Can alternative metrics provide new insights from Net-Promoter data?

A thesis presented in partial fulfilment of the requirements for the degree of Master of Business Studies in Marketing at Massey University, Palmerston North, New Zealand

Philip James Mecredy

2016

ABSTRACT

Marketers regularly use loyalty measures to better understand consumers' purchase behaviour. In commercial market research the loyalty metric, Net Promoter Score (NPS), is commonly used due to its simplicity, and because there are claims that increases in NPS relate to increases in company revenue. However, the connection between NPS and revenue growth rates is widely criticised by scholars, casting doubt on the wisdom of implementing strategies that focus on increasing the numbers of highly loyal customers.

This research considers whether alternative metrics, derived from Net-Promoter data, can provide new insights into customer loyalty. It examines whether the NPS, likelihood mean, and Polarization Index measure different aspects of loyalty in the real estate (n=1,818) and agricultural (n=2,785) sectors. It then evaluates the ability of the three measures to predict changes in same customer spend and company revenue using data from the agricultural sector.

The findings show that the NPS and likelihood mean measure similar aspects of loyalty and that the Polarization Index measures a different aspect of loyalty when applied to 11-point Net-Promoter data. Longitudinal comparisons suggests that the NPS and likelihood mean are poor predictors of the current (t) and future (t+1) spend by the same customers, compared with the Polarization Index which provides a more accurate prediction. In contrast, the NPS and likelihood mean are found to have a strong relationship with current (t) and future (t+1) company revenue, while negative relationships were observed for the Polarization Index.

These findings suggest that loyal customers increase their spending less than disloyal customers, as they have likely reached saturation point with the company's products. However, loyal customers still contribute to company revenue growth by attracting new customers, presumably through Word-of-mouth (WOM). Therefore growth comes through penetration and increasing the amount spent by the least loyal customers, rather than through increasing spend by loyal customers.

ACKNOWLEDGEMENTS

First, I would like to sincerely thank my supervisors, Professor Malcolm Wright and Pam Feetham for the wealth of knowledge you have installed in me, and for constantly providing me with the motivation to continually improve. I have achieved immense personal growth and developed a passion for marketing under your guidance, for which I am truly grateful.

I would also like to acknowledge my parents who have continually supported me in every facet of life and provided me with the financial support to pursue my academic endeavours.

To my friends, thank you for providing me with a breath of fresh air and calming distraction from the rigors of academic study.

Lastly, thank you to all the staff and post graduate students in the School of Communication, Journalism and Marketing for the friendly encouragement and support you have provided.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION

1.1 Introduction to the Net-Promoter Metric

Customer loyalty measures are used extensively in commercial market research because of the perceived benefits of having highly loyal customers. Additionally, academic marketers use loyalty measures to better understand consumers and their purchasing behaviour. Investigating the accuracy of customer loyalty measures is important due to the substantial investments made by companies attempting to attract and maintain a highly loyal customer base.

One of the most commonly adopted loyalty measures, the Net-Promoter Score (NPS) developed by Frederick Reichheld, was introduced due to the increased academic and commercial interest in customer satisfaction and loyalty measures. Reichheld developed the Net-Promoter metric in the belief that current non-financial measures provide a poor gauge of company growth (Reichheld, 2003). To rectify this problem, Reichheld tested 20 loyalty questions to discover that one “would recommend” question had the strongest relationship with repeat purchases and referrals. Following this discovery, Reichheld tested the “would recommend” question across a dozen industries, finding that a strong correlation with growth rates existed in the majority of industries (2003). Given these findings, Reichheld claims that only a single “would recommend” question is needed to determine loyalty and predict growth. Additionally, he claims that this “would recommend” question is superior to other customer satisfaction and loyalty metrics.

The “would recommend” question, known as NPS, is determined by asking customers, “How likely is it that you would recommend [company X] to a friend or colleague?” Using an 11-point scale customers report their recommendation likelihood, with a score of 0 indicating “not at all likely” and 10 indicating “extremely likely” to recommend. Those scoring 9–10 are classified as “promoters”, 7–8 as “passively satisfied”, and 0–6 as “detractors”. To determine the NPS, only a simplistic calculation of the percentage of “promoters” minus the percentage of “detractors” is required. Due to its simplicity, the Net-Promoter metric has been adopted extensively in commercial market research (Creamer, 2006; Reichheld, 2006b). This extensive commercial adoption has stimulated increased academic attention to test and verify Reichheld’s (2003) claims on the superiority of the NPS.

1.2 Research Relevance and Purpose

A number of researchers have attempted to replicate the relationship between NPS and company growth and found considerable variations in results (Keiningham, Cooil, Andreassen, & Askoy, 2007a; Marsden, Samson & Upton, 2005; Morgan & Rego, 2006; Van Doorn, Leeflang & Tijs, 2013). Some researchers have compared the ability of the Net-Promoter metric to predict growth with other customer satisfaction measures (Morgan & Rego, 2006; Van Doorn et al., 2013). Additionally, Keiningham et al. (2007a) compared correlations of both NPS and the American Customer Satisfaction Index (ACSI) with company growth rates. These studies found that the Net-Promoter metric did not outperform the customer satisfaction and ACSI metrics, putting Reichheld’s claims in serious doubt.

Additionally, numerous academics have criticised the metric (Grisaffe, 2007; Keiningham et al., 2007a; Keiningham, Cooil, Askoy, Andreassen & Weiner, 2007b; Keiningham, Askoy Cooil, Andreassen & Williams, 2008a; Kristensen & Westlund, 2004; Morgan & Rego, 2006; Pingitore, Morgan, Rego, Gigliotti & Meyers, 2007; Rego & Morgan, 2004; Sharp, 2008). In particular, Grisaffe (2007) criticises the calculation of the NPS, using “promoters” minus “detractors”. Adding further concern to the NPS calculation, research found that “detractors” engage in both positive and negative word-of-mouth (East, Romaniuk & Lomax, 2011). These criticisms will be discussed in detail in section 3.7.

Despite significant academic criticism and limited validation of the relationship between NPS and company growth, NPS is still widely used commercially. Clearly, this commercial adoption presents a significant problem as companies are using a metric that may not accurately capture company growth. Therefore, the objective of this research is to explore whether the Net-Promoter question can be analysed using alternative methods to the NPS calculation of “promoters” minus “detractors” to predict company growth rates with more accuracy.

Specifically, this research aims to determine whether applying a mean similar to the Juster scale (Juster, 1966) or applying a Polarization Index (Sabavala & Morrison, 1977) to the Net-Promoter question will provide a better predictor of company growth rates than the calculation of NPS. As meta-analysis has shown large similarities between the predictive ability of 11-point likelihood scales and probability scales (Wright & MacRae, 2007), Net-Promoter data can be analysed similar to that of probability data. Therefore, it is possible to apply alternative measures of loyalty to the data, such as an average like the Juster scale or polarization (Juster, 1966; Sabavala & Morrison, 1977).

1.3 Organisation of this Thesis

This thesis begins with a literature review consisting of chapters’ two to five. Chapter two provides a discussion of the contrasting academic views regarding customer loyalty. The following chapter introduces the Net-Promoter metric, provides an overview of the NPS literature, and critically examines the strengths and weaknesses of the metric. Chapters four and five include an examination of the likelihood mean and Polarization Index respectively.

Following the literature review, the objectives and hypothesis of this research are identified in chapter six and the research methodology undertaken is described in chapter seven. In chapter eight, the findings of this research are examined in-depth and a critical discussion of these findings, their limitations, and the conclusion of this research is covered in chapter nine.

2.0 CUSTOMER LOYALTY

2.1 Introduction

The adoption of customer satisfaction and loyalty metrics is based on the belief that they provide an accurate indicator of future company performance. A large body of research claims that customer satisfaction and loyalty is linked to profitability (see Section 2.3). Furthermore, Reichheld and Sasser (1990), along with other researchers, have identified a variety of benefits from retaining loyal customers. However, more recent academic studies have found that the benefits of retaining loyal customers are overstated (East, Gendall, Hammond & Lomax, 2005; East, Hammond & Gendall, 2006; Reinartz & Kumar, 2000, 2002).

