(0.20)	(0.21)	(0.96)	(0.37)	(0.04)
2015	-0.00	-0.11	-0.10+	-0.13+	-0.08*	-0.06+
	(0.93)	(0.14)	(0.09)	(0.07)	(0.03)	(0.05)
Constant	0.18**	0.48**	0.27**	0.36**	0.11*	0.23**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Adjusted R^2	0.20	0.42	0.27	0.23	0.21	0.27
Observations	53	53	53	50	53	50

Notes. Unstandardized regression coefficients from six OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.
$+p < 0.1$ $*p < 0.05$ $**p < 0.01$.

Table 5: Important predictors in BART models of percent revenue growth.

NPS	Important predictors
Eq. 2.1. The year of the survey to one year after the survey	Year = 2008 Year = 2009 Industry = Automobile manufacturer
Eq. 2.2. One year before the survey to one year after the survey	Year = 2006 Industry = Automobile manufacturer
Eq. 2.3. The year before the survey to two years after the survey	Industry = Automobile manufacturer
Eq. 2.4. Two years before the survey to one year after the survey	Year = 2006 Industry= Automobile manufacturer
Frequency of likelihood to recommend	
Eq. 2.1. The year of the survey to one year after the survey	Year = 2008 Year = 2009
Eq. 2.2. One year before the survey to one year after the survey	Year = 2006 Industry = Automobile manufacturer
Eq. 2.3. The year before the survey to two years after the survey	Industry = Automobile manufacturer
Eq. 2.4. Two years before the survey to one year after the survey	Year = 2006 Industry = Automobile manufacturer

Table 6. Predicting the percent revenue growth with different NPS calculations.

	Relative revenue growth					
	Eq. 2.0. Average growth two years before to one year after the survey	Eq. 2.1. The year of the survey to one year after the survey	Eq. 2.2. One year before the survey to one year after the survey	Eq. 2.3. The year before the survey to two years after the survey	Eq. 2.4. Two years before the survey to one year after the survey	Eq. 2.5. The year of the survey to two years after the survey
Company evaluation						
Net score	0.16 (0.17)	-0.15 (0.54)	0.13 (0.32)	-0.03 (0.88)	0.15+ (0.08)	-0.11 (0.44)
Experiment group						
NPS 10 Promoters - 0-6 Detractors	-0.01 (0.13)	-0.00 (0.97)	-0.00 (0.85)	0.00 (0.68)	-0.00 (0.62)	0.00 (0.93)
NPS: 10-7 Promoters, 0-6 Detractors	-0.01 (0.54)	0.00 (0.93)	0.01 (0.32)	0.00 (0.77)	0.01 (0.12)	-0.00 (0.86)
Interactions						
NPS 10 Promoters - 0-6 Detractors **x** Net score	0.05+ (0.10)	-0.01 (0.84)	0.01 (0.60)	-0.01 (0.74)	0.02 (0.32)	-0.01 (0.71)
NPS: 10-7 Promoters, 0-6 Detractors **x** Net score	-0.03 (0.38)	0.04 (0.62)	-0.05 (0.26)	0.00 (0.95)	-0.05* (0.04)	0.03 (0.44)
Company industries						
Automobile manufacturer	-0.08+ (0.09)	-0.07 (0.28)	-0.10+ (0.05)	-0.18* (0.01)	-0.06+ (0.09)	-0.06 (0.11)
Computer retail	-0.05 (0.29)	0.13* (0.02)	-0.00 (0.96)	0.02 (0.68)	-0.03 (0.43)	0.05+ (0.06)
Pharmacy	-0.05 (0.22)	0.05 (0.29)	-0.04 (0.41)	0.01 (0.89)	-0.03 (0.26)	0.05 (0.12)
Food retail	-0.06* (0.01)	0.01 (0.71)	-0.07* (0.03)	-0.09* (0.01)	-0.05* (0.02)	-0.01 (0.45)
Home improvement	-0.01 (0.76)	0.03 (0.54)	-0.03 (0.59)	-0.06 (0.43)	-0.01 (0.81)	-0.02 (0.60)
Restaurants	0.02 (0.53)	0.07 (0.40)	0.02 (0.75)	-0.08 (0.29)	0.02 (0.52)	-0.03 (0.33)
Pet stores	-0.00 (1.00)	0.06 (0.27)	-0.00 (0.90)	0.05 (0.14)	0.00 (0.92)	0.06** (0.00)
Telecommunications	-0.02 (0.34)	0.03 (0.34)	-0.04 (0.18)	-0.10** (0.00)	-0.02 (0.39)	-0.03* (0.05)
Year						
2008	-0.07 (0.10)	-0.23** (0.00)	-0.11* (0.03)	-0.07 (0.17)	-0.05 (0.12)	-0.04 (0.15)
2009	-0.04 (0.41)	0.10 (0.11)	-0.07 (0.23)	0.01 (0.85)	-0.03 (0.37)	0.07* (0.02)
2015	-0.01 (0.89)	-0.11 (0.16)	-0.10+ (0.09)	-0.13+ (0.08)	-0.08* (0.02)	-0.06+ (0.06)
Constant	0.18** (0.00)	0.47** (0.00)	0.27** (0.00)	0.36** (0.00)	0.11** (0.01)	0.23** (0.00)

Adjusted R^2	0.31	0.51	0.38	0.35	0.32	0.38
Observations	153	153	153	144	153	144

Notes. Unstandardized regression coefficients from six OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.
$+p < 0.1$ $*p < 0.05$ $**p < 0.01$.

Table 7. Predicting the percent revenue growth with different customer satisfaction measures.

	Relative revenue growth					
	Eq. 2.0. Average growth two years before to one year after the survey	Eq. 2.1. The year of the survey to one year after the survey	Eq. 2.2. One year before the survey to one year after the survey	Eq. 2.3. The year before the survey to two years after the survey	Eq. 2.4. Two years before the survey to one year after the survey	Eq. 2.5. The year of the survey to two years after the survey
Company evaluation						
Customer satisfaction metrics	0.13 (0.20)	-0.14 (0.52)	0.11 (0.35)	-0.03 (0.89)	0.13 (0.10)	0.16 (0.17)
Experiment group (NPS as reference)						
Satisfaction with the company	-0.12 (0.22)	0.20 (0.27)	-0.05 (0.66)	0.06 (0.69)	-0.08 (0.29)	-0.11 (0.29)
Liking of the company	-0.10 (0.21)	0.06 (0.64)	-0.06 (0.51)	-0.00 (0.98)	-0.06 (0.32)	-0.08 (0.36)
Interactions						
Satisfaction with the company **x** Customer satisfaction metrics	0.11 (0.28)	-0.22 (0.19)	0.02 (0.87)	-0.08 (0.59)	0.05 (0.49)	0.09 (0.41)
Liking of the company **x** Customer satisfaction metrics	0.08 (0.29)	-0.02 (0.86)	0.03 (0.72)	0.02 (0.85)	0.02 (0.70)	0.03 (0.67)
Company industries						
Automobile manufacturer	-0.09+ (0.09)	-0.07 (0.31)	-0.10* (0.05)	-0.18* (0.01)	-0.06+ (0.09)	-0.07+ (0.07)
Computer retail	-0.04 (0.34)	0.13* (0.02)	-0.00 (0.99)	0.02 (0.68)	-0.02 (0.50)	-0.11 (0.25)
Pharmacy	-0.05 (0.19)	0.05 (0.27)	-0.04 (0.38)	0.01 (0.88)	-0.03 (0.23)	-0.12 (0.20)
Food retail	-0.07** (0.01)	0.01 (0.53)	-0.07* (0.02)	-0.09* (0.01)	-0.05* (0.01)	-0.14 (0.14)
Home improvement	-0.01 (0.78)	0.03 (0.50)	-0.03 (0.62)	-0.06 (0.42)	-0.01 (0.86)	-0.10 (0.33)
Restaurants	0.01 (0.70)	0.07 (0.34)	0.01 (0.87)	-0.08 (0.24)	0.01 (0.71)	0.01 (0.70)
Pet stores	-0.00 (0.85)	0.07 (0.25)	-0.01 (0.84)	0.05 (0.15)	-0.00 (0.96)	-0.09 (0.34)
Telecommunications	-0.02 (0.29)	0.03 (0.30)	-0.04 (0.16)	-0.10** (0.00)	-0.02 (0.33)	-0.11 (0.25)
Year						
2008	-0.08+ (0.05)	-0.23** (0.00)	-0.12* (0.02)	-0.07 (0.18)	-0.06+ (0.06)	-0.14 (0.12)
2009	-0.05 (0.25)	0.11*	-0.08	0.02	-0.04	-0.13

		(0.04)	(0.14)	(0.82)	(0.21)	(0.19)
2015	-0.01 (0.82)	-0.10	-0.11+	-0.13+	-0.08*	-0.14*
		(0.16)	(0.08)	(0.08)	(0.02)	(0.05)
Constant	0.20**	0.46**	0.28**	0.35**	0.13**	0.20*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.04)
Adjusted R^2	0.30	0.51	0.38	0.35	0.31	0.15
Observations	153	153	153	144	153	150

Notes. Unstandardized regression coefficients from six OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.
$+p < 0.1$ $*p < 0.05$ $**p < 0.01$.

Table 8. Predicting the percent revenue growth with the different likelihood to recommend versions.

	Relative revenue growth					
	Eq. 2.0. Average growth two years before to one year after the survey	Eq. 2.1. The year of the survey to one year after the survey	Eq. 2.2. One year before the survey to one year after the survey	Eq. 2.3. The year before the survey to two years after the survey	Eq. 2.4. Two years before the survey to one year after the survey	Eq. 2.5. The year of the survey to two years after the survey
Company evaluation						
Likelihood to recommend score	0.13 (0.15)	-0.11 (0.54)	0.10 (0.38)	-0.04 (0.83)	0.11 (0.12)	-0.09 (0.41)
Experiment group (NPS as reference)						
Eleven unipolar partially labeled	-0.06 (0.24)	0.06 (0.62)	-0.05 (0.52)	0.05 (0.65)	-0.05 (0.23)	0.06 (0.33)
Seven bipolar partially labeled	-0.20* (0.02)	0.06 (0.77)	-0.09 (0.36)	0.02 (0.92)	-0.11+ (0.08)	0.13 (0.17)
Five unipolar fully labeled	-0.01 (0.81)	0.02 (0.87)	-0.00 (0.94)	0.05 (0.59)	-0.01 (0.82)	0.06 (0.20)
Seven bipolar fully labeled	0.01 (0.86)	0.05 (0.79)	0.02 (0.80)	-0.02 (0.86)	-0.00 (0.99)	-0.04 (0.65)
Percent	0.08 (0.10)	0.27* (0.03)	0.11+ (0.06)	0.23* (0.01)	0.05 (0.13)	0.11+ (0.07)
Number of recommendations	0.05 (0.17)	-0.05 (0.49)	0.03 (0.42)	-0.02 (0.78)	0.04 (0.12)	-0.04 (0.38)
Interactions						
Eleven unipolar partially labeled **x** Likelihood to recommend	0.04 (0.45)	-0.04 (0.73)	0.04 (0.66)	-0.06 (0.55)	0.03 (0.48)	-0.05 (0.33)
Seven bipolar partially labeled **x** Likelihood to recommend	0.23* (0.02)	-0.02 (0.94)	0.10 (0.43)	0.01 (0.98)	0.11 (0.11)	-0.14 (0.23)
Five unipolar fully labeled **x** Likelihood to recommend	-0.01 (0.89)	0.02 (0.88)	-0.02 (0.87)	-0.05 (0.73)	-0.02 (0.78)	-0.06 (0.36)
Seven bipolar fully labeled **x** Likelihood to recommend	-0.10 (0.16)	-0.04 (0.87)	-0.10 (0.26)	0.03 (0.84)	-0.07 (0.20)	0.10 (0.44)
Percent **x** Likelihood to recommend	-0.24* (0.03)	-0.56* (0.01)	-0.29* (0.01)	-0.51* (0.01)	-0.16* (0.03)	-0.22+ (0.06)

Number of Recommendations **x** Likelihood to recommend	-0.12 (0.21)	0.14 (0.44)	-0.08 (0.46)	0.06 (0.74)	-0.10 (0.13)	0.10 (0.35)

Continued.

Table 8. Continued.