As the Net-Promoter metric relies on the relationship between customer loyalty and profitability, it is important that customer loyalty is discussed in depth. This chapter begins by identifying various loyalty definitions and a discussion of the links between customer satisfaction, loyalty and profitability. This is followed by a discussion of the benefits of customer loyalty, and whether such benefits are overstated. Finally, the double jeopardy law and divided loyalty are discussed as they reinforce that for a company to grow, attracting new customers is more important than customer retention (Riebe, Wright, Stern & Sharp, 2014).

2.2 Defining customer loyalty

The concept of customer loyalty has been subject to considerable conceptual debate. Early research used a variety of behavioural measures to capture loyalty (Cunningham, 1956; Kahn, Kalwani & Morrison, 1986; Lipstein, 1959; Kuehn, 1962; Massey, Montgomery & Morrison, 1970). However, according to Dick and Basu (1994), behavioural measures, such as purchase frequency, are unable to differentiate between actual loyalty and spurious loyalty, where customers continually purchase a good or service without having a strong attitude or commitment to it. For example, a consumer may continue to purchase a product because there is a lack of alternative options available.

Due to this criticism, many academics consider customer loyalty a multidimensional construct requiring both behavioural and attitudinal measures. Jacoby and Kyner define customer loyalty as having six necessary conditions; "the biased (i.e., non-random), behavioural response (i.e., purchase), expressed over time, by some decision-making unit, with respect to one or more alternative brands out of a set of such brands, and is the function of a psychological (decision-making, evaluative) processes" (1973, p. 2). Alternatively, Dick and Basu (1994) provide a more simplistic definition of customer loyalty as "the strength of the relationship between an individual's relative attitude and repeat patronage" (p. 99). While these definitions differ, they highlight the view that loyalty is much more than repeat purchases.

Reichheld (2003) agrees that loyalty is about more than simple behavioural measures such as repeat purchases. Reichheld defines it as the "willingness of someone – a customer, employee, or a friend – to make an investment or personal sacrifice in order to strengthen a relationship" (2003, p. 3). He argues that a recommendation from a customer is a personal sacrifice because they are risking their reputation, therefore, indicating that they must have strong attitudinal loyalty. Thus, while the Net-Promoter metric does not measure loyalty through actual purchase behaviour, it does capture loyalty through consumers' attitudes and intentions to recommend.

2.3 Links between customer satisfaction, loyalty and profitability

The positive link between customer satisfaction and increased loyalty has been discussed or observed in a variety of studies (Anderson and Sullivan, 1993; Fornell, 1992; Gronholdt, Martensen & Kristensen, 2000; Hallowell, 1996; LaBarbera & Mazursky, 1983). Additionally, many researchers claim customer satisfaction and/or loyalty is linked to profitability (Anderson, Fornell & Lehman, 1994; Bernhardt, Donthu, & Kennett, 2000; Bowen & Chen, 2001; Hallowell, 1996; Heskett, Jones, Loveman, Sasser & Schlesinger, 1994; Ittner & Larcker, 1998; LaBarbera & Mazursky, 1983; Reichheld & Sasser, 1990; Yeung, Ging, & Ennew, 2002). The considerable amount of literature in this area highlights significant academic support for a relationship between customer satisfaction, loyalty, and company profit.

Using 342,308 consumer responses, Bernhardt et al. (2000) conducted a longitudinal analysis of the relationship between satisfaction and profitability for a national chain of fast-food restaurants. Their cross-sectional analysis determined that no relationship existed between customer satisfaction and profit/sales. However, their analysis of time series data over a period of 12 months discovered a positive relationship between changes in customer satisfaction and changes in profit/sales (Bernhardt et al., 2000). Therefore, Bernhardt et al's. (2000) research indicates that increases in customer satisfaction will positively affect profitability in the long run; however, in the short run this relationship may not be observable.

Anderson et al. (1994) used the Swedish counterpart of the ACSI, the Swedish Customer Satisfaction Barometer (SCSB), to measure the financial impact of customer satisfaction. This study used data from 77 firms that hold the largest market shares in a range of industries to discover that an annual one-point increase in customer satisfaction is correlated with a $7.48million net present value over 5 years (Anderson et al., 1994). Similar to Bernhardt et al. (2000), this study determined that customer satisfaction affects future purchasing behaviour and long run profitability. Notably, this is due to satisfied customers becoming increasingly loyal and engaging in future purchasing.

Hallowell (1996) discovered a relationship between customer satisfaction and loyalty as well as customer loyalty and profitability in the banking sector. A 7-point Likert-type scale ranging from "very satisfied" to "very dissatisfied" was used to collect a variety of satisfaction data from 12,000 randomly selected retail-banking customers. This study found that customer satisfaction had a positive correlation with all measures of customer loyalty used, which included customer retention and relationship tenure. Additionally, seven out of eight regressions between customer loyalty and various profitability measures supported the positive relationship between customer loyalty and profitability.

Yeung et al. (2002) used the American Customer Satisfaction Index (ACSI) to examine whether the positive relationship between customer satisfaction and performance is linear. Over the observed ranges of satisfaction scores calculated by the ACSI, it was found that both linear and non-linear models fit the data, indicating that there is no reason to reject the assumption that a linear relationship exists between customer satisfaction and profitability.

These studies, among others, highlight the growing literature supporting the relationship between customer satisfaction, loyalty, and profitability. The existence of this link is of importance as non-financial measures, such as the Net-Promoter metric and the ACSI, are based on it. If the relationship does not exist, than these metrics will fail to provide an accurate measure on which managers can rely to assess whether a company's profitability will grow in response to improvements in loyalty.

2.4 Consequences of customer loyalty

Some researchers claim that there are various positive behavioural benefits of high customer loyalty. However, a growing body of research has contradicted and criticised these claims. This section discusses the alleged behavioural consequences of achieving high customer loyalty claimed by some researchers, followed by a critique of these claims in section 2.5.

While some research has found a relationship between customer loyalty and profitability, other research has studied the behavioural consequences of customer loyalty. One of the most expected behavioural consequence of loyalty is repeat patronage. Some studies claim that customer satisfaction and loyalty has a positive effect on customer retention and market share (Anderson & Sullivan 1993; Bolton, 1998; Rust & Zahorik, 1993). Additionally, some researchers believe that retaining customers is very important as it can significantly impact company profitability (Dawkins & Reichheld, 1990; Fornell & Wernerfelt, 1987; Reichheld & Sasser 1990). Highlighting the importance of the relationship, Reichheld and Sasser (1990) claim a 5% increase in retention rates corresponds to a 25–85% increase in profits.

Other research claims that a large increase in profitability can be achieved because the cost of keeping current customers is lower than attracting new customers (Fornell & Wernerfelt, 1987; Reichheld & Sasser, 1990). Ultimately, customer retention is a defensive strategy, as opposed to an offensive strategy focused on obtaining new customers and encouraging brand switching (Fornell & Wernerfelt, 1987). A defensive strategy involves fewer costs than an offensive strategy, as less marketing and promotions are required to attract new customers. This strategy is critical for mature industries and competitive markets as slow customer growth rates highlight the importance of customer retention (Fornell & Wernerfelt, 1987; Fornell, 1992).

Additionally, some researchers believe that companies who can retain loyal customers are able to charge higher prices, whilst customers are more likely to increase purchases and buy other goods and services from the company (Reichheld & Sasser, 1990; Zeithaml, 2000). Loyal customers are also thought to be more resistant to competitive efforts and less likely to leave should product quality temporarily decline (Anderson & Sullivan, 1993; Dick & Basu, 1994; Fornell, 1992).

Furthermore, research has found that as customers become loyal, their search for alternative brands decreases (Dick & Basu, 1994; Furse, Punj & Stewart 1984; Moore & Lehmann, 1980). In particular, Furse et al. (1984) investigated the amount of time spent searching for information when purchasing a new car using a sample of 1,031 respondents. Using cluster analysis they discovered that consumers who spent less time searching for information, purchased more cars and were more satisfied with past car purchases in comparison with those who conducted a more extensive search for information. Similarly, a study exploring information search for bread purchases found that respondents who repeatedly purchased a brand engage in a lower search for information (Moore & Lehmann, 1980). These two studies indicate that more satisfied and loyal customers are less likely to search for information about competitor brands, therefore decreasing their likelihood of switching brands.