	Relative revenue growth					
	Eq. 2.0. Average growth two years before to one year after the survey	Eq. 2.1. The year of the survey to one year after the survey	Eq. 2.2. One year before the survey to one year after the survey	Eq. 2.3. The year before the survey to two years after the survey	Eq. 2.4. Two years before the survey to one year after the survey	Eq. 2.5. The year of the survey to two years after the survey
Company industries						
Automobile manufacturer	-0.08+ (0.09)	-0.06 (0.27)	-0.09+ (0.08)	-0.17* (0.02)	-0.05 (0.16)	-0.06+ (0.08)
Computer retail	-0.06 (0.12)	0.08 (0.21)	-0.01 (0.88)	0.01 (0.91)	-0.02 (0.47)	0.03 (0.25)
Pharmacy	-0.06 (0.11)	0.04 (0.30)	-0.04 (0.32)	0.00 (0.94)	-0.04 (0.19)	0.05+ (0.09)
Food retail	-0.07** (0.00)	0.00 (0.94)	-0.07* (0.02)	-0.09** (0.01)	-0.05* (0.01)	-0.01 (0.35)
Home improvement	-0.01 (0.73)	0.02 (0.71)	-0.03 (0.61)	-0.06 (0.38)	-0.00 (0.91)	-0.02 (0.50)
Restaurants	0.01 (0.77)	0.06 (0.43)	0.01 (0.84)	-0.09 (0.25)	0.01 (0.64)	-0.03 (0.31)
Pet stores	-0.01 (0.74)	0.06 (0.30)	-0.01 (0.83)	0.05 (0.14)	-0.00 (1.00)	0.06** (0.00)
Telecommunications	-0.03 (0.20)	0.02 (0.43)	-0.05 (0.16)	-0.11** (0.00)	-0.02 (0.34)	-0.03* (0.03)
Year						
2008	-0.08* (0.04)	-0.23** (0.00)	-0.13* (0.01)	-0.08 (0.12)	-0.06* (0.04)	-0.04+ (0.06)
2009	-0.05 (0.25)	0.11* (0.02)	-0.08 (0.14)	0.01 (0.87)	-0.04 (0.20)	0.08** (0.00)
2015	0.01 (0.89)	-0.10 (0.17)	-0.09 (0.10)	-0.12+ (0.07)	-0.07* (0.03)	-0.06* (0.04)
Constant	0.20** (0.00)	0.46** (0.00)	0.29** (0.00)	0.36** (0.00)	0.13** (0.00)	0.22** (0.00)
Adjusted R^2	0.32	0.51	0.38	0.38	0.32	0.41
Observations	268	268	268	253	268	253

Notes. Unstandardized regression coefficients from six OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.
$+p < 0.1$ $*p < 0.05$ $**p < 0.01$.

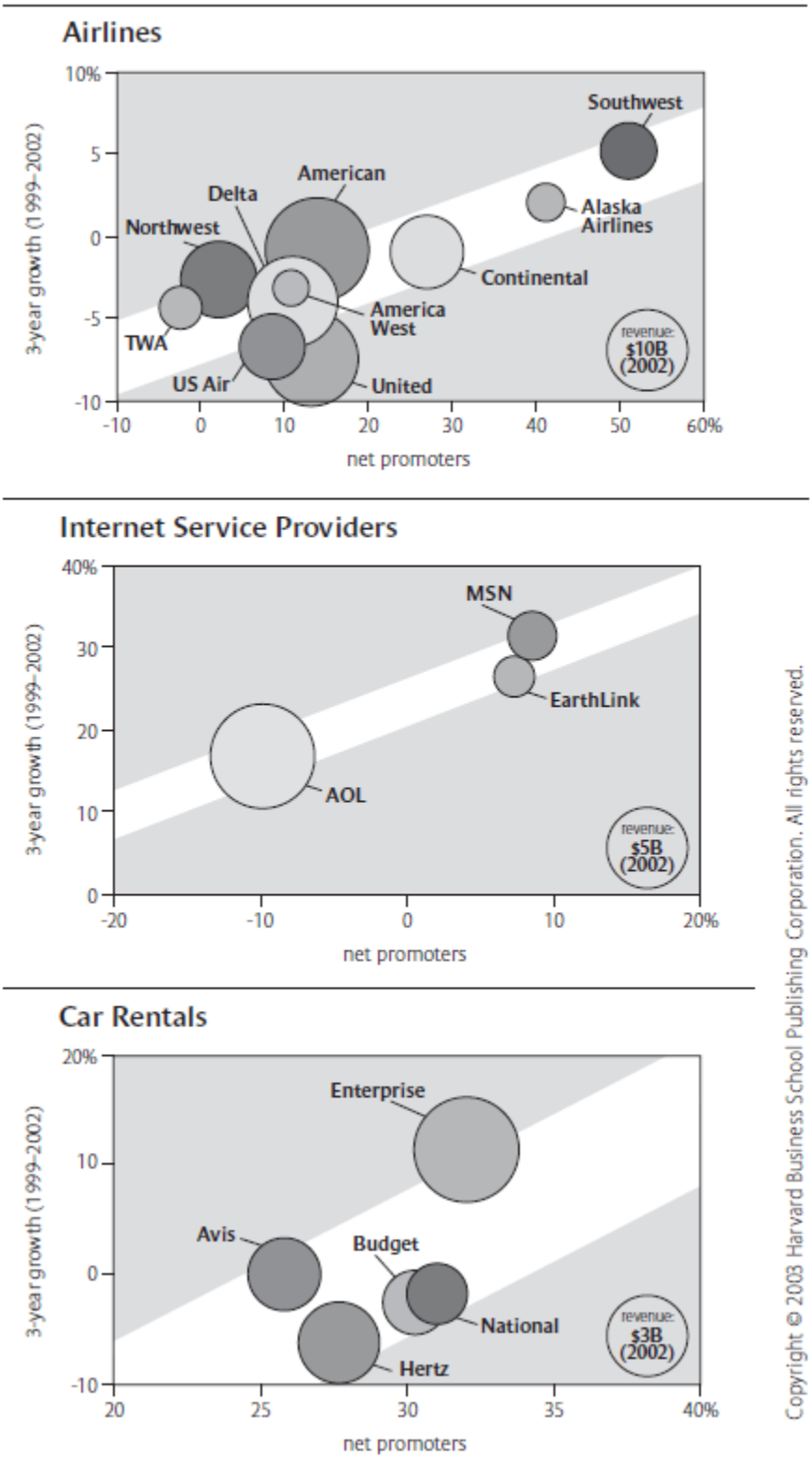

Figure 1. Graph called 'Growth by Word of Mouth' published in Reichheld (2003, p. 7).

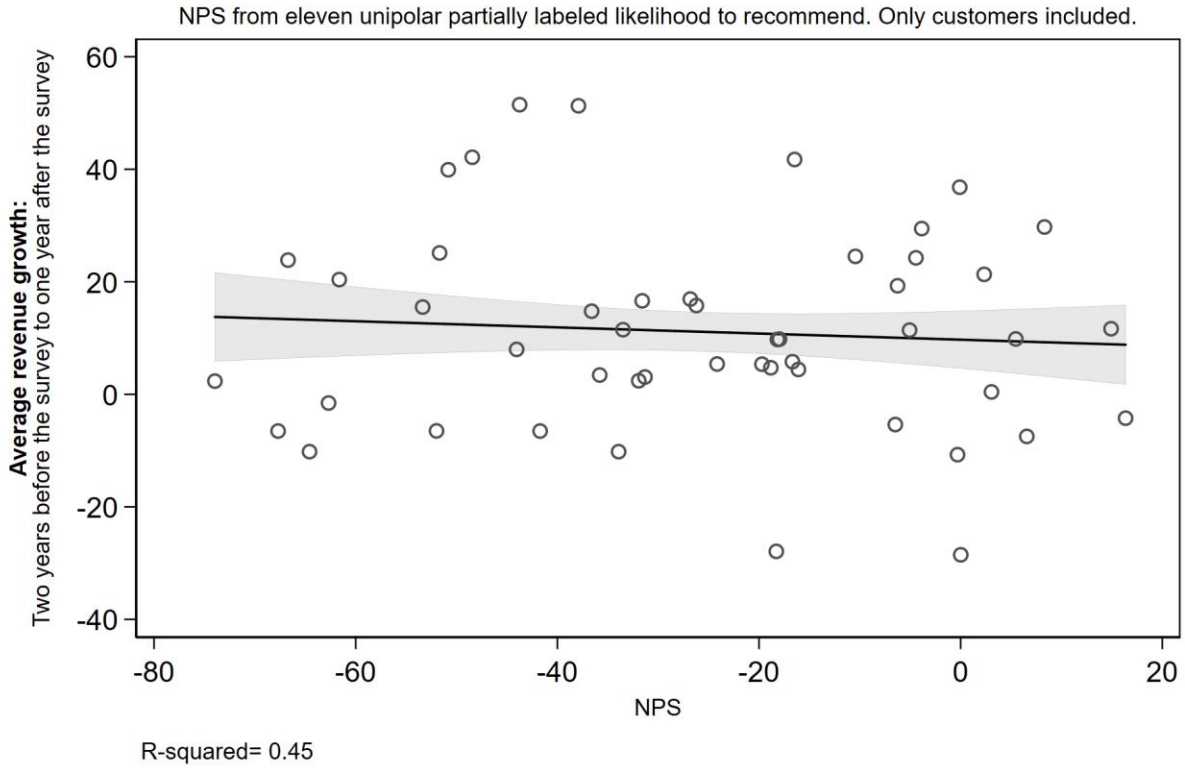

Figure 1. Scatterplot of average percent revenue growth from two years before the survey to one year after survey and NPS (Eq. 2.0). Fitted regression line predicting predicted values from a regression equation predicting percent revenue growth from the year of the survey to one year after the survey with the NPS from eleven unipolar partially labeled likelihood to recommend, the company industries, and the year the questionnaire was administered. Outliers in percent revenue growth above three standard deviations from the mean are excluded. Study 1, 2, 3, and 4.

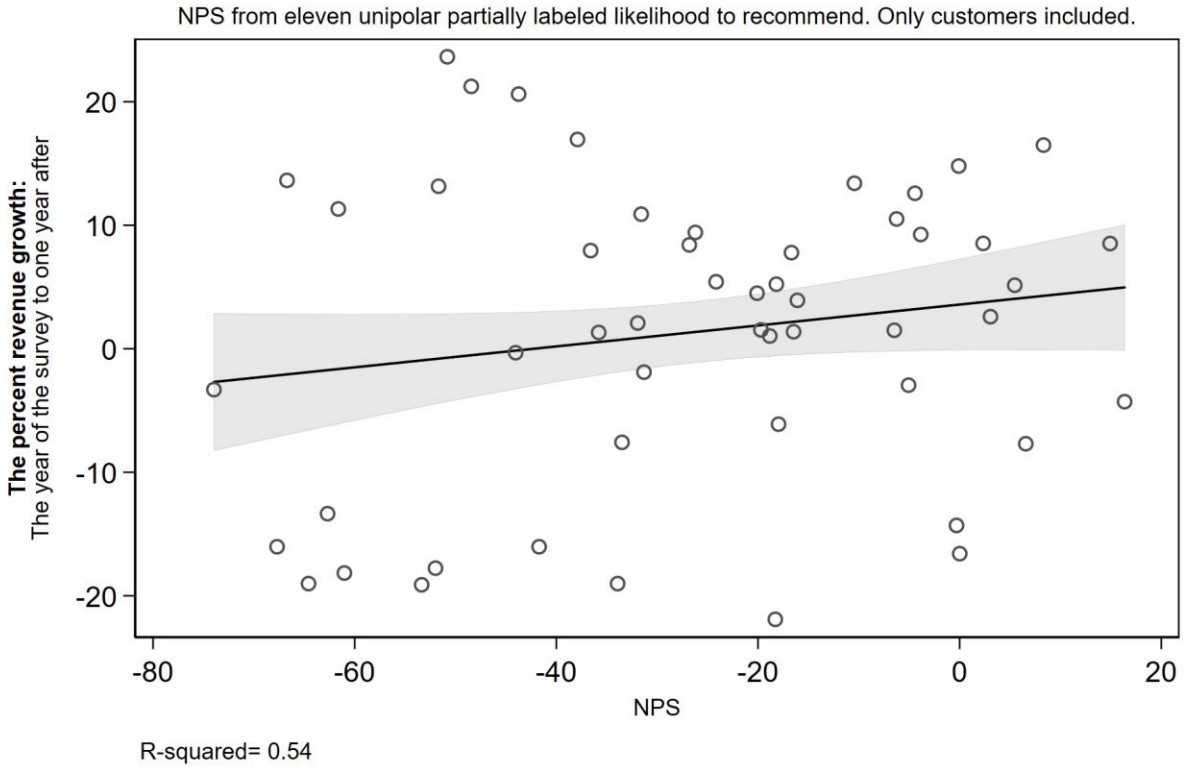

Figure 2. Scatterplot of percent revenue growth from the year of the survey to one year after survey and NPS (Eq. 2.1). Fitted regression line predicting predicted values from a regression equation predicting percent revenue growth from the year of the survey to one year after the survey with the NPS from eleven unipolar partially labeled likelihood to recommend, the company industries, and the year the questionnaire was administered. Outliers in percent revenue growth above three standard deviations from the mean are excluded. Study 1, 2, 3, and 4.

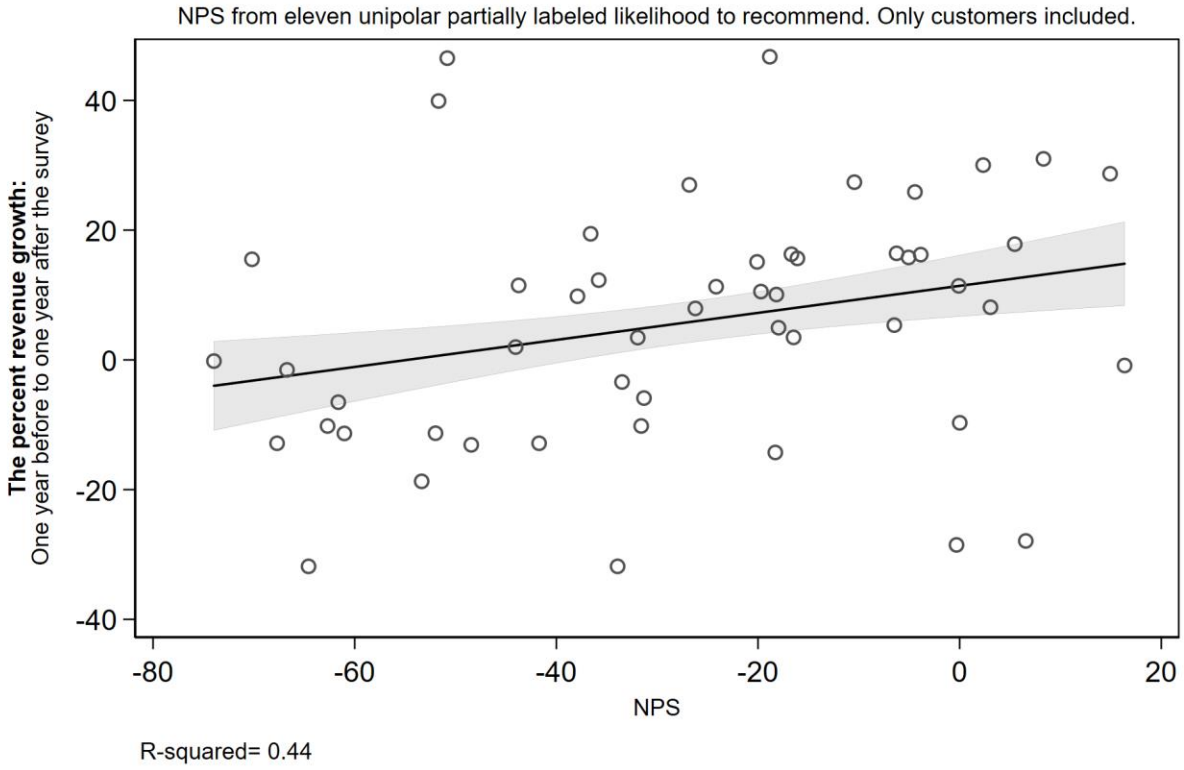

Figure 3. Scatterplot of percent revenue growth from the year before the survey to one year after the survey and NPS (Eq. 2.2). Fitted regression line predicting predicted values from a regression equation predicting percent revenue growth from the year before the survey to one year after the survey with the NPS from eleven unipolar partially labeled likelihood to recommend, the company industries, and the year the questionnaire was administered. Outliers in percent revenue growth above three standard deviations from the mean are excluded. Study 1, 2, 3, and 4.

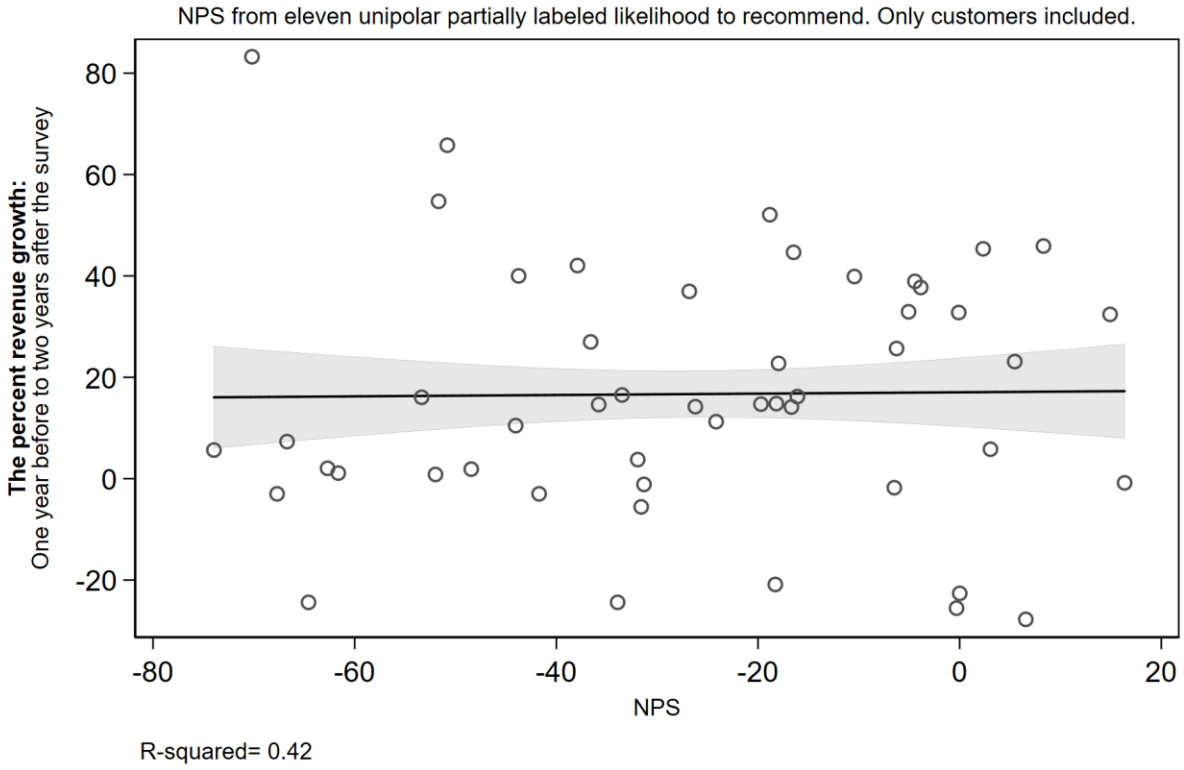

Figure 4. Scatterplot of percent revenue growth from the year before the survey to two years after the survey and NPS (Eq. 2.3). Fitted regression line predicting predicted values from a regression equation predicting percent revenue growth from the year before the survey to two years after the survey with the NPS from eleven unipolar partially labeled likelihood to recommend, the company industries, and the year the questionnaire was administered. Outliers in percent revenue growth above three standard deviations from the mean are excluded. Study 1, 2, 3, and 4.

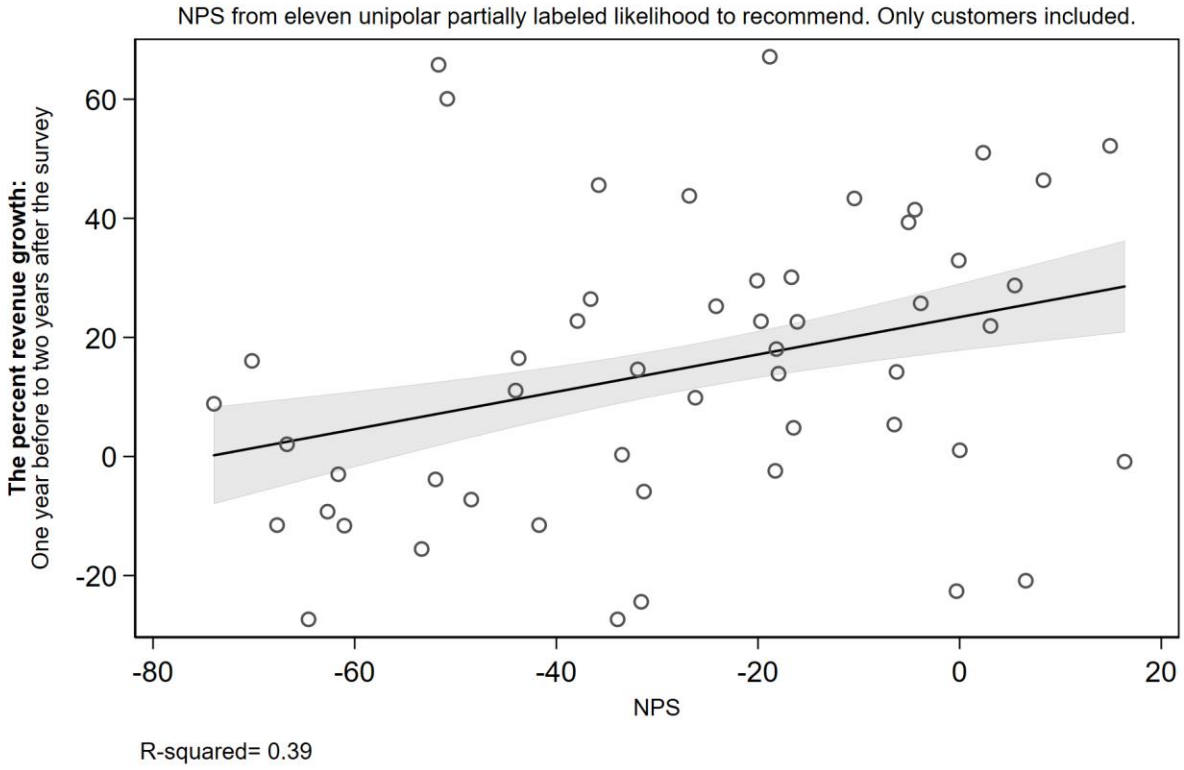

Figure 5. Scatterplot of percent revenue growth from two years before the survey to one year after the survey and NPS (Eq. 2.4). Fitted regression line predicting predicted values from a regression equation predicting percent revenue growth from two years before the survey to one year after the survey with the NPS from eleven unipolar partially labeled likelihood to recommend, the company industries, and the year the questionnaire was administered. Outliers in percent revenue growth above three standard deviations from the mean are excluded. Study 1, 2, 3, and 4.

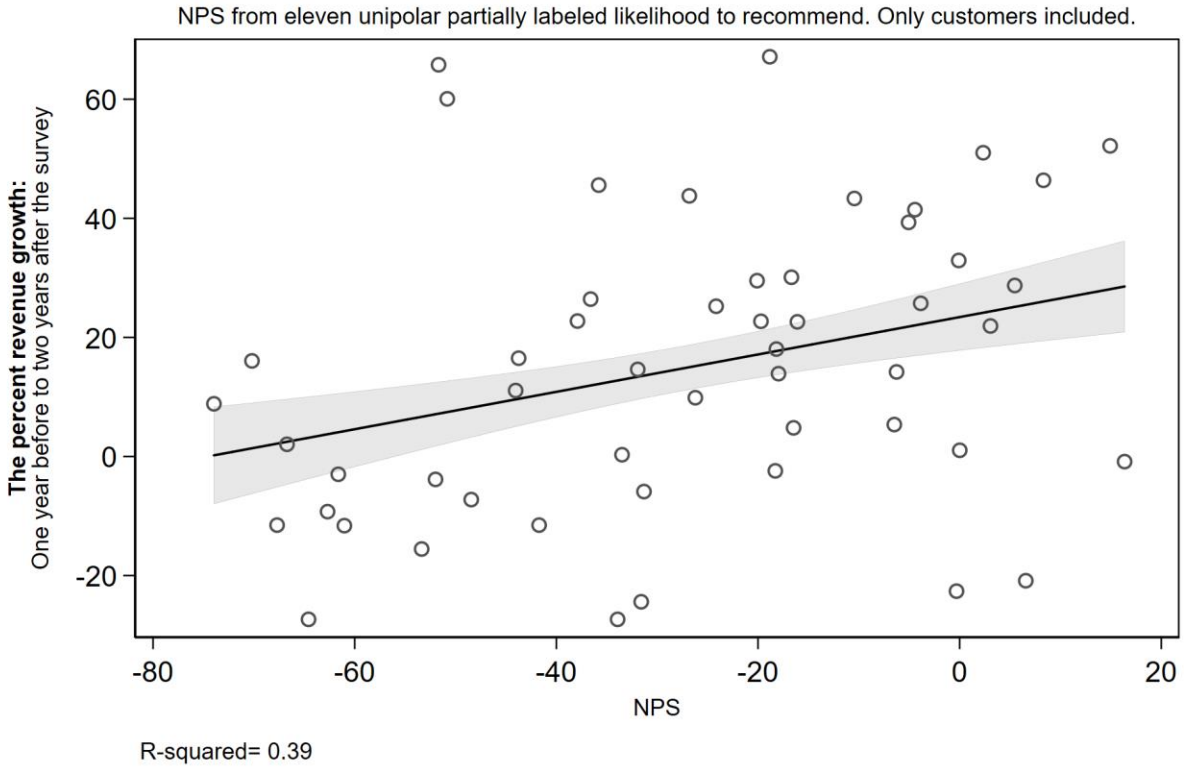

Figure 6. Scatterplot of percent revenue growth from the year of the survey to two years after the survey and NPS (Eq. 2.5). Fitted regression line predicting predicted values from a regression equation predicting percent revenue growth from two years before the survey to one year after the survey with the NPS from eleven unipolar partially labeled likelihood to recommend, the company industries, and the year the questionnaire was administered. Outliers in percent revenue growth above three standard deviations from the mean are excluded. Study 1, 2, 3, and 4.

Appendix - Data

Study 1. 32 companies dataset

Study 1 was administered to panelists of Lightspeed GMI's opt-in sample in February of 2006. Lightspeed's respondent pool was recruited through several methods including co-registration (the practice of referring leads concurrent with another registration process), and web banner ads on websites. Recruited participants were sent e-mails and electronic newsletters soliciting participation in the online surveys. Organizations sent emails to people on their mailing lists encouraging them to visit the Lightspeed GMI's enrollment webpage. Panel members encouraged other people to go to Lightspeed GMI's enrollment webpage. Lightspeed Research advertised on general topic websites with broad appeal and on special interest sites, which created a diversity of profiles and provided the ability to target-recruit certain demographics when required. People who clicked on the ads were taken to Lightspeed GMI's webpage to enroll in the panel. People using search engines might have seen a link to Lightspeed GMI's enrollment webpage.