Of particular importance, research has discovered that retained loyal customers are more likely to engage in positive WOM (Dick & Basu, 1994; Fornell, 1992; Reichheld & Sasser, 1990; Zeithaml, 2000). Positive WOM is thought to act as an offensive marketing tool; reducing the acquisition costs the company must spend to attract new customers to the firm (Reichheld & Sasser, 1990). Engagement in positive WOM is an important consequence of loyalty because it provides companies with free advertising and the ability to grow their market share.

In summary, research has highlighted a variety of behavioural consequences of achieving customer loyalty. These claims are that retained loyal customers require a lower cost to serve, are more likely to increase purchases, have a greater willingness to pay more for products, are less likely to search for information, are less vulnerable to competitors, and are more likely to engage in positive WOM. These proposed benefits support the belief that achieving customer loyalty will increase profitability and company growth. These perceived behavioural benefits of high loyalty indicate the reason behind the extensive adoption of customer loyalty measures in commercial market research, such as the Net-Promoter Score. However, as a large body of research disputes these claims, a critique of the benefits of customer loyalty is provided in the following section.

2.5 Misconceptions about loyalty

While many academics have identified the potential benefits of customer loyalty, a growing body of academics dispute the importance of achieving loyalty. Research provides counter claims that long-term customers do not always spend more (East et al., 2006), are not cheaper to serve (Reinartz & Kumar, 2000, 2002), do not pay higher prices (Reinartz & Kumar, 2002), and are not more likely to recommend a brand than short-term customers (East et al., 2005). Additionally, research has found that to maintain or improve market share, a focus on acquiring new customers is more important than customer retention (Riebe et al., 2014).

Between 1998 and 2004, East et al. (2006) tested the association between customer spending and relationship tenure for 17 services. On average, the correlation between relationship tenure and spending was 0.09, indicating a very weak positive relationship. Only credit card (0.32), outdoor clothing (0.21), and mobile phone airtime (0.19) had statistically significant correlations. While this does indicate that in some industries long-term loyal customers spend more, there are other industries where this relationship may not exist at all.

Additionally, Reinartz and Kumar (2000) tested the relationship of customer tenure with profitability and costs to serve. Using the purchase data of 9,167 households from a well-known U.S catalogue retailer, moderate correlations (0.179 and 0.219) were found between customer tenure and lifetime profits. The results indicate that long-term customers spend more; however, the strength of this relationship is relatively small, compared with earlier claims of the importance of retaining customers by Riechheld and Sasser (1990). Furthermore, this study found that long- and short-term customers did not have significantly different costs to serve (Reinartz & Kumar, 2000).

To build on earlier findings, Reinartz and Kumar (2002) studied the effects of customer loyalty across 16,000 customers from four companies: a U.S. corporate service provider, a U.S. mail-order company, a French food retailer, and a German direct brokerage house. Again, small to medium associations between tenure and profitability were found, with correlations of 0.30, 0.20, 0.45, and 0.29 respectively. Among the four companies studied, long-term customers were not found to be cheaper to serve than short-term customers (Reinartz & Kumar, 2002). In particular, long-term customers of the corporate service provider were found to be more expensive to serve. Additionally, in the four companies studied long-term customers were found to pay between 5% and 7% lower prices than short-term customers. In contrast to previous findings, this research suggests that long-term customers are less profitable than short-term customers.

East et al. (2005) analysed the relationship between customer tenure and recommendations across 23 service categories. The research found correlations for each service ranged from –0.44 to 0.25 with an average correlation of –0.01, indicating that long-term customers do not engage in more positive WOM by giving a recommendation. Of the 23 services studied, only four services – three car servicing and one main fashion store – had significant positive correlations between customer tenure and recommendations. In contrast, four categories – two credit cards, one car insurance, and one cheque book service – had a significant negative correlation. Overall, this research highlights that the relationship between customer tenure and recommendations is dependent on the industry studied.

Furthermore, Riebe et al. (2014) examined whether customer acquisition or retention had the greatest impact on market share. This study collected data from 100 doctors' anti-depressants prescriptions over a 1-year period in the pharmaceutical industry and company choice data for close to 10,000 respondents over a 4½-year period in the financial sector. A comparison of company acquisition and defection rates with stochastic benchmarks, calculated using the Dirichlet model, discovered that in both industries, customer acquisitions had the largest impact on market share (Riebe et al., 2014). A further finding from this research indicates that a change in customer retention rates had a slightly greater impact on profitability than a similar change in acquisition rates. Despite this, growing customer acquisitions were found to have a greater impact on changes in profitability as companies observe greater variation in acquisition rates (Riebe et al., 2014). Therefore, this research demonstrates that companies should not abandon strategies to acquire new customers by focusing solely on loyalty strategies to increase customer retention.

The research discussed in this section contradicts claims by Reichheld and Sasser (1990) that retaining loyal customers has a variety of consequential benefits that in turn affect a company's profitability. These studies highlight that across all industries, loyalty may not result in increased company profitability and customers who are cheaper to serve, more willing to pay higher prices, and more likely to recommend. It appears that the commercial impetus put on customer retention and loyalty is potentially a result of overestimated outcomes of loyalty and thus, the importance of customer acquisition should not be ignored. Consequently, companies need a better understanding of the importance loyalty in their particular industry before investing in expensive loyalty programmes.

2.6 Double Jeopardy

Patterns of loyalty are important in explaining company growth. Double jeopardy, a regularly occurring phenomenon that is receiving increasing academic and commercial attention, provides an important explanation on the pattern of loyalty. Double jeopardy states that brands with low market share have fewer buyers who tend to make purchases less frequently than brands with high market share (Ehrenberg, Goodhardt & Barwise, 1990). Therefore, smaller brands have less loyalty as their buyers make purchases at a less frequent-rate (Ehrenberg et al., 1990; Sharp, 2010).

Ehrenberg et al. (1990) provide a review of both behavioural and attitudinal cases of double jeopardy across frequently purchases consumer goods, durables, television, and newspapers. One case from the review showed that U.S. coffee brands with low market share had a lower market penetration, as well as a slightly lower purchase frequency than brands with high market share. A correlation of 0.65 was found between market penetration (the number of buyers), and the purchase frequency (the number of purchases per buyer). This example highlights that brands with low market share are hit with two negatives – lower market penetration and purchase frequency – hence the name double jeopardy.

The double jeopardy phenomenon has been observed for television programme choice (Barwise, 1986; Barwise & Ehrenberg, 1987; Donthu, 1994) and radio station choice (Lees & Wright, 2013; McDowell & Dick, 2005). Additionally, a clear pattern was found among past and current automobile purchases in Thailand (Bennett & Graham, 2010). Seeking to determine whether double jeopardy occurred in Australia and New Zealand, Wright, Sharp and Sharp (1998) uncovered similar results to previous studies across supermarkets, department stores, and petrol stations. These studies provide only a small reflection of the multiple industries in which double jeopardy is observed. However, it shows that the pattern is applicable to the majority of industries and therefore needs to be clearly recognised by companies as their loyalty measures may fail to reveal the naturally occurring pattern of double jeopardy.

According to Ehrenberg et al. (1990), double jeopardy arises whenever competitive brands have different market shares because larger brands conduct more advertising and are likely to be available in more places. For example, Coca-Cola is a brand with high market share across the world, therefore, Coca-Cola are able to spend more on advertising compared with smaller brands in the soft drink market. Additionally, due to its popularity, Coca-Cola is available at the majority of cafes, restaurants, supermarkets, dairies, etc. These two factors are likely influences on the higher purchase penetration and frequency observed in popular brands, such as Coca-Cola.

The double jeopardy phenomenon contrasts with views held by Reichheld and Sasser (1990) on the importance of retaining loyal customers as double jeopardy highlights that company growth is typically achieved through increases in market penetration. The double jeopardy pattern indicates that a focus should be put on increasing market share rather than loyalty, because loyalty does not differ much between brands. According to Ehrenberg and Goodhardt (2002), high market share is achieved by increasing brand penetration through growth in the customers who purchase the brand. As the number of brand buyers grows, the buyers will make more frequent purchases and market share will increase.

2.7 Divided loyalty

Many companies attempt to create sole loyalty among their customers, where 100% of category purchases are made from one brand. However, consumers infrequently purchase only one particular brand. Instead, most consumers in fast moving consumer goods markets display divided or polygamous loyalty where purchases are made from multiple brands in a consumer’s repertoire (Sharp, Wright & Goodhardt, 2002). While some buyers do make 100% of their purchases from one brand, they tend to have a significantly lower purchase frequency (Uncles, Ehrenberg & Hammond, 1995).

Divided loyalty occurs because buyers most often have various brands in a consideration set from which they make a purchase decision (Roberts, 1989). Marketers have understood for decades that most buyers display divided loyalty by purchasing more than one brand in the consideration set. For this reason, various behavioural measures of loyalty were developed such as proportion of purchases (Cunningham, 1956) and purchase sequence (Kahn et al., 1986). Share-of-category requirement (SCR), the proportion of category purchase for each brand, is now being used regularly to measure brand loyalty in academic market research (Bennett & Graham, 2010; Ehrenberg, Uncles & Goodhardt, 2004; Uncles et al., 1995).

Additionally, many studies that witnessed a double jeopardy pattern also observed limited sole loyalty (Bennett & Graham, 2010; Ehrenberg et al., 2004; Sharp et al., 2002; Uncles et al., 1995 Wright et al., 1998). Research by Ehrenberg et al. (2004) showed that only 12% of buyers of the U.S. instant coffee brand “Folgers” were 100% loyal. In this study, the 12% of sole buyers purchased coffee less frequently than those who purchased two or more brands. Furthermore, Wright et al. (1998) discovered that there is limited sole loyalty among supermarkets, department stores, and petrol stations brands and those who are loyal to one brand, purchase less frequently than the average buyer.

Divided loyalty highlights that brand growth does not necessarily come through generating 100% loyal customers as these buyers are minimal and usually purchase less frequently from a category. Similar to double jeopardy, divided loyalty shows that in order for market share and profitability to grow, brands must grow the number of consumers who purchase their product. Therefore, for brands to grow, companies must ensure that their brand is among a consumer’s consideration set (Roberts, 1989).

2.8 Customer Loyalty Summary

The research in this chapter highlights the contrasting academic views regarding customer loyalty. Some researchers view customer loyalty from a behavioural perspective and have used behavioural measures to capture loyalty (Cunningham, 1956; Kahn et al., 1986; Kuehn, 1962; Lipstein, 1959; Massey et al., 1970). Alternatively, other researchers have argued that customer loyalty is a combination of a strong attitude and repeat patronage (Dick & Basu, 1994; Jacoby & Kyner, 1973). Furthermore, when introducing the Net-Promoter metric, Reichheld (2003) claimed that loyalty requires a strong attitude towards a brand or company in order for a consumer to give a recommendation.

Despite the conceptual debate, many studies have observed a positive relationship between customer satisfaction and loyalty (Anderson & Sullivan, 1993; Fornell, 1992; Gronholdt et al., 2000; Hallowell, 1996; LaBarbera & Mazursky, 1983). Other studies have also observed that customer satisfaction and/or loyalty has a positive impact on profitability and company growth (Anderson et al., 1994; Bernhardt et al., 2000; Bowen & Chen, 2001; Hallowell, 1996; Heskett et al., 1994; Ittner & Larcker, 1998; LaBarbera & Mazursky, 1983; Reichheld & Sasser, 1990; Yeung et al., 2002).

Additionally, some researchers agree that the behavioural consequences of loyalty are customer retention, their willingness to pay higher prices, customers cost less to serve, are less likely to search for information on alternative brands, are less vulnerable to competitors, and are more likely to recommend the company or brand (Dick & Basu, 1994; Fornell & Wernerfelt, 1987; Furse et al., 1984; Moore & Lehmann, 1980; Reichheld & Sasser, 1990; Zeithaml, 2000). However, more recent findings by researchers have disputed these claims (East et al., 2005, 2006; Reinartz & Kumar 2000, 2002).

The effect of loyalty patterns must be accounted for when examining the impact of loyalty on company growth. Double jeopardy, a phenomenon that is receiving increasing commercial attention, describes a pattern of loyalty observed by many marketers. Double jeopardy states that brands with lower market share suffer from lower market penetration and slightly lower purchase frequency, which is a measure of loyalty (Ehrenberg et al., 1990). Therefore, there are two major consequences of brands with low market share. This pattern may occur because brands with high market share are sold in more locations and have the financial ability to conduct more advertising (Ehrenberg et al., 1990). As loyalty is found to differ only slightly between brands with different market shares, double jeopardy indicates that the focus should be on increasing a company's market share by growing the number of customers who purchase from the brand, rather than focusing on loyalty.

Many studies that have observed double jeopardy have also witnessed divided loyalty (Bennett & Graham, 2010; Ehrenberg et al., 2004; Sharp et al., 2002; Uncles et al., 1995 Wright et al., 1998), where consumers purchase from a repertoire of various brands (Sharp et al., 2002). This finding, along with double jeopardy, highlights that defensive loyalty strategies focused on customer retention are insufficient and companies need to focus on the quantity of consumers who purchase. This point was reinforced by Riebe et al. (2014), as it was discovered that customer acquisition had a bigger impact on changes in market share and profitability than customer retention.

These studies have all highlighted the abundance of academic debate surrounding customer loyalty. Nonetheless, customer loyalty measures are still used extensively in commercial and academic market research. Therefore, it is important that research focuses on how these measures can be improved and provide a more accurate estimate of purchase rates and company growth.

3.0 NET-PROMOTER SCORE

3.1 Introduction

This chapter explores Net-Promoter Score (NPS), first introduced by Reichheld (2003) as an alternative loyalty metric to predict brand growth. Reichheld believed that current loyalty measures, used both academically and commercially, provided a poor measure of company growth and that a better alternative was needed. Therefore, after testing a variety of loyalty questions, Reichheld proposed that only one "would recommend" question is needed to determine loyalty and predict company growth. To date, Net-Promoter is widely used in commercial market research despite limited academic support for the metric.

This section will give context to the research objectives and will highlight the reasons why alternative measures need to be considered for analysis of Net-Promoter data. The chapter begins by discussing how NPS is calculated and the theoretical foundations that underpin the metric. This is followed by a discussion of the extensive commercial adoption and the studies that have attempted to replicate or test Reichheld's Net-Promoter findings. Finally, the strengths and weaknesses of the metric are discussed.

3.2 How to calculate the NPS

Reichheld (2003) proposed that by asking a single "would recommend" question, loyalty is determined and growth predicted. To determine the Net-Promoter Score (NPS), customers are asked, "How likely is it that you would recommend [company X] to a friend or colleague?" Customers report their likelihood of recommendation on an 11-point scale ranging from 0 to 10. Scores of 10 indicate "extremely likely" to recommend, 0 indicates "not at all likely", and 5 indicates neutral. The respondents who give a score of 9–10 are classified as "promoters", 7–8 are "passively satisfied", and 0–6 are "detractors". The NPS is calculated by subtracting the percentage of "promoters" by the percentage of "detractors". This relatively basic calculation highlights the simplicity of NPS which has resulted in its widespread adoption in commercial market research despite its limited academic support.

3.3 Theoretical foundations

Academics and practitioners recognise the significant impact customer satisfaction and loyalty has on a firm's performance. However, Reichheld (2003) believed that current customer loyalty metrics provided an inadequate measure of company growth. For this reason, Reichheld introduced Net-Promoter as an alternative loyalty measure based on consumers' intentions to engage in positive word-of-mouth (WOM) by recommending a company or brand to a friend or colleague. Reichheld believed that intentions to recommend are one of the most prominent signs of customer loyalty because consumers are putting their reputation on the line. Despite inconclusive evidence that intentions to recommend is a main driver of loyalty, Reichheld argues that his loyalty metric measures consumers' intention to engage in positive WOM and will provide the most accurate indication of company growth.

For Reichheld's Net-Promoter metric to predict company growth accurately two assumptions must be met. The first assumption is that WOM impacts purchasing decisions, and the second assumption is that intentions to recommend correlate to actual behaviour. Both assumptions are discussed in this section, beginning with the first assumption. WOM is the informal advice between consumers about a product, service, organisation or brand (Anderson, 1998; East, Hammond & Lomax, 2008; Westbrook 1987). Some researchers have explored the effect WOM has on the adoption of new products. An early study by Arndt (1967) found that positive WOM increased the probability of purchasing a new food product, while negative WOM decreased it. Additionally, an exploratory study for a hypothetical personal computer indicated that both positive and negative WOM impact on attitudes and predicted purchase behaviour (Charlett, Garland & Marr, 1995).