People who successfully enrolled completed the following steps:

• The person completed a panel registration form, including his/her contact and demographic information;

• The person's postal code passed a validation check, which compared the postal code to the postal address files;

• A confirmation email was sent to the potential panel member who must click on it to complete his/her registration;

• The potential panel member agreed to be contacted by Lightspeed GMI's by email and/or SMS and agreed to the Lightspeed GMI's terms and conditions.

Lightspeed GMI's ensured the quality of panelists through the following automated and manual quality checks:

• Automatically prevented duplicate email addresses from registering a second time;

• Automatically required the person to type letters that were displayed in a distorted image to prevent "bots" and auto-scripts from joining;

• Automatically checked a panelist's age so that it fits within a designated range per country/culture (i.e., for the U.S., the minimum age to join the panel was 13);

• Automatically validated that the combination of the last name, zip code, birthdate, and country was unique and that no other panel member had it;

• Automatically ensured that all new panelists provided at least the following information: country, language, first and last name, e-mail, gender, password, username, birthdate, and zip code;

• Automatically checked that a panelist entered a postal code format according to the country's definition;

• Automatically checked that the IP address corresponded to the appropriate country;

• Automatically prevented a user from having multiple accounts based on a combination of factors (name, zip code, country, IP address, etc.);

• Manually removed new panelists with suspicious patterns of responses, poor quality responses, fraudulent registrations, or any other abusive and fraudulent behavior to the community.

The member terms that all members of Lightspeed GMI's panels were required to agree to specify that participation in any survey was voluntary. A survey invitation to panel members included information regarding how long the survey was expected to take, the number of points to be awarded for qualifying and completing the survey, and a topic for the survey.

In total, 2,277 respondents completed the questionnaire and were sampled using quotas for age, gender, and region to reflect the U.S. adult population according to the U.S Census Bureau's Current Population Survey. The data were weighted using Pasek's (2015) ANESrake raking algorithm to match the Current Population Survey 2006 – March Supplement statistics of age, sex, education, and household income.

Each respondent evaluated 32 different companies and was randomly assigned to receive one out of four versions of the 'likelihood to recommend' question.[8] The four versions were: eleven unipolar partially labeled response options (suggested by Reichheld, 2003), seven unipolar partially labeled response options, seven unipolar fully labeled response options, and five unipolar fully labeled response options. Respondents also reported how much they liked each company and how satisfied they were with each company.[9] The sizes of the four groups varied between 558 and 578 respondents.

The companies in Study 1 spanned six industries: airlines, computer retail, pharmacies, food retail, home retail, pet stores, and car rental companies (see Table 3 for the names of the companies).

Study 2. Sixteen companies dataset

Study 2 was administered between January 23, 2008, and February 8, 2008, with members of the Harris Interactive Online Panel, using a quota sampling strategy based on age, sex, region, income, education, and ethnicity. The Harris Interactive Online Panel had over 6

[8] Each experimental version of the 'likelihood to recommend' question is annotated with a number within a parenthesis, with each number representing one of the version (see Table A2 for the exact wording and response options).

[9] *Satisfaction:* "Overall, how satisfied are you with each of the following companies?" with eleven bipolar partially labeled response options numbered from 0 to 10 with verbal labels on 0 "Extremely dissatisfied," 5 "Neutral," and 10 "Extremely satisfied." *Liking:* "How much do you like or dislike each of the following companies?" with seven bipolar fully labeled response options: Dislike a great deal, dislike a moderate amount, dislike a little, neither like nor dislike, like a little, like a moderate amount, like a great deal.

million self-selected members who were recruited through various websites and online panel enrollment campaigns. In total, 28,089 respondents were invited via e-mail to a password-protected web-based survey on political and consumer issues. Respondents were sent one reminder. Of the invited respondents, 4,883 started the survey, and 4,326 completed the survey (88.6% completion rate). The data were weighted using raking to match the Current Population Survey 2007 – March Supplement statistics of age, sex, region, race-ethnicity, and income.

The respondents were randomly assigned to answer one out of five versions of the 'likelihood to recommend' question: eleven unipolar partially labeled response options, seven unipolar partially labeled response options, seven unipolar fully labeled response options, and five unipolar fully labeled response options[10]. The group sizes varied between 649 and 749 respondents. Respondents also reported how much they liked and how satisfied they were with each company.[11]

Prior to the 'likelihood to recommend question,' the respondents were randomly assigned to report if they were familiar with eight automobile-manufacturers or eight airline companies. Respondents were then randomly assigned to evaluate one of the brands with which they were familiar with. If the respondent was familiar with only one brand, he/she evaluated that brand, and if no brand was familiar, he/she did not evaluate any brands (see Table 3 for the company names).

[10] Respondents in this group were also randomly assigned to reported their likelihood to recommend against the company either before or after reporting their likelihood to recommend the company. The order did not affect the distribution to the likelihood to recommend.

[11] *Satisfaction:* "How dissatisfied or satisfied were you with the [PRODUCT] by [COMPANY]?" with seven bipolar fully labeled response options: Extremely dissatisfied, moderately dissatisfied, slightly dissatisfied, neither dissatisfied nor satisfied, slightly satisfied, moderately satisfied, and extremely satisfied.
Liking: "How much do you like or dislike [PRODUCT] by [COMPANY]?" with seven bipolar fully labeled response options: Dislike a great deal, dislike a moderate amount, dislike a little, neither like or dislike, like a little, like a moderate amount, like a great deal.

Study 3. Three companies dataset and fifteen companies dataset

Study 3 was part of the Face-to-Face Recruited Internet Survey Project (FFRISP) panel conducted during the years of 2008 to 2009. The FFRISP panelists were a probability sample of the American population of 18 years or older. A representative area probability sample of English-speaking adult residents of households in the United States area was drawn using a multi-stage design. The sample was built using the University of Michigan Survey Research Center's (SRC) national sampling frame of counties for PSU (primary sampling unit) selection. All phases of fieldwork were managed by Abt SRBI.

During the summer of 2008, interviewers visited randomly selected households, randomly selected one adult resident per household and recruited 1,000 such people to answer 12 monthly 30-minute questionnaires online. All potential panel members were offered a free laptop computer in exchange for agreeing to complete the 12 questionnaires. People who already owned a computer were offered an alternative cash incentive of $200 paid up-front, followed by up to four quarterly payments of $75 as long as they completed the monthly questionnaires. Participants who did not have high-speed Internet service in their homes (26%) were given high-speed Internet service at no cost to them. A total of 47% (AAPOR RR4) agreed to participate in the FFRISP panel.

For all surveys, an individualized, PIN-embedded survey link was created for each respondent. Once a survey was ready to launch, each respondent was sent an e-mail invitation with the link. Another way respondents could access the survey link was by logging into the NSP website. For reminder follow-ups, respondents were contacted via different modes in a tiered approach. The first choice for mode of contact was e-mail. If after multiple e-mails, a respondent could not be reached to complete a monthly survey, Abt SRBI attempted to reach him or her by phone for up to two weeks. Finally, a postal letter was sent to respondents who had not

completed a monthly survey and had not yet been reached by e-mail or phone. At the end of the project, most enrolled members were still part of the panel and were, therefore, being invited to participate, and more than 90% of panelists completed all 12 monthly surveys.

The 'likelihood to recommend' questions were included in the fifth (February 2008) and twelfth (September 2009) waves of the panel.

Study 3: Wave 5. In the fifth wave, 971 of the 991 invited respondents (97% participation rate, 47% RR1) evaluated three companies in a three-by-two full factorial experimental design. The respondents were randomly assigned to one out of three versions of the 'likelihood to recommend' question: eleven unipolar partially labeled response options, seven bipolar fully labeled response options, and the percent chance of recommending the company.

Each group was randomly assigned to receive the question with a reversed order of the response options (i.e., 'Extremely likely' as the first response option instead of the last). However, the order of the response options did not statistically significantly affect the distributions and will thus be collapsed into one group per 'likelihood to recommend' version. The group sizes varied between 233 and 255 respondents. Respondents also reported how much they liked each company and how satisfied they were with each company.[12] The data were weighted using post-stratification weights provided by FFRISP.

[12] *Satisfaction:* "Are you satisfied with [PRODUCT BY COMPANY], dissatisfied with it, or neither satisfied nor dissatisfied?" with three bipolar fully labeled response options: satisfied, neither dissatisfied nor satisfied, dissatisfied. Respondents who answered 'satisfied' answered "Are you extremely satisfied, moderately satisfied, or slightly satisfied with [PRODUCT BY COMPANY]?", and respondents who answered 'dissatisfied' answered "Are you extremely dissatisfied, moderately dissatisfied, or slightly dissatisfied with [PRODUCT BY COMPANY]?" *Liking:* "Overall, do you like [PRODUCT BY COMPANY], dislike it, or neither like nor dislike it?" with three bipolar fully labeled response options: like, dislike, and neither like nor dislike. Respondents who answered 'like' answered, "Do you like [PRODUCT BY COMPANY] a great deal, a moderate amount, or a little?", and respondents who answered 'dislike' answered "Do you dislike [PRODUCT BY COMPANY] a great deal, a moderate amount, or a little?"

Study 3: Wave 12. In the twelfth wave of the FFRISP, 904 out of 979 invited respondents (92% participation rate, 43% RR1) reported their familiarity with eight airline companies or with seven automobile manufacturers (see Table 3). Respondents were randomly assigned to evaluate one of the brands he or she reported familiarity with, and respondents not familiar with any of the brands were not asked the 'likelihood to recommend' question. The respondents were then randomly assigned to the eleven unipolar partially labeled response options,[13] seven bipolar fully labeled response options, and the percent chance of recommending the company. The group sizes varied between 217 and 229 respondents. Respondents also reported how much they liked each company and how satisfied they were with each company.[14] The data were weighted using post-stratification weights provided by FFRISP.

Study 4. Three Companies Dataset

Study 4 was administered from August 20 to August 27, 2015. The participants were a non-probability sample of American adults aged 18 or older. Participants were drawn from commercial online panels owned by Toluna or integrated into Toluna's network of panels (which Toluna refers to as integrated partner panels). The panels included people living in all U.S. states and territories. The difference between panelists from a Toluna-owned panel and an integrated partner panel is that the integrated partner, rather than Toluna themselves, invited the participant and paid the participant incentives for completing a questionnaire. The Toluna-owned panels were comprised of over 500,000 members in the United States. These panels and Toluna's integrated partner panels consisted of convenience samples of individuals who elected to

[13] Respondents who were assigned to the eleven-point version were randomly assigned to see the response options in the order 'extremely likely' (10) to 'not at all likely' (0) from left to right on the screen or 'not at all likely' (0) to 'extremely likely' (10) from left to right on the screen. The order of the response options did not affect the distribution of answers.

[14] ***Satisfaction:*** The same wording as in Wave 5, but response options presented in the order satisfied, neither satisfied nor dissatisfied, and dissatisfied. ***Liking:*** The same wording as in Wave 5, but response options presented in the order like, neither like nor dislike, and dislike.

participate in surveys in exchange for points, which they may exchange for gift cards from retail merchants, for cash, to enter raffles, for gift cards, or for products through Toluna's website or through the website of Toluna's integrated partner.

Toluna-owned panels and integrated partner panels recruited panel members by placing ads on websites; people who clicked on the ads were taken to a Toluna webpage to enroll in the panel. People using search engines might see a link to that Toluna enrollment webpage. Panel members encouraged other people to go to that enrollment webpage. Organizations sent emails to people on their mailing lists, encouraging them to visit the Toluna enrollment webpage. And people saw posts on social media, attracting them to the enrollment webpage.

People who successfully enrolled completed the following steps:

- The person completed a panel registration form, including his/her contact and demographic information;
- The person's postal code passed a validation check, which compared the postal code to the postal address files;
- A confirmation email was sent to the potential panel member who must click on it to complete his/her registration;
- The potential panel member agreed to be contacted by Toluna by email and/or SMS and agreed to the Toluna terms and conditions.

Toluna ensured the quality of panelists through the following automated and manual quality checks:

- Automatically prevented duplicate email addresses from registering a second time;
- Automatically required the person to type letters that were displayed in a distorted image to prevent "bots" and auto-scripts from joining;

• Automatically checked a panelist's age so that it fits within a designated range per country/culture (i.e., for the U.S., the minimum age to join the panel was 13);

• Automatically validated that the combination of the last name, zip code, birthdate, and country was unique and that no other panel member had it;

• Automatically ensured that all new panelists provided at least the following information: country, language, first and last name, e-mail, gender, password, username, birthdate, and zip code;

• Automatically checked that a panelist entered a postal code format according to the country's definition;

• Automatically checked that the IP address corresponded to the appropriate country;

• Automatically prevented a user from having multiple accounts based on a combination of factors (name, zip code, country, IP address, etc.);

• Manually removed new panelists with suspicious patterns of responses, poor quality responses, fraudulent registrations, or any other abusive and fraudulent behavior to the community.

The member terms state that all members of Toluna's panel or Toluna's integrated partner panels were required to agree to specify that participation in any survey was voluntary. A survey invitation to panel members included information regarding how long the survey was expected to take, the number of points to be awarded for qualifying and completing the survey, and a topic for the survey.

Participants were invited based on the profiling information we provided. Toluna employed sample stratification in drawing the sample for the study: six sex and age groups (males 18-34, 35-54, 55+ and females 18-34, 35-54, 55+), 4 region groups (Northeast, Midwest, South, West), ethnicity (Hispanic/Latino or not), and 7 race groups (Asian, Black or African-

Americans, Native American or Alaska Native, Native Hawaiian or other Pacific Island, other ethnicity, multi-racial, and White). Sample members were selected from the members of Toluna-owned panels and Toluna's integrated partner panels.