Furthermore, research has investigated the effect WOM has on consumers' choice of service providers. Engel, Blackwell, and Kegerreis (1969) found that WOM was the most influential type of information sought by consumers when conducting a search for information into a new automotive diagnostic centre. Work by Keaveney (1995) also supports the idea that WOM affects consumer choices, as she investigated brand switching among service providers, discovering that half the respondents chose a new provider through WOM. This research indicates that WOM is an influential information channel when making a purchase decision.

There is a general consensus that WOM about a product or brand can be either positive or negative (Charlett et al., 1995; East et al., 2008). However, despite the impact and the incidence of positive and negative WOM differing (East et al., 2008; East, Hammond & Wright, 2007), the main goal for companies is to increase positive WOM and decrease negative WOM. This is consistent with the NPS calculation of "promoters" minus "detractors". To increase the NPS, you need to increase the "promoters" who engage in positive WOM or decrease the "detractors" who engage in negative WOM or do both (Reichheld, 2003).

Despite acknowledgement that WOM has a significant impact on an organisation there is limited research that examines the relationship between intentions to recommend and actual recommendations (Keiningham et al., 2007b). As Net-Promoter measures intentions to recommend, it is important that there is a strong link between intended behaviour and actual behaviour. If this relationship is weak then it is unlikely that the Net-Promoter will be accurate in forecasting whether a company will grow. A body of literature in consumer behaviour argues that a strong relationship between intentions and actual purchase behaviour does not exist (Lee, Elango & Schnaars, 1997; Juster, 1966; Pickering & Isherwood, 1974; Theil & Kosobud, 1968; Young, Desarbo & Morwitz, 1998). However, a limited number of studies have explored the relationship amongst intentions to recommend and actual recommendations.

One study that has examined the relationship between stated intentions to recommend and actual behaviour was conducted by Kumar, Petersen and Leone, (2007). They used a sample of 9,900 and 6,700 customers of a telecommunication and financial company respectively. In total, 81% of customers from the telecommunications company stated an intention to recommend, while only 30% actually recommended the company. For the financial company, 68% stated an intention to recommend; however, only 30% did so (Kumar et al., 2007). This indicates that fewer than half the consumers who stated an intention to recommend actually recommended the company.

Additionally, Romaniuk, Nguyen, and East (2011) investigated how accurately respondents' intentions to recommend six television programmes predicted actual recommendation behaviour. An 11-point probability scale was used for participants to rate their likelihood to recommend the six television programmes in the following week. This study found that only 30% of respondents who were classified as 'intenders' (7–10) gave a recommendation, while 95% of those who were classified as 'non-intenders' (0–4) did not recommend. Ultimately, due to the discrepancies among 'intenders', Romaniuk et al. (2011) concluded that recommend intentions provide an inaccurate measure of actual recommendations. However, this study focused solely on television programmes. Should similar results be found across multiple industries this would suggest that a majority of "promoters" will not engage in giving positive WOM. To date, little research has filled this gap; therefore, it is uncertain whether the limited relationship between intentions to recommend and actual recommendations will be observed in other industries.

3.4 Commercial adoption of Net-Promoter

Since its introduction, Net-Promoter has been extensively adopted in commercial market research because it is simple to use and easy to understand. NPS is used by multi-national corporations including GE, American Express, and Microsoft, to report to investors and to determine pay for senior staff (Creamer, 2006). Companies that have adopted the metric have an average NPS of between 5 and 10 percent, showing that "promoters" only slightly outweigh the detractors (Reichheld, 2006b).

Prominent companies performing exceptionally well may have a significantly higher NPS, indicating substantial differences between “promoters” and “detractors”. USAA (82%), Harley-Davidson (81%), Costco (79%), and Amazon (73%) are leading companies in terms of NPS (Creamer, 2006; Reichheld 2006b).

3.5 Extensions of the Net-Promoter metric

As noted earlier, Reichheld (2003) collected information on 20 questions thought to capture customer loyalty. From 4,000 survey responses Reichheld determined that one “would recommend” question had the strongest correlation with two drivers of growth – repeat purchases and referrals. The “would recommend” question outperformed all others by ranking first or second in 11 of the 14 case studies. To validate this finding, Reichheld decided to test the “would recommend” question more rigorously. Examining the relationship between the “would recommend” question and company growth from a dozen industries, he discovered a strong correlation between NPS and growth rates in most industries (Reichheld, 2003). Reichheld also claimed that despite variations between industries, a 12-point increase in NPS corresponds to a doubling of a company’s growth rate (Reichheld, 2006a). However, this variation is expected as Reichheld (2003) acknowledges that NPS predictability of growth differs between industries and fails to predict growth in industries where a company has a near monopoly share of the market. Nonetheless, his findings of a strong correlation between NPS and company growth allowed Reichheld to claim that the NPS is the only number needed to measure loyalty and predict company growth.

One study in the UK attempted to replicate the relationship between NPS and company expansion (Marsden et al., 2005). In total, 1,256 UK consumers reported their likelihood to recommend their current bank, mobile phone network, supermarket, and make of car to a friend or colleague via a telephone interview in 2005 (Marsden et al., 2005). By calculating the consumers NPSs and correlating them with 2003 and 2004 sales figures for each respective industry, Marsden et al. found that the NPS was a statistically significant predictor of sales and growth. Their study indicated that brands with high NPS grew faster than their competitors, and brands with low NPS grew slower than their competitors. Additionally, these researchers claim that a 7-point increase in NPS produced a one percent increase in brand growth (Marsden et al., 2005).

However, other researchers dispute Reichheld's claim that Net-Promoter is the most accurate method to measure growth and manage loyalty. Morgan and Rego (2006) compared six different customer satisfaction and loyalty metrics and their ability to predict future company performance. The six measures compared included an American Customer Satisfaction Index (ACSI) score, top 2 box satisfaction score, NPS, repurchase likelihood, and number of recommendations. By comparing the measures with various indicators of firm performance they discovered that the NPS is not significantly associated with a company's sales growth. This indicated that increasing the number of Net-Promoters will not have a positive effect on business performance. Therefore, Morgan and Rego (2006) recommend that managers should not abandon other methods of monitoring customer satisfaction and should not rely solely on NPS.

Furthermore, the findings of Reichheld (2003) and Marsden et al. (2005) are considered flawed because their NPSs were correlated with past growth rates (Keiningham et al., 2007a; Sharp 2008). A close inspection of Reichheld's (2003) study indicates that the research firm Satmetrix began collecting NPSs in the first quarter of 2001. These scores were then correlated with average company growth rates over a 3-year period from 1999 to 2002. Additionally, the replication of Reichheld's findings in the UK by Marsden et al. (2005) correlated 2005 NPSs with past growth rates from 2003 and 2004 (Keiningham et al., 2007a). The comparison of NPS with past growth rates indicates caution is needed when suggesting that a causal relationship exists between NPS and future sales growth.

To address the concerns about Net-Promoter, Keiningham et al. (2007a) examined the relationship between NPS and company growth rates in a cross-industry longitudinal study. They replicated Reichheld's study, but instead of using past growth rates they correlated NPS with company growth rates from identical time periods. Using longitudinal data from the Norwegian Customer Satisfaction Barometer (NCSB), which included 21 firms and approximately 16,000 interviews, they found no support for Reichheld's claim that Net-Promoter is the only question required to measure growth in customer surveys. By comparing correlations of both NPS and ACSI scores with company growth rates, they determined that Net-Promoter performance is not superior to the ACSI. This discovery signals that the commercial over-reliance of NPS is unwise, given this evidence contradicts Reichheld's claims.