Participants were invited by seeing a link to the questionnaire on the panel member's homepage on a Toluna owned panel website or a website of Toluna's integrated partners, by being directed to our questionnaire through a survey router that redirects participants to our questionnaire after they had completed another questionnaire, or by email.

For panelists to earn points, they had to complete the questionnaire and pass the following attention check question, which was placed toward the end of the survey:

"To help us be sure that your computer is working properly with ours, please select 'Slightly disagree' below." with seven response choices presented vertically Strongly agree, Agree, Slightly agree, Neither agree nor disagree, slightly disagree, Disagree, and Strongly disagree.

The number of panelists who completed the questionnaire and successfully passed the attention check was 1,208. The total number of people who were invited to complete the questionnaire was 29,440, which equated to a participation rate of 4.1%.

The data were weighted using Pasek's (2015) ANESrake raking algorithm to match the Current Population Survey 2015 – March Supplement statistics of age, sex, education, and household income.

The respondents evaluated three companies and were randomly assigned to one out of four versions of the 'likelihood to recommend' question. The four versions included the eleven unipolar partially labeled response options, five unipolar fully labeled response options, seven bipolar fully labeled response options, and the percent chance of recommending the company.

The group sizes varied between 188 and 220 respondents. Respondents did not report how satisfied nor how much they liked each company.

Appendix - Measurements.

Measurements.

Dependent variable *The number of recommendations.* Respondents in Study 1 answered, "During the last 6 months, how many times did you recommend each of these companies?" Participants in Study 2 answered, "During the past 2 years, how many times did you recommend [buying a car made by / flying on] [company] to a friend or a colleague?" Respondents in Study 3 Wave 5 answered, "During the PAST 2 years, how many times did you recommend [Company's main product]?" Respondents in Study 3 Wave 12 answered, "During the PAST 2 years, how many times did you recommend [buying a car made by / flying on] [company]?" Respondents in Study 4 answered, "During the PAST [2 / 5] years, how many times did you recommend [company] to a friend or a colleague? Respondents in Study 4 reported the times they recommended the company over a 2-year period cell phone and airline companies, and 5 years for the automobile manufacturers.[15]

The self-reported numbers of recommendations were coded to range from 0 to 10, with all responses between 10 and 100 truncated at 10 in order to decrease the impact outliers had on the regression coefficients. Recommendations above 100 were coded as missing.

[15] The assessment of validity for the number of recommendations assumed that likelihood of recommending a company and the reported actual recommendations should be moderately correlated. A more valid 'likelihood to recommend' version will therefore produce stronger concurrent validity with the number of self-reported number of recommendations.

Appendix – Figures and Tables.

Table A1. Total revenue, percent revenue growth, detractors, passive, promoters, and NPS for each company and survey year. Study 1, 2, 3, and 4.

	Total revenue in millions, U.S. dollars						Percent growth				Likelihood to recommend			
Company and survey year	Two years before the survey	Year before survey	Year of the survey	Year after the survey	Two years after the survey	Average growth (Eq. 2.0)	2-year (Eq. 2.1)	3-year (Eq. 2.2)	4-year (Eq. 2.3)	4-year (Eq. 2.4)	Detractors	Passive	Promoters	NPS
Alaska Airlines 2006	2724	2975	3334	3506	3663	9%	5%	18%	23%	29%	27%	41%	32%	5
Albertsons Supermarkets 2006	35019	35019	39810	Became private 2007		-	1%	15%	-	-	-	30%	29%	-12
America West Airlines 2006	2757	5069	11557	11700	12118	71%*	1%	131%*	139%*	324%	39%	36%	25%	-14
American Airlines 2006	18608	20657	22490	22833	23696	7%	2%	11%	15%	23%	42%	35%	22%	-20
American Airlines 2008	22490	22833	23696	19898	22150	-4%	-16%	-13%	-3%	-12%	53%	35%	12%	-42
American Airlines 2008	22490	22833	23696	19898	22150	-4%	-16%	-13%	-3%	-12%	79%	9%	12%	-68
American Airlines 2009	22833	23696	19898	22150	23957	-0%	11%	-7%	1%	-3%	71%	19%	9%	-62
American Airlines 2015	26701	42676	40938	40163	42195	18%	-2%	-6%	-1%	50%	55%	21%	24%	-31
An-Jan Feed and Pet Supply 2006			Private			-	-	-	-	-	54%	8%	39%	-
Avis Rent a Car 2006	4780	5379	5678	5986	5984	8%	5%	11%	11%	25%	46%	32%	22%	-24
Best Buy 2006	2458	27433	30848	35934	40023	14%	16%	31%	46%	46%	26%	40%	34%	8
Budget Rent a Car 2006			Private			-	-	-	-	-	55%	24%	22%	-
Chevrolet 2008	104589	135592	150276	152256	155427	14%	1%	12%	15%	46%	54%	28%	18%	-36
Chevrolet 2009	135592	150276	152256	155427	155929	5%	2%	3%	4%	15%	51%	31%	19%	-32
Circuit City 2006	10082	9890	10419	11514	12430	5%	11%	16%	26%	14%	35%	36%	29%	-6
CompUSA 2006	Data unavailable, acquired by Systemax 2007						-	-	-	-	35%	38%	26%	-
Continental Airlines 2006	16391	17379	19340	20098	20194	7%	4%	16%	16%	23%	41%	34%	25%	-16
Continental Airlines 2008	19340	20098	20194	16335	23325	-5%	-19%	-19%	16%	-16%	57%	40%	3%	-53
Continental Airlines 2009	20098	20194	16335	23325	37003	8%	43%	16%	83%*	16%	77%	16%	7%	-70
Delta Airlines 2006	15235	16480	17532	19154	22697	8%	9%	16%	38%	26%	30%	44%	26%	-4
Delta Airlines 2008	17532	19154	22697	28063	31755	17%	24%	47%	66%	60%	62%	27%	11%	-51
Delta Airlines 2009	19154	22697	28063	31755	35115	18%	13%	40%	55%	66%	59%	35%	7%	-52
Dodge 2008	Merged with Daimler 2010, then subsidiary only						-	-	-	-	53%	30%	17%	-36
Dodge 2009	Merged with Daimler 2010, then subsidiary only						-	-	-	-	62%	23%	15%	-47
Drug Emporium 2006	Bankrupt 2003						-	-	-	-	50%	29%	21%	-29

Enterprise Rent-A-Car 2006			Private				-	-	-	-	45%	33%	22%	-23
Ford 2008	160065	170572	143584	116283	128954	-9%	-19%	-32%	-24%	-27%	49%	37%	15%	-34
Ford 2008	160065	170572	143584	116283	128954	-9%	-19%	-32%	-24%	-27%	76%	12%	12%	-65
Ford 2009	170572	143584	116283	128954	135605	-8%	11%	-10%	-6%	-24%	52%	28%	20%	-32
Ford 2015	146917	144077	149558	151800	14518	1%	1%	5%	-90%	3%	39%	28%	33%	-6
Hertz Rent a Car 2006	6676	7469	8058	8686	8525	9%	8%	16%	14%	30%	40%	36%	23%	-17
Home Depot 2006	64816	73094	77019	79022	77349	7%	3%	8%	6%	22%	29%	39%	32%	3
Honda 2008	90841	101652	110047	91787	78658	1%	-17%	-10%	-23%	1%	18%	65%	18%	0
Honda 2009	101652	110047	91787	78658	81937	-8%	-14%	-29%	-26%	-23%	34%	32%	34%	0
Jet Blue 2008	2363	2843	3392	3292	3779	12%	-3%	16%	33%	39%	27%	51%	22%	-5
Jet Blue 2009	2843	3392	3292	3779	4504	10%	15%	11%	33%	33%	48%	4%	48%	0
Long's Drug Stores 2006	4527	4608	4545	4973	5263	3%	9%	8%	14%	10%	48%	30%	22%	-26
Lowes Home Improvement 2006	30838	36464	43243	46927	48283	15%	9%	29%	32%	52%	22%	42%	37%	15
McDonalds 2008	20896	22787	23522	22745	24075	3%	-3%	0%	6%	9%	82%	10%	8%	-74
National Car Rental 2006			Private				-	-	-	-	42%	36%	22%	-
Northwest Airlines 2006	11279	12286	12568	12528	13572	4%	0%	2%	10%	11%	55%	35%	11%	-44
Northwest Airlines 2008	12568	12528	13572	11108	-	-3%	-18%	-11%	-	-12%	61%	39%	0%	-61
Northwest Airlines 2009	12528	13572	11108	Bought in 2010		-	-	-	-	-	-	37%	1%	-60
Orchard Supply Hardware 2006			Private			-	-	-	-	-	34%	24%	42%	-
P W Supermarkets 2006			Private			-	-	-	-	-	45%	42%	13%	-
Petco Pet Store 2006	1610	1812	1996	2086	-	9%	4%	15%	-	30%	43%	33%	23%	-20
PetSmart Pet Store 2006	2993	3363	3760	4234	4673	12%	13%	26%	39%	41%	39%	27%	34%	-4
Rite Aid 2006	16600	16816	17163	17399	24327	2%	1%	3%	45%	5%	41%	35%	24%	-16
Safeway Food & Drug 2006	35823	38416	40185	42286	44104	6%	5%	10%	15%	18%	39%	40%	21%	-18
Save Mart Supermarkets 2006			Private			-	-	-	-	-	45%	35%	20%	-
Southwest Airlines 2006	6530	7584	9086	9861	11023	15%	9%	30%	45%	51%	29%	39%	32%	2
Southwest Airlines 2008	9086	9861	11023	10350	12104	5%	-6%	5%	23%	14%	38%	43%	20%	-18
Southwest Airlines 2009	9861	11023	10350	12104	15658	8%	17%	10%	42%	23%	55%	27%	17%	-38
Toyota 2008	192875	219566	241031	188224	173751	1%	-22%	-14%	-21%	-2%	44%	30%	26%	-18
Toyota 2009	219566	241031	188224	173751	174142	-7%	-8%	-28%	-28%	-21%	28%	38%	34%	7
US Airways 2006	7068	8051	11692	11813	12244	20%	1%	47%	52%	67%	42%	34%	24%	-19
US Airways 2008	11692	11813	12244	10609	12055	-3%	-13%	-10%	2%	-9%	69%	25%	6%	-63

US Airways 2009	11813	12244	10609	12055	13140	1%	14%	-2%	7%	2%	72%	22%	6%	-67
United Airlines 2006	9899	11208	13128	14232	15350	13%	8%	27%	37%	44%	52%	23%	25%	-27
United Airlines 2008	13128	14232	15350	12623	14351	-1%	-18%	-11%	1%	-4%	55%	41%	3%	-52
United Airlines 2009	14232	15350	12623	14351	37012	1%	14%	-7%	141%*	1%	75%	18%	7%	-69
United Drugs 2006	1516	1605	1776	1917	2038	8%	8%	19%	27%	26%	49%	39%	12%	-37
Verizon 2015		127079	131620	125980	126034	2%	-4%	-1%	-1%	5%	32%	20%	48%	16
Volkswagen 2008	111569	115848	121072	111901	134974	0%	-8%	-3%	17%	0%	46%	41%	13%	-33
Volkswagen 2009	115848	121072	111901	134974	169508	6%	21%	11%	40%	17%	56%	31%	13%	-44
Volvo 2009	32891	35108	25165	30510	35768	-0%	21%	-13%	2%	-7%	59%	29%	11%	-48
Walgreen Drug Stores 2006	37508	42202	47409	53762	59034	13%	13%	27%	40%	43%	39%	32%	29%	-10

Comment: * Three standard deviations greater than average revenue growth and excluded from analyses.

Table A2. Predicting the percent revenue growth with the NPS.

	Relative revenue growth					
	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.
Company evaluation						
NPS	0.22* (0.05)	0.19 (0.36)	0.26+ (0.08)	0.02 (0.94)	0.18* (0.04)	-0.05 (0.59)
Constant	0.10* (0.03)	0.29** (0.00)	0.14* (0.02)	0.26** (0.01)	0.05 (0.12)	0.20** (0.00)
Adjusted R^2	0.07	-0.00	0.06	-0.02	0.09	-0.01
Observations	53	53	53	50	53	50

Notes. Unstandardized regression coefficients from six OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.
$+p < 0.1$ $*p < 0.05$ $**p < 0.01$.

Table A3. Predicting the percent revenue growth with different NPS calculations.

	10 as Promoters 0-6 as detractors						10-7 promoters, 0-6 detractors					
	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.
Company evaluation												
NPS equation	0.20	-0.14	0.14	-0.04	0.16+	-0.11	0.14	-0.13	0.10	-0.03	0.11	-0.09
	(0.12)	(0.62)	(0.34)	(0.89)	(0.09)	(0.51)	(0.19)	(0.52)	(0.40)	(0.88)	(0.16)	(0.48)
Company industries												
Automobile manufacturer	-0.09	-0.08	-0.11+	-0.19*	-0.06	-0.07	-0.08	-0.08	-0.11+	-0.19*	-0.06	-0.06
	(0.11)	(0.28)	(0.07)	(0.02)	(0.11)	(0.13)	(0.14)	(0.32)	(0.08)	(0.02)	(0.13)	(0.14)
Computer retail	-0.03	0.15**	0.01	0.04	-0.02	0.07*	-0.03	0.15**	0.01	0.04	-0.02	0.07*
	(0.54)	(0.00)	(0.87)	(0.56)	(0.57)	(0.03)	(0.57)	(0.01)	(0.85)	(0.55)	(0.60)	(0.04)
Pharmacy	-0.04	0.05	-0.03	0.01	-0.03	0.05	-0.04	0.05	-0.03	0.01	-0.03	0.05
	(0.29)	(0.27)	(0.47)	(0.86)	(0.31)	(0.13)	(0.30)	(0.28)	(0.48)	(0.86)	(0.32)	(0.14)
Food retail	-0.06*	0.01	-0.06*	-0.09*	-0.04*	-0.01	-0.06*	0.02	-0.07*	-0.09*	-0.05*	-0.01
	(0.03)	(0.59)	(0.05)	(0.02)	(0.03)	(0.55)	(0.03)	(0.52)	(0.05)	(0.02)	(0.03)	(0.63)
Home improvement	-0.01	0.03	-0.03	-0.06	-0.01	-0.02	-0.01	0.04	-0.03	-0.06	-0.01	-0.01
	(0.83)	(0.53)	(0.65)	(0.47)	(0.86)	(0.62)	(0.81)	(0.49)	(0.64)	(0.49)	(0.85)	(0.71)
Restaurants	0.03	0.06	0.02	-0.09	0.02	-0.04	0.03	0.06	0.01	-0.09	0.02	-0.04
	(0.51)	(0.45)	(0.77)	(0.33)	(0.53)	(0.37)	(0.55)	(0.48)	(0.79)	(0.31)	(0.58)	(0.36)
Pet stores	0.00	0.07	-0.00	0.06	0.00	0.06**	0.01	0.06	-0.00	0.06	0.01	0.06**
	(0.94)	(0.26)	(0.93)	(0.15)	(0.93)	(0.00)	(0.83)	(0.29)	(0.99)	(0.14)	(0.82)	(0.00)
Telecommunications	-0.02	0.03	-0.04	-0.10**	-0.02	-0.03+	-0.02	0.03	-0.04	-0.10**	-0.01	-0.03+
	(0.46)	(0.27)	(0.24)	(0.01)	(0.45)	(0.08)	(0.49)	(0.32)	(0.25)	(0.01)	(0.47)	(0.09)
Year												
2008	-0.07	-0.22**	-0.11+	-0.07	-0.05	-0.03	-0.06	-0.23**	-0.10+	-0.07	-0.04	-0.04
	(0.12)	(0.00)	(0.05)	(0.26)	(0.14)	(0.23)	(0.23)	(0.00)	(0.09)	(0.26)	(0.24)	(0.22)
2009	-0.04	0.12*	-0.07	0.02	-0.04	0.08**	-0.02	0.10	-0.06	0.02	-0.02	0.07+
	(0.40)	(0.05)	(0.26)	(0.78)	(0.35)	(0.01)	(0.68)	(0.24)	(0.41)	(0.83)	(0.61)	(0.08)
Constant	0.17**	0.46**	0.27**	0.35**	0.11**	0.22**	0.16*	0.48**	0.26**	0.36**	0.11*	0.23**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.02)	(0.00)	(0.00)	(0.01)	(0.04)	(0.00)
Adjusted R^2	0.21	0.44	0.29	0.24	0.21	0.28	0.20	0.45	0.29	0.24	0.20	0.28
Observations	50	50	50	47	50	47	50	50	50	47	50	47

Notes. Unstandardized regression coefficients from twelve OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.

+$p < 0.1$ *$p < 0.05$ **$p < 0.01$.

Table A4. Predicting the percent revenue growth with different customer satisfaction measures.

	Satisfaction with the company						Liking of the company					
	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.
Company evaluation												
Satisfaction metric	0.23	-0.39	0.13	-0.11	0.18	-0.23	0.22	-0.12	0.17	0.02	0.16	-0.08
	(0.25)	(0.35)	(0.61)	(0.76)	(0.25)	(0.36)	(0.24)	(0.70)	(0.43)	(0.93)	(0.27)	(0.69)
Company industries												
Automobile manufacturer	-0.08	-0.06	-0.11+	-0.18*	-0.06	-0.06	-0.09	-0.08	-0.12*	-0.20*	-0.07	-0.07
	(0.12)	(0.41)	(0.07)	(0.03)	(0.11)	(0.21)	(0.12)	(0.30)	(0.05)	(0.01)	(0.12)	(0.13)
Computer retail	-0.02	0.16**	0.01	0.04	-0.02	0.07*	-0.03	0.15**	0.01	0.04	-0.02	0.06*
	(0.65)	(0.00)	(0.78)	(0.53)	(0.68)	(0.02)	(0.57)	(0.00)	(0.90)	(0.59)	(0.62)	(0.04)
Pharmacy	-0.05	0.06	-0.04	0.01	-0.04	0.06	-0.04	0.05	-0.03	0.01	-0.03	0.05
	(0.21)	(0.23)	(0.40)	(0.84)	(0.22)	(0.11)	(0.31)	(0.27)	(0.48)	(0.83)	(0.31)	(0.12)
Food retail	-0.07*	0.03	-0.07*	-0.08*	-0.05*	-0.00	-0.06*	0.02	-0.07*	-0.09*	-0.05*	-0.01
	(0.02)	(0.32)	(0.05)	(0.04)	(0.02)	(0.97)	(0.02)	(0.46)	(0.04)	(0.02)	(0.02)	(0.71)
Home improvement	-0.01	0.05	-0.02	-0.05	-0.00	-0.01	-0.01	0.03	-0.03	-0.06	-0.01	-0.02
	(0.87)	(0.40)	(0.71)	(0.51)	(0.91)	(0.80)	(0.81)	(0.57)	(0.63)	(0.42)	(0.87)	(0.58)
Restaurants	0.02	0.05	0.01	-0.09	0.01	-0.04	0.01	0.08	0.00	-0.08	0.00	-0.02
	(0.66)	(0.53)	(0.91)	(0.28)	(0.71)	(0.28)	(0.88)	(0.26)	(0.99)	(0.22)	(0.94)	(0.45)
Pet stores	-0.00	0.07	-0.00	0.06	-0.00	0.07**	-0.01	0.07	-0.01	0.05	-0.00	0.06**
	(0.98)	(0.25)	(0.92)	(0.15)	(1.00)	(0.00)	(0.84)	(0.25)	(0.82)	(0.22)	(0.89)	(0.01)
Telecommunications	-0.02	0.03	-0.04	-0.10*	-0.02	-0.03+	-0.02	0.04	-0.04	-0.10**	-0.02	-0.03+
	(0.43)	(0.30)	(0.23)	(0.01)	(0.41)	(0.07)	(0.35)	(0.21)	(0.19)	(0.01)	(0.34)	(0.09)
Year												
2008	-0.08+	-0.22**	-0.12*	-0.07	-0.06+	-0.03	-0.07+	-0.21**	-0.11*	-0.06	-0.06+	-0.03
	(0.06)	(0.00)	(0.03)	(0.29)	(0.06)	(0.23)	(0.09)	(0.00)	(0.03)	(0.30)	(0.09)	(0.29)
2009	-0.05	0.12*	-0.08	0.02	-0.05	0.09**	-0.04	0.12*	-0.07	0.03	-0.04	0.09**
	(0.23)	(0.03)	(0.17)	(0.79)	(0.20)	(0.01)	(0.36)	(0.03)	(0.23)	(0.73)	(0.29)	(0.01)
Constant	0.08	0.68*	0.24	0.42	0.05	0.34+	0.08	0.50*	0.20	0.32	0.06	0.24+
	(0.54)	(0.02)	(0.17)	(0.10)	(0.61)	(0.06)	(0.54)	(0.04)	(0.21)	(0.10)	(0.58)	(0.09)
Adjusted R^2	0.19	0.46	0.28	0.24	0.19	0.30	0.20	0.44	0.29	0.24	0.20	0.26
Observations	50	50	50	47	50	47	50	50	50	47	50	47

Notes. Unstandardized regression coefficients from twelve OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.

+$p < 0.1$ *$p < 0.05$ **$p < 0.01$.

Table A5.1. Predicting the percent revenue growth with the different likelihood to recommend versions.

	Eleven unipolar partially labeled						Seven bipolar partially labeled					
	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.
Company evaluation												
NPS	0.20	-0.17	0.17	-0.12	0.18	-0.18	0.42	-0.19	0.29	-0.00	0.29	-0.28
	(0.24)	(0.64)	(0.48)	(0.73)	(0.19)	(0.38)	(0.10)	(0.69)	(0.32)	(0.99)	(0.14)	(0.22)
Company industries												
Automobile manufacturer	-0.08	-0.07	-0.09+	-0.17*	-0.05	-0.06	-0.06	-0.04	-0.09	-0.16+	-0.04	-0.06
	(0.12)	(0.30)	(0.10)	(0.03)	(0.17)	(0.11)	(0.43)	(0.68)	(0.24)	(0.08)	(0.42)	(0.18)
Computer retail	-0.06	0.08	-0.01	0.01	-0.03	0.04	-0.04	0.15**	0.00	0.04	-0.03	0.07*
	(0.17)	(0.25)	(0.79)	(0.88)	(0.42)	(0.26)	(0.38)	(0.01)	(0.95)	(0.59)	(0.47)	(0.05)
Pharmacy	-0.05	0.04	-0.03	-0.00	-0.03	0.04	-0.07+	0.07	-0.05	0.01	-0.05	0.07*
	(0.26)	(0.48)	(0.48)	(0.98)	(0.36)	(0.26)	(0.09)	(0.21)	(0.28)	(0.86)	(0.12)	(0.04)
Food retail	-0.06*	-0.01	-0.06+	-0.10*	-0.04*	-0.02	-0.07**	0.02	-0.08*	-0.09*	-0.06*	0.00
	(0.02)	(0.86)	(0.06)	(0.01)	(0.05)	(0.26)	(0.01)	(0.41)	(0.04)	(0.05)	(0.01)	(0.83)
Home improvement	-0.01	0.02	-0.03	-0.06	-0.01	-0.02	-0.04	0.04	-0.04	-0.06	-0.02	0.00
	(0.77)	(0.71)	(0.61)	(0.48)	(0.87)	(0.60)	(0.50)	(0.55)	(0.52)	(0.50)	(0.61)	(0.90)
Restaurants	0.05	0.05	0.04	-0.10	0.04	-0.06	-	-	-	-	-	-
	(0.40)	(0.71)	(0.60)	(0.42)	(0.34)	(0.32)						
Pet stores	-0.01	0.06	-0.01	0.05	-0.00	0.06**	-0.01	0.07	-0.01	0.06	-0.01	0.07**
	(0.79)	(0.35)	(0.82)	(0.20)	(0.97)	(0.00)	(0.68)	(0.23)	(0.77)	(0.21)	(0.76)	(0.00)
Telecommunications	-0.03	0.02	-0.05	-0.11**	-0.02	-0.03*	-0.03	0.04	-0.05	-0.10*	-0.02	-0.02
	(0.24)	(0.48)	(0.17)	(0.01)	(0.36)	(0.05)	(0.30)	(0.13)	(0.21)	(0.02)	(0.30)	(0.20)
Year												
2008	-0.08+	-0.24**	-0.12*	-0.08	-0.06+	-0.04	-0.07	-0.21**	-0.10+	-0.06	-0.05	-0.03
	(0.07)	(0.00)	(0.03)	(0.15)	(0.08)	(0.11)	(0.11)	(0.00)	(0.07)	(0.34)	(0.11)	(0.30)
2009	-0.05	0.09	-0.08	-0.00	-0.04	0.07*	-	-	-	-	-	-
	(0.31)	(0.11)	(0.19)	(0.98)	(0.28)	(0.02)						
2015	0.00	-0.11	-0.10	-0.14+	-0.07*	-0.06+	-	-	-	-	-	-
	(0.97)	(0.18)	(0.13)	(0.09)	(0.04)	(0.07)						
Constant	0.12	0.54*	0.21	0.43+	0.06	0.31*	-0.05	0.53	0.12	0.34	-0.03	0.37*
	(0.33)	(0.04)	(0.21)	(0.09)	(0.55)	(0.04)	(0.79)	(0.12)	(0.53)	(0.22)	(0.84)	(0.02)
Adjusted R^2	0.19	0.41	0.26	0.24	0.20	0.28	0.15	0.39	0.21	0.19	0.13	0.33
Observations	53	53	53	50	53	50	34	35	34	32	34	33

Continued.

Table A5.2. Continued.