Recently, Van Doorn et al. (2013) compared the performance of Net-Promoter and customer satisfaction metrics as predictors of growth rates. Using a market research agency, 11,967 responses were collected for 46 companies in the banking, insurance, utilities, and telecom industries via an online survey in the Netherlands. Customers were asked about their satisfaction, loyalty, and willingness to recommend. The customer satisfaction and loyalty data, collected in 2008, were correlated with sales revenue growth, gross margin, and net operating cash flow from the years 2008–2010. This ensured that the metric's ability to predict future performance and growth was being analysed (Van Doorn et al., 2013). The study found little distinction between the metrics' performance in predicting current revenue growth. The performance of all metrics to predict future sales growth, gross margins, and current and future net cash flows were poor. This indicates that Reichheld's (2003) claim that Net-Promoter is superior to other metrics in predicting growth rates is not supported. However, as other metrics did not outperform Net-Promoter, there is no evidence to suggest that companies measuring NPS should abandon the metric in favour of a particular alternative.

Since the development of Net-Promoter, radical commercial adoption has occurred without significant validation or replication. While a few studies (Marsden et al., 2005; Morgan & Rego, 2006; Keiningham et al., 2007a; Van Doorn et al., 2013) have attempted to replicate Reichheld's work, there has been a significant variation in findings. With limited academic support for the superiority of Net-Promoter there are concerns about its extensive adoption in commercial market research. Findings from Keningham et al. (2007a) and Van Doorn et al. (2013) indicate it could be unfavourable for a firm to focus solely on using NPS to evaluate performance and as the basis for company decision making. It is clear that other strengths and weaknesses of the Net-Promoter metric need to be scrutinised before determining whether NPS is an accurate predictor of company growth.

3.6 Strengths of the NPS metric

The underlining strength of the Net-Promoter metric is its simplicity of use (Grisaffe, 2007; Keiningham, Askoy, Cooil & Andreassen, 2008b). Calculating the NPS only requires a simple subtraction of “promoters” by “detractors”. This simplicity also makes the NPS very easy for managers and shareholders to understand and interpret. Because it is simple to use and easy to understand, significant commercially adoption has occurred (Creamer, 2006). With numerous companies using the Net-Promoter metric, it has become very useful for comparisons across organisations and industries.

Another advantage of the metric is that companies are now able to reduce their long and complex satisfaction and loyalty surveys to one concise would-recommend question, decreasing the amount spent on research (Keiningham et al., 2008b). Additionally, shortening complex surveys down to one question is likely to increase response rates as dropout rates are likely to decrease. However, ease of use and convenience are not the only considerations needed when using a loyalty metric – the relevance of the measure to the current global economy is also important.

Net-Promoter is one of the many customer satisfaction and loyalty measures developed in recent decades as financial measures have become less relevant (Ittner & Larcker, 2003; Ittner, Larcker & Rajan, 1997). In an attempt to measure loyalty, Net-Promoter includes a measure of WOM, a measure some researchers agree can significantly affect company performance. The WOM foundations on which Net-Promoter is built provide strength to the metric as the 21^{st} century has seen increased WOM communication through an exponential rise in internet and social media (Mangold & Faulds, 2009). Essentially, the Net-Promoter metric is in line with the changes taking place in the 21st century that are resulting in increased WOM communication (Grisaffe, 2007). It would seem that Net-Promoter is an important customer loyalty measure aligning with new technological developments; however, questions remain about its accuracy in measuring loyalty and predicting company growth.

Overall, Net-Promoter has been adopted because of its simplicity to use and ease of comparison across companies and industries. As the calculation of NPS is very basic, companies are able to use this method to calculate loyalty without needing to conduct expensive market research. Additionally, the metric has significant relevance to technological developments in the 21st century. However, the strengths of Net-Promoter must be weighed up against the weaknesses when making a decision on whether to adopt the metric as a predictor of company growth.

3.7 Weaknesses of the NPS metric

Since its introduction, NPS has received significant criticism in academic research (East et al., 2011; Grisaffe, 2007; Keiningham et al., 2007a, 2007b, 2008a; Kristensen & Westlund, 2004; Morgan & Rego, 2006; Pingitore et al., 2007; Rego & Morgan, 2004; Sharp, 2008). The first concern, as identified earlier, is the lack of research that supports the superiority of Net-Promoter in predicting growth rates. Outside of studies by Reichheld, only Marsden et al. (2005) find strong support for the positive relationship NPS has with sales growth. Alternative findings (Keiningham et al., 2007a; Morgan & Rego, 2006; Van Doorn et al., 2013) indicate a lack of validation of the superiority of the Net-Promoter metric.

Reichheld (2003) identified that Net-Promoter was not applicable in a few of the industries he studied. With only a dozen or more industries tested, discovering that a few industries NPSs could not predict growth rates illustrates a significant proportion of industry is unable to use the metric. In particular, Net-Promoter was shown to be ineffective in determining growth rates in monopolistic industry and for niche companies (Reichheld, 2003). These inconsistencies indicate clearly that the NPS may not be applicable to many industries. Reichheld's claim that NPS is the only number companies require to predict growth is therefore not universally accepted (Grisaffe, 2007).

Further concerns with Reichheld's methodology are that the 11-point Net-Promoter scale is broken into three categories and that the "passively satisfied" category is excluded from his calculations (Grisaffe, 2007). Currently, a customer's NPS is grouped into one of three categories, "promoters" (9–10), "passively satisfied" (7–8), or "detractors" (0–6). Grisaffe (2007) raises the concern that clustering the Net-Promoter scale results in different scenarios requiring diverse managerial actions being seen as similar. For example, Company X may have 100 "promoters", 0 "passively satisfied", and 100 "detractors". Alternatively, Company Y may have 0 "promoters", 200 "passively satisfied", and 0 "detractors". Despite significantly different scenarios, when "detractors" are subtracted from "promoters", both companies have a NPS of zero. In this case, relying on NPS is likely to mislead marketing decisions as it is unclear which customer group requires more focus (Grisaffe, 2007).

Reichheld not only reduced the 11-point likelihood scale to three categories, but also excluded the "passively satisfied" from the calculation of NPS altogether. While debate has continued over the optimal number of scale categories to use, other researchers have found that validity and reliability worsens when the number of scale points is reduced, especially when reduced to 4-point scales (Green & Rao, 1970; Lozano, García-Cueto & Muñiz, 2008; Preston & Colman, 2000). Due to this, Grisaffe (2007) questions why only "promoters" and "detractors" are used to calculate the NPS when it is expected that more scale points would result in a more accurate prediction of growth. This criticism represents a valid argument and highlights the need to apply different methods of analysis to the Net-Promoter question to assess whether they provide a better indicator of company growth.

Additionally, concern is raised over including rating scores of six in the “detractor” category (Grisaffe, 2007). The main concern is that the Net-Promoter scale considers a score of five as being “neutral”. However, by grouping the zero to six ratings together, consumers who rate their likelihood to recommend as “neutral” or slightly above are assumed to engage in negative word-of-mouth (WOM) and have the same chances of defecting as those that give a rating of zero. This criticism adds further weight to the NPS calculation of “promoters” minus ‘detractors”, indicating that other methods of analysing the ‘would recommend’ question need to be explored.

Furthermore, East et al. (2008) suggest there are four weaknesses in the Net-Promoter metric. The first weakness is that the likelihood to recommend question requires self-prediction. By asking a consumer their likelihood to recommend, requires a prediction of when a situation would arise that would allow a recommendation about a specific organization to be given. Research has found that self-predicted recommendation intentions are an inaccurate predictor of actual recommendations (Romaniuk et al., 2011).

Second, the Net-Promoter metric assumes that ‘one size fits all’ (East et al., 2008). For this reason, East et al. (2008) claim that NPS does not allow for variation in the impact of word-of-mouth (WOM) across brands. For example, depending on a variety of factors, positive word-of-mouth (WOM) about one particular brand may have a larger impact on purchase probability as opposed to an alternative brand. Additionally, the ‘one size fits all’ theory assumes that a given likelihood to recommend score has identical growth consequences for companies with varying market share (Uncles, East & Lomax, 2010).

Third, NPS does not directly measure negative word-of-mouth (WOM). Instead, the metric infers negative WOM from “detractors”, who give low scores on the willingness to engage in positive WOM. A study by East et al. (2011) discovered that NPS is poor at capturing negative WOM because “detractors” were shown to engage in both positive and negative WOM. Furthermore, a study by East, Hammond, and Wright (2007) found that 49% of negative WOM was given by ex-users of a brand and 30% was given by customers who have never used the brand. The findings of

East et al. (2011) support this statement: 55% and 22% of negative WOM was given by ex-users and never-users, respectively. However, only current users of a brand are typically sampled when calculating a brand's NPS, highlighting the ineffectiveness of the metric in measuring negative WOM.

The fourth weakness of the NPS is that it measures given WOM, as opposed to received WOM. Received WOM is likely to be more effective in measuring the impact of WOM, because a consumer may give a recommendation to more than one consumer at a time (East et al., 2008, 2011). Therefore, the quantity of consumers who have received a recommendation will provide a better measure of WOM.

Due to the numerous weaknesses and poor predictive performance of the Net-Promoter metric, East et al. (2008) introduce an alternative method for measuring WOM. The new metric measures the net effect of positive and negative WOM by combining its measure of WOM impact with East et al.'s (2007) measure of WOM incidence. East et al.'s (2008) proposed metric measures the incidence and impact of received positive and negative WOM on all brands in the category. The impact and incidence measures are multiplied together to calculate the combined effect of positive and negative WOM on each category brand. Following this, East et al. (2008) divided the combined effect by the market share to evaluate the proportionate effect on market share. However, it was acknowledged that the performance of this metric will need to be tested against NPS to ascertain whether it provides a better measure of WOM (East et al., 2008). While this measure will not be investigated in this research it provides an opportunity for future comparisons with NPS.

3.8 Net-Promoter summary

This chapter has highlighted the relative simplicity of administering and calculating the NPS. It is this simplicity and ease of comparison across industries and companies that have resulted in its extensive adoption in commercial market research. Despite the commercial adoption of the Net-Promoter metric, there is little academic support for its use over other performance measures.

While Reichheld (2003) acknowledges that NPS does not predict growth in every industry, he claims that NPS is the only number required to measure loyalty and predict whether a company will grow. Reichheld believes that, on average, a 12-point increase in NPS results in a company doubling its growth rate (Reichheld, 2006a). Since Reichheld's initial claim, one study cited a 7-point increase in NPS resulted in a one percent increase in company growth (Marsden et al., 2005). Alternatively, other studies replicating Reichheld's work or comparing NPS with other performance measures have found contrasting results that do not support claims of Net-Promoter's superiority at predicting company growth (Keiningham et al., 2007a; Morgan & Rego., 2006; Van Doorn et al., 2013).

In particular, Keiningham et al. (2007a) disputed the NPS growth claims of Reichheld (2003) and Marsden et al. (2005) because both these studies correlated NPS with past growth rates. In order to rectify this issue, the first cross-industry longitudinal Net-Promoter study took place that correlated NPS with company growth from identical years, as opposed to past years (Keiningham et al., 2007a). From this study, no support was found for Reichheld's NPS claims and NPS was not found to outperform the ACSI.

Additionally, a significant body of literature has criticised and identified weaknesses in the Net-Promoter metric. Specifically, the clustering of the Net-Promoter scale, the calculation of "promoters" minus "detractors" and the inclusion of respondents who indicate a likelihood to recommend of "neutral" or slight above as "detractors" are highly criticised (Grisaffe, 2007). Furthermore, it has been discovered that "detractors" engage in both positive and negative WOM and that the majority of negative WOM is given by ex-users and never-users, which are not typically sampled by NPS (East et al., 2007, 2011).

A further point that is relevant to this research is that Reichheld does not reveal whether any other analysis was undertaken with the Net-Promoter data to ascertain whether predictions of company growth could be improved in comparison to the NPS calculation. Ultimately, this uncertainty coupled with conflicting NPS results indicates an opportunity to predict company growth rates from Net-Promoter data more accurately. As there is little distinction between 11-point likelihood and probability scales (Wright & MacRae, 2007), alternative methods of analysis such as a likelihood mean similar to the Juster mean (Juster, 1966) and Polarization Index (Sabavala & Morrison, 1977) can be applied to the data. The following chapter discusses the Juster scale and the associated alternative analysis method.

4.0 JUSTER SCALE

4. 1 Introduction

The Juster scale, an 11-point purchase probability scale developed by Thomas Juster, is used to estimate actual purchase rates of a population from a sample of consumers within that population (Juster, 1966). The scale predicts actual purchase rates by using consumers' probability of purchasing a selected brand or product category over a given time period (Uncles & Lee, 2006).

Before the measurement of purchase probabilities, forecasts of purchase rates were usually calculated using surveys of purchase intentions. Intention surveys ask consumers about their plans to purchase a good or service. However, intention scales have been found to be a poor predictor of actual purchase rates (Juster, 1966; Lee et al., 1997; Pickering & Isherwood, 1974; Theil & Kosobod, 1968; Young et al., 1998). In an attempt to provide more accurate estimates of actual purchase rates, early studies into the performance of purchase probabilities were undertaken by Byrnes (1964) and Ferber and Piskie (1965). These two studies provided limited support for the use of purchase probabilities to predict purchase rates. However, they did provide the basis for the development of the Juster scale, which improved predictions of household automobile purchases in comparison to intention surveys (Juster, 1966).

This section begins by discussing the Juster mean calculation and examines how the Juster scale was developed. Following this, an overview of the research that extends the scale to various products and services is discussed. Finally, the strengths and weaknesses of the metric are examined.

4.2 How to calculate the Juster Mean

The exact wording of the Juster scale question depends heavily on the purchase behaviour and category being measured. Wright, Sharp, and Sharp (2002) outline the general formation of the Juster scale question, “Now, using the Juster scale, and taking everything into account, what are the chances that you, personally, will <buy/shop at> <brand j> in the next <period>” (p. 84). Responses are indicated on an 11-point scale ranging from 0 to 10, with 0 indicating “no chance, almost no chance (1 in 100)” and 10 indicating “certain, practically certain (99 in 100)”. As shown in Figure 1, each scale point is indicated by “point descriptors that were both qualitative and quantitative” (Day, Gan, Gendall & Esslemont, 1991 p. 21).

Figure 1: The Juster scale

10	Certain, practically certain	(99 in 100)
9	Almost sure	(9 in 10)
8	Very probable	(8 in 10)
7	Probable	(7 in 10)
6	Good possibility	(6 in 10)
5	Fairly good possibility	(5 in 10)
4	Fair possibility	(4 in 10)
3	Some possibility	(3 in 10)
2	Slight possibility	(2 in 10)
1	Very slight possibility	(1 in 10)
0	No chance, almost no chance	(1 in 100)

Source: Day, D., Gan, B., Gendall, P., & Esslemont, D. (1991). Predicting purchase behaviour. Marketing Bulletin, 2(3), 18-30.

The Juster mean is calculated from the sample of respondents who report their probability of purchase. The mean provides an estimate of the average proportion of a population that will purchase the particular brand or product in question. To calculate this, each scale point's corresponding probability of purchase is multiplied by the number of respondents that have selected that scale point. For example, if 30 respondents select "Probable (7 in 10)", then 0.7 is multiplied by 30. Once this is calculated for each scale point, the values are then summed and divided by the number of respondents in the sample to calculate the Juster mean. Therefore, if the Juster mean for a particular brand is calculated as 0.62, it is estimated that 62% of the population will purchase this brand during the stated timeframe.

An important aspect of the Juster scale is that it includes an 11-point scale ranging from 0 to 10, similar to the Net-Promoter scale. As it has been found that there is little predictive distinction between 11-point likelihood and probability scales (Wright & MacRae, 2007), a Juster mean can be applied to the Net-Promoter question. This is done by applying the Juster probability scale to the Net-Promoter likelihood scores. For example, if a respondent reports that their likelihood to recommend a brand is a 6, then by applying the Juster scale their likelihood of recommending is 0.6. Thus, a likelihood mean can be calculated following the same procedure used to calculate the Juster mean. Consequently, if the likelihood mean is calculated to be 0.4 for a particular brand, then it is estimated that around 40% of the population is likely to recommend this brand to a friend or colleague.

Due to similarities between 11-point probability and likelihood scales, a likelihood mean is able to be calculated from Net-Promoter data to determine whether it is more predictive of company growth in comparison to the calculation of NPS. However, before this is carried out an evaluation of the Juster scale is necessary to determine whether it can accurately predict purchase behaviour.