	Five unipolar fully labeled						Seven bipolar fully labeled					
	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.	Eq. 2.0.	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.
Company evaluation												
Likelihood to recommend version	0.12	-0.14	0.08	-0.17	0.10	-0.21	0.01	-0.26	-0.14	-0.08	0.00	0.03
	(0.61)	(0.71)	(0.76)	(0.63)	(0.62)	(0.24)	(0.97)	(0.63)	(0.63)	(0.89)	(0.99)	(0.93)
Company industries												
Automobile manufacturer	-0.04	-0.02	-0.04	-0.12+	-0.01	-0.06+	-0.12*	-0.08	-0.09+	-0.18+	-0.06+	-0.06
	(0.56)	(0.79)	(0.46)	(0.07)	(0.78)	(0.07)	(0.01)	(0.37)	(0.09)	(0.07)	(0.06)	(0.32)
Computer retail	-0.05	0.09	-0.00	0.02	-0.02	0.04	-0.16**	-0.09	-0.04	-0.09	-0.03	-0.05
	(0.28)	(0.19)	(1.00)	(0.75)	(0.64)	(0.21)	(0.00)	(0.27)	(0.61)	(0.43)	(0.53)	(0.26)
Pharmacy	-0.06	0.05	-0.04	0.01	-0.04	0.06	-	-	-	-	-	-
	(0.17)	(0.31)	(0.37)	(0.90)	(0.24)	(0.12)						
Food retail	-0.06*	0.00	-0.07*	-0.09*	-0.05*	-0.01	-	-	-	-	-	-
	(0.01)	(0.87)	(0.04)	(0.02)	(0.03)	(0.50)						
Home improvement	-0.01	0.02	-0.03	-0.05	-0.00	-0.01	-	-	-	-	-	-
	(0.84)	(0.69)	(0.68)	(0.57)	(0.97)	(0.87)						
Pet stores	-0.01	0.06	-0.01	0.06	0.00	0.06**	-	-	-	-	-	-
	(0.82)	(0.31)	(0.86)	(0.15)	(0.99)	(0.00)						
Telecommunications	-0.02	0.02	-0.04	-0.12*	-0.01	-0.04+	-	-	-	-	-	-
	(0.51)	(0.66)	(0.34)	(0.02)	(0.63)	(0.06)						
Restaurants	-	-	-	-	-	-	0.06	0.15+	0.07	0.01	0.05+	0.02
							(0.17)	(0.08)	(0.12)	(0.89)	(0.09)	(0.65)
Year												
2008	-0.08	-0.24**	-0.13+	-0.09	-0.07	-0.04+	-	-	-	-	-	-
	(0.15)	(0.00)	(0.06)	(0.15)	(0.13)	(0.09)						
2009	-	-	-	-	-	-	0.11+	0.46**	0.15*	0.20	0.07+	0.14*
							(0.07)	(0.00)	(0.04)	(0.12)	(0.07)	(0.04)
2015	0.00	-0.14	-0.11	-0.18+	-0.08	-0.09+	0.20**	0.29**	0.15	0.08	0.05	0.03
	(1.00)	(0.15)	(0.19)	(0.06)	(0.15)	(0.05)	(0.01)	(0.00)	(0.16)	(0.37)	(0.17)	(0.41)
Constant	0.17	0.50*	0.27	0.45+	0.11	0.31**	0.09	0.25	0.18	0.22	0.05	0.10
	(0.28)	(0.04)	(0.14)	(0.05)	(0.36)	(0.01)	(0.47)	(0.40)	(0.27)	(0.48)	(0.56)	(0.58)
Adjusted R^2	0.04	0.36	0.18	0.23	0.11	0.29	0.34	0.37	0.07	0.12	0.03	0.13

Observations	37	38	37	35	37	36	19	18	19	18	19	17

Continued.

Table A5.3. Continued.

	Percent						Number of recommendations					
	Eq. 2.0	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.	Eq. 2.0	Eq. 2.1.	Eq. 2.2.	Eq. 2.3.	Eq. 2.4.	Eq. 2.5.
Company evaluation												
Likelihood to recommend version	-0.19	-1.15*	-0.40+	-0.97*	-0.14	-0.55+	0.08	0.16**	0.09	0.12+	0.05	0.05+
	(0.23)	(0.01)	(0.05)	(0.02)	(0.23)	(0.06)	(0.16)	(0.00)	(0.13)	(0.09)	(0.33)	(0.06)
Company industries												
Automobile manufacturer	-0.10*	-0.04	-0.07	-0.13+	-0.05	-0.04	-0.08	-0.10	-0.10	-0.19*	-0.05	-0.08+
	(0.03)	(0.56)	(0.14)	(0.09)	(0.12)	(0.44)	(0.12)	(0.18)	(0.10)	(0.02)	(0.24)	(0.07)
Computer retail	-0.13**	0.06	0.01	0.06	-0.01	0.03	-0.02	0.14**	0.04	0.05	0.01	0.05+
	(0.01)	(0.50)	(0.91)	(0.61)	(0.85)	(0.55)	(0.72)	(0.00)	(0.47)	(0.44)	(0.90)	(0.07)
Restaurants	0.04	0.06	0.05	-0.07	0.04+	-0.04+	-0.02	0.06	-0.02	-0.10	-0.01	-0.03
	(0.12)	(0.25)	(0.22)	(0.14)	(0.06)	(0.10)	(0.66)	(0.37)	(0.73)	(0.11)	(0.79)	(0.31)
Pharmacy	-	-	-	-	-		-0.05	0.06	-0.03	0.02	-0.03	0.06
							(0.23)	(0.18)	(0.48)	(0.79)	(0.31)	(0.11)
Food retail	-	-	-	-	-		-0.05*	0.02	-0.06	-0.08+	-0.04+	-0.01
							(0.05)	(0.31)	(0.11)	(0.06)	(0.07)	(0.64)
Home improvement	-	-	-	-	-		0.01	0.02	-0.01	-0.06	0.01	-0.03
							(0.85)	(0.63)	(0.88)	(0.45)	(0.77)	(0.31)
Pet stores	-	-	-	-	-		0.01	0.07	0.01	0.06	0.01	0.06**
							(0.81)	(0.20)	(0.87)	(0.11)	(0.71)	(0.00)
Telecommunications	-	-	-	-	-		-0.02	0.03	-0.04	-0.10*	-0.02	-0.03+
							(0.37)	(0.25)	(0.28)	(0.02)	(0.48)	(0.06)
Year												
2008	-	-	-	-	-		-0.08+	-0.19**	-0.12*	-0.06	-0.07+	-0.02
							(0.08)	(0.00)	(0.04)	(0.39)	(0.08)	(0.29)
2009	0.12**	0.51**	0.16**	0.25**	0.08**	0.19**	-0.06	0.14**	-0.08	0.03	-0.05	0.09**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.24)	(0.00)	(0.18)	(0.73)	(0.19)	(0.00)
2015	0.23**	0.45**	0.20+	0.23*	0.08+	0.12*	-0.09	-0.27**	-0.20**	-0.26**	-0.13*	-0.11**
	(0.00)	(0.00)	(0.05)	(0.03)	(0.06)	(0.01)	(0.13)	(0.00)	(0.00)	(0.00)	(0.02)	(0.00)
Constant	0.16*	0.55**	0.25**	0.54**	0.10*	0.33**	0.23**	0.38**	0.30**	0.32**	0.16**	0.17**
	(0.01)	(0.00)	(0.01)	(0.00)	(0.04)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Adjusted R^2	0.38	0.73	0.23	0.52	0.09	0.48	0.18	0.43	0.27	0.25	0.18	0.27
Observations	19	18	19	18	19	17	53	53	53	50	52	50

Notes. Unstandardized regression coefficients from twelve OLS equations, p-values in parentheses. Only the participants who were assigned to the original 11-point scale included in the analyses. Observations nested in companies. All control variables were coded to range from 0 to 1. Revenue was estimated in U.S. dollars (currency conversion

rate from April 27, 2018). The following dummy categories served as reference categories: Industry: Airlines, Year: 2006.
$+p < 0.1$ $*p < 0.05$ $**p < 0.01$.

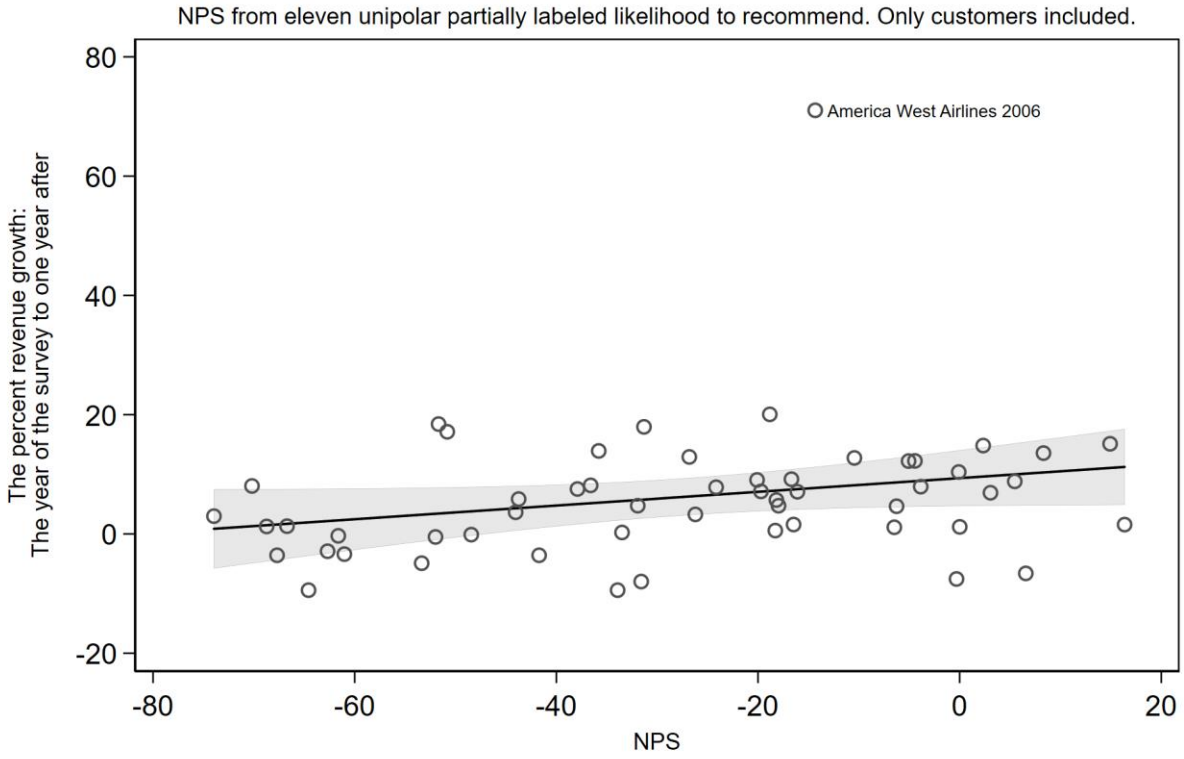

Figure A1. Scatterplot with fitted regression line predicting the percent average revenue growth over a three-year period from two years before the survey to one year after the survey (Eq. 2.0.) with the NPS from eleven unipolar partially labeled likelihood to recommend. Outliers in percent revenue growth above three standard deviations from the mean are labeled with company name and year. Study 1, 2, 3, and 4.

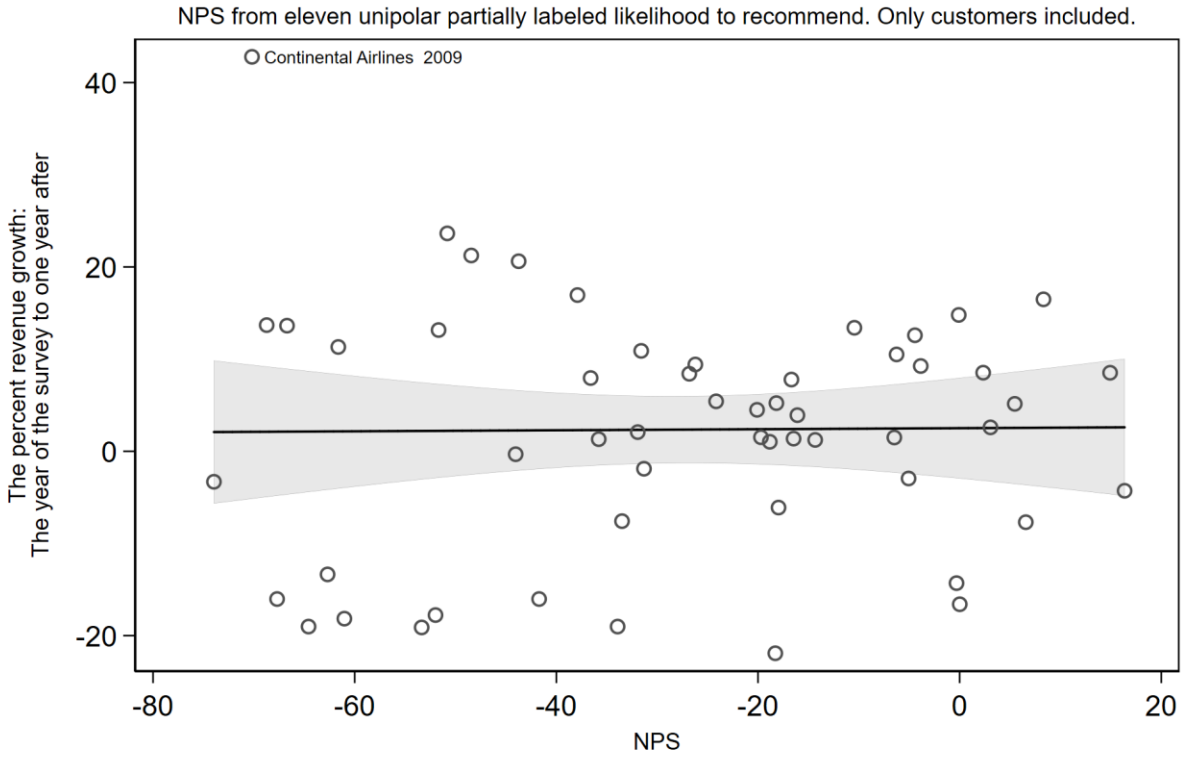

Figure A2. Scatterplot with fitted regression line predicting the percent revenue growth from the year of the survey to one year after (Eq. 2.1.) with the NPS from eleven unipolar partially labeled likelihood to recommend. Outliers in percent revenue growth above three standard deviations from the mean are labeled with company name and year. Study 1, 2, 3, and 4.

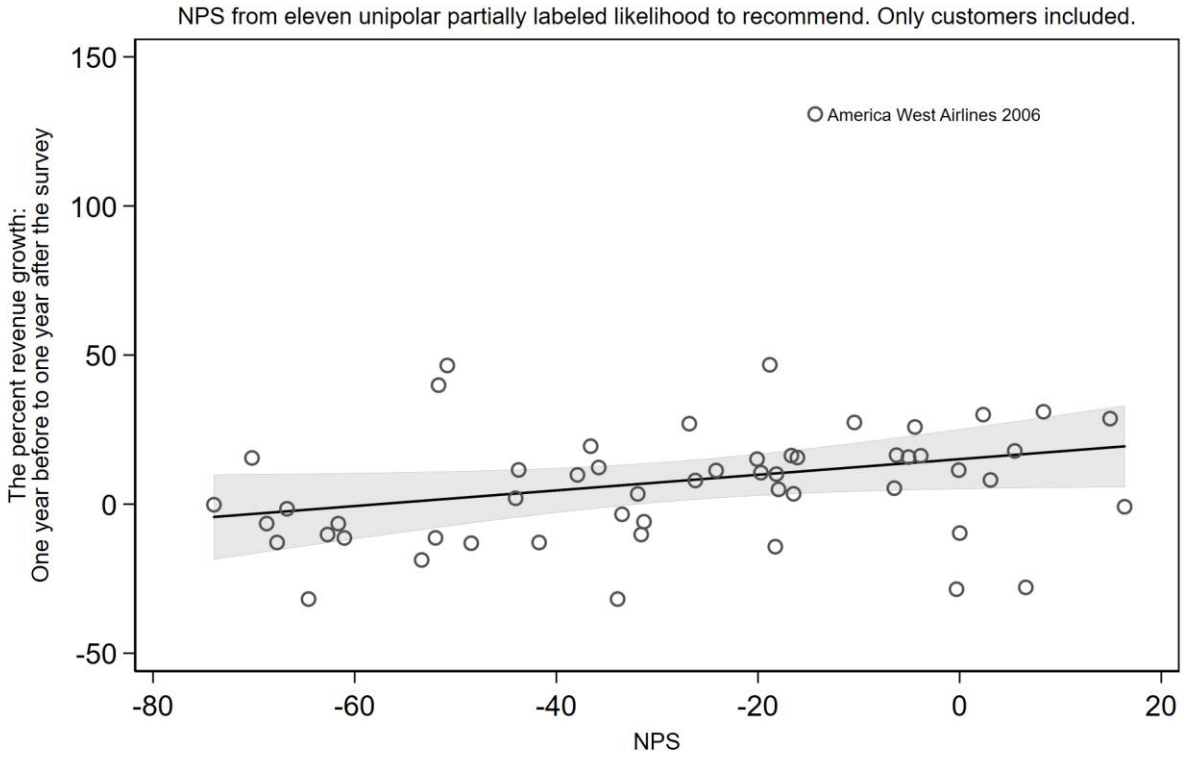

Figure A3. Scatterplot with fitted regression line predicting the percent revenue growth from one year before the survey to one year after the survey was administered (Eq. 2.2.) with the NPS from eleven unipolar partially labeled likelihood to recommend. Outliers in percent revenue growth above three standard deviations from the mean are labeled with company name and year. Study 1, 2, 3, and 4.

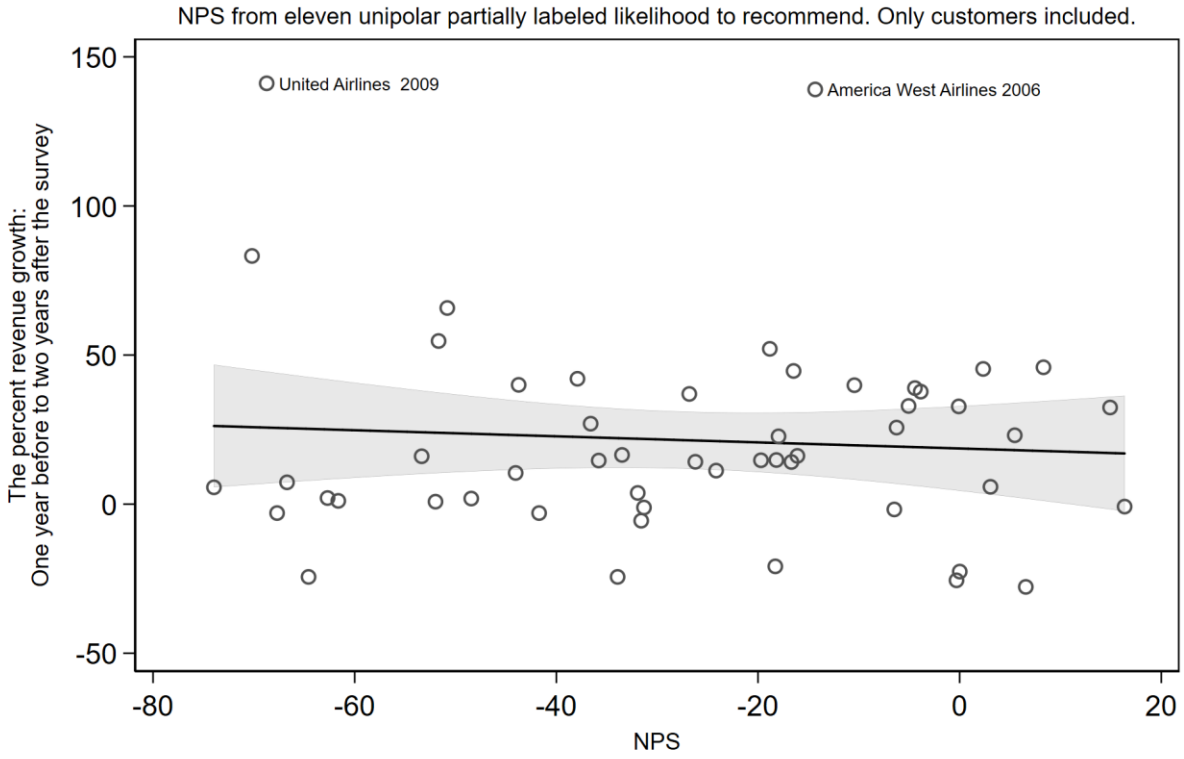

Figure A4. Scatterplot with fitted regression line predicting the percent revenue growth from one year before the survey to two years after the survey was administered (Eq. 2.3.) with the NPS from eleven unipolar partially labeled likelihood to recommend. Outliers above three standard deviations from the mean percent revenue growth excluded. Study 1, 2, 3, and 4.

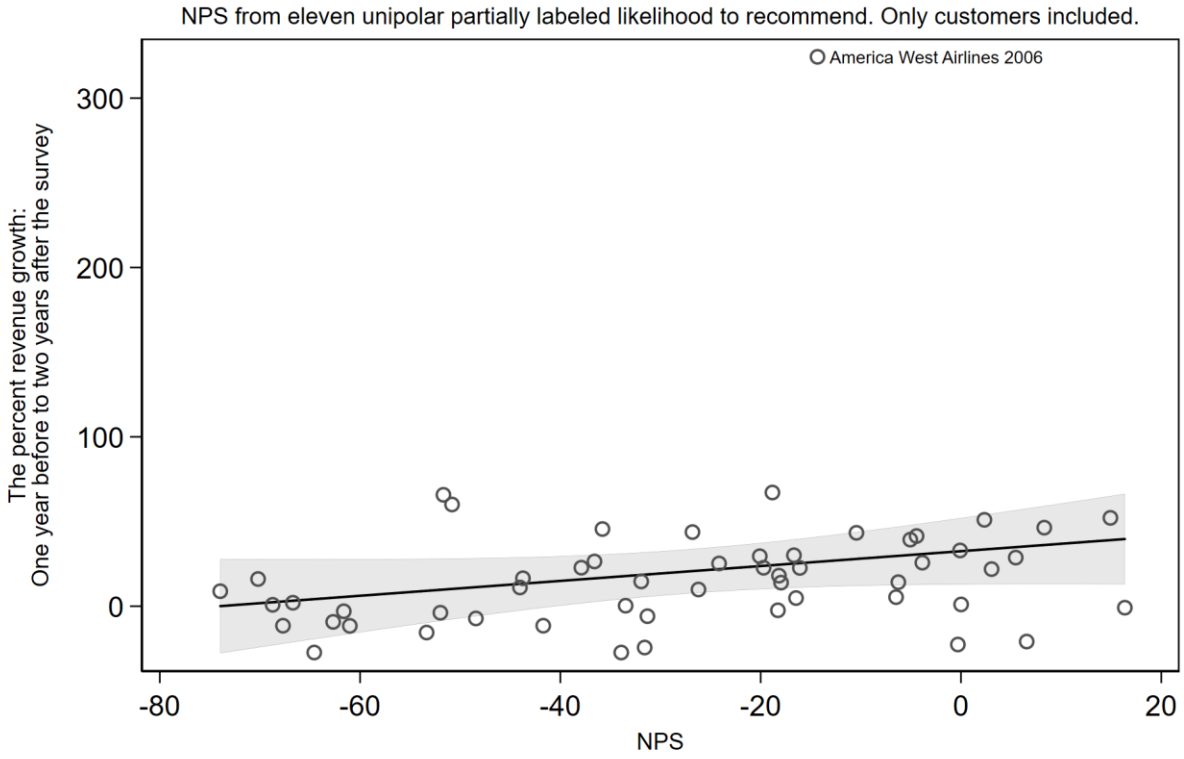

Figure A5. Scatterplot with fitted regression line predicting the percent revenue growth from two years before the survey to one year after the survey was administered (Eq. 2.4.) with the NPS from eleven unipolar partially labeled likelihood to recommend. Outliers above three standard deviations from the mean percent revenue growth excluded. Study 1, 2, 3, and 4.

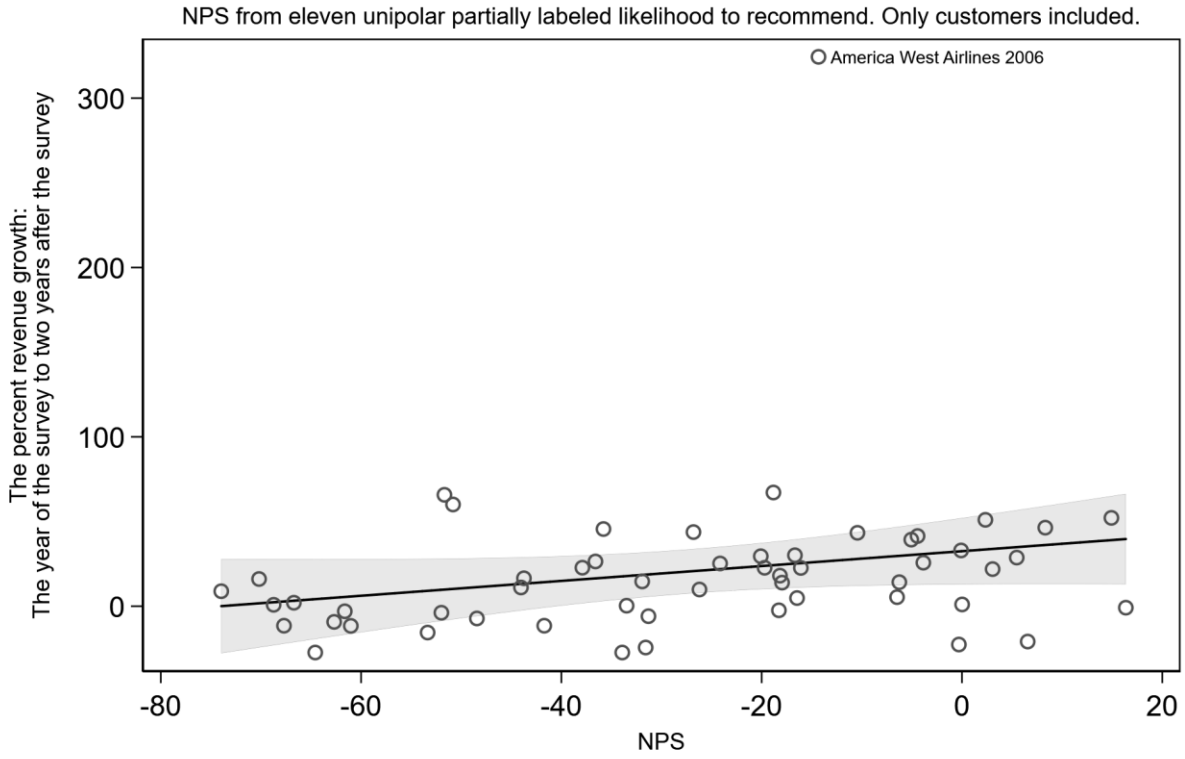

Figure 5. Scatterplot with fitted regression line predicting the percent revenue growth from the year of the survey to two years after the survey was administered with the NPS from eleven unipolar partially labeled likelihood to recommend. Outliers above three standard deviations from the mean percent revenue growth excluded. Study 1, 2, 3, and 4.