4.3 The Development of the Juster Scale

Due to mixed performances of purchase intentions and the amount of actual purchases made by respondents with no intention to purchase, researchers have investigated measuring purchase probabilities and their ability to predict actual purchase rates. The Juster scale was developed following two earlier purchase probabilities studies carried out by Byrnes (1964) and Ferber and Piskie (1965). These purchase probability studies undertaken during the 1960s primarily focused on major consumer durables, such as automobiles and took place in the United States.

One of these early studies carried out was the Consumer Savings Project conducted in St Louis, Missouri, in 1958–1959 that investigated the application of subjective probabilities to purchase intentions (Ferber & Piskie, 1965). In total, 250 households were asked whether they planned to purchase various goods or services during a 6-month period. Respondents were then asked to indicate how likely they were to purchase those goods or services on a "plan-o-meter" card. The "plan-o-meter" card consisted of a 0–10 scale, with only three verbal descriptors (0 = no plans at all, 5 = fifty-fifty and 10 = certain). However, this exploratory study found limited support for the use of purchase probabilities over buying intentions, as they did not provide any predictive ability above what is acquired in purchase intention surveys (Ferber & Piskie, 1965). However, Juster (1966) criticises the Consumer Savings Project and provides two reasons for the lack of distinction between purchase probabilities and purchase intentions. Firstly, Juster identifies that only three verbal descriptors were included on the scale, causing a strong trimodal distribution at these points, therefore a strong mid-point bias exists. Additionally, respondents were asked their plans to buy, which is very similar to intentions, with the only difference being that they were provided with a response scale to indicate their plans (Day et al., 1991; Juster, 1966). Respondents were therefore not actually asked their probability of purchase.

The second study, from which the Juster scale is formed, known as the Detroit Experiment, was conducted by Byrnes (1964). Respondents from 192 households were asked about their chances of purchasing an automobile, and other durable goods over a 6-, 12- and 24-month period (Juster, 1966). As with the experiment by Ferber and Piskie (1965), respondents were given an answer sheet containing an 11-point probability scale from which to select their answer (see Fig. 2). However, the Detroit Experiment contained verbal labels for every scale point, in contrast with only three that were used in the Consumer Savings Project. Importantly, it was discovered that a substantial number of households that reported no buying intentions selected a purchase probability greater than zero and actually purchased a product (Day et al., 1991; Juster, 1966). Furthermore, according to Juster (1966), although the Detroit Experiment has an improved distribution compared with the Consumer Savings Project, strong trimodal distribution with peaks was still evident at 0, 5, and 10. Juster suggests the reason behind the midpoint peak (about even chance) is attributed to points 2 through to 8 being verbally conveyed relative to the midpoint (Day et al., 1991; Juster, 1966).

Figure 2: The Detroit Experiment probability scale

10	Absolutely certain to buy	10
9	Almost certain to buy	9
8	Much better than even chance	8
7	Somewhat better than even chance	7
6	Slightly better than even chance	6
5	About even chance (50-50)	5
4	Slightly less than even chance	4
3	Somewhat less than even chance	3
2	Much less than even chance	2
1	Almost no chance	1
0	No chance	0

Source: Day, D., Gan, B., Gendall, P., & Esslemont, D. (1991). Predicting purchase behaviour. Marketing Bulletin, 2(3), 18-30.

Juster (1966) conducted multiple experiments at the US Department of the Census through Quarterly Surveys of Intentions (QSI) to test whether a probability scale was better than intention scales at predicting purchase rates. Purchase probabilities were obtained for automobiles and a variety of other durables from 451 personal interviews over periods of 6, 12, and 24 months (Day et al., 1991). In contrast to the Detroit Experiment, the QSI Experiment used different qualitative and quantitative descriptors for each scale point (see Fig. 1). The change in the scale descriptors resulted in a much smoother distribution of purchase probabilities shaped like an inverse *J*, with peaks at 0 and 1.0 and a trough in between (Juster, 1966; Day et al., 1991).

By avoiding strong trimodal peaks, the Juster scale was able to develop construct validity, as respondents were more likely to give accurate estimates of their probability of purchase (Wright et al., 2002). Developing construct validity allows the Juster scale to predict actual purchase rates from purchase probabilities more accurately. Furthermore, comparisons with actual purchase behaviour allowed the Juster scale to evaluate predictive validity (Wright et al., 2002). By comparison with actual purchase behaviour, the QSI Experiment found that the Juster scale was more accurate in predicting household purchase rates of automobiles compared with other durables (Juster, 1966; Day et al., 1991). For automobile purchases, probability surveys were found to improve predictions of household purchase rates, relative to intention surveys (Juster, 1966).

Purchase probability studies during the 1960s have highlighted that many non-intenders have a purchase probability above zero and that purchase probabilities were more predictive than intention surveys in predicting purchases of automobiles. Nonetheless, due to the large focus on consumer durables in these studies, it is important that similar results are observed with non-durables to ensure purchase probabilities provide an accurate measure across products and industries. Therefore, it is imperative that the extensions of the Juster scale are discussed.

4.4 Extensions of the Juster Scale

Following its development, the Juster scale has been extended, with varying degrees of success, to predict the purchase rates of a variety of goods and services (Brennan & Esslemont, 1994; Brennan, Esslemont & Clarence, 1995; Clawson, 1971; Day et al., 1991; Gan, Esslemont & Gendall, 1986; Hamilton-Gibbs, Esslemont & McGuinness, 1992; Juster, 1966). In all these studies, purchase probabilities are collected and used to predict the proportion of households or individuals that will likely make a purchase over a future period.

The Juster scale was successfully applied to predict the purchase rates of services (Clawson, 1971; Gan et al., 1986). Clawson (1971) found that in a variety of consumer services such as travel, savings, investment, and recreation, the Juster mean provided the most reliable prediction of actual purchase behaviour over a 3-month time period. Furthermore, Gan et al. (1986) discovered that the mean probability scores, obtained using the Juster scale, were highly correlated with actual purchases of multiple services. However, in contrast to Juster's (1966) findings, predicted purchases were overestimated compared with actual purchase.

Additionally, studies have explored the use of purchase probabilities, obtained using the Juster scale, to predict actual purchase rates of fast-moving consumer goods (FMCGs). With similar results to the services studied, Gan et al. (1986) discovered that the Juster scale was able to predict accurately the purchase rates of an LP record, a pair of shoes, and a hard cover book. Additionally, Brennan and Esslemont (1994) found from a survey of 417 households in Palmerston North, that the purchase probabilities obtained in the Juster scale allow for reasonably accurate predictions of purchase rates and market shares of branded soup and yogurt. These studies highlight that the Juster scale is an accurate predictor of actual purchases of FMCG, with only slight systematic overestimation of predicted purchases observed.

The application of the Juster scale in predicting purchases of FMCGs is important as these products are purchased more frequently than durables. Despite it being sufficient to measure only the purchase penetration of durables, it is also important to measure the purchase frequency for FMCGs as they can be purchased multiple times in a period (Hamilton-Gibbs et al., 1992). To address this issue, Hamilton-Gibbs et al. (1992) tested two methods – the Constant Sum Method and the Multiple Question Method –, to determine whether the Juster scale could be used to estimate multiple purchase quantities for FMCGs. The main grocery shopper for each household was asked to estimate the household's purchases of seven regularly purchased supermarket products over a period of 1 month. The Constant Sum Method was found to be more accurate than the Multiple Question Method at estimating actual purchase frequency. The Constant Sum Methods provides respondents with a board of various FMCGs, each with a row of twelve squares indicating the purchase frequency ranging from zero to twelve. For each product, respondents are required to place 10 counters, each representing a probability of 0.1, to indicate their probability of purchasing various frequencies of the product. By altering the Juster scale using this method, it was discovered that purchases frequency of FMCGs can be accurately predicted.

Following this, Brennan et al. (1995) applied the Constant Sum Method and Multiple Question Method to predict the purchase penetration and purchase frequency at various price points for two branded FMCGs, Coca Cola, and Campbell's Red and White Label soup. Brennan et al. (1995) discovered that the estimated purchase rates for Coca Cola were very accurate. In contrast, predictions made for Campbell's Red and White Label soup were inaccurate. While this studied only included two products, it clearly demonstrates that discrepancies in the performance of the Juster scale exist when attempting to measure the purchase frequency of FMCGs. However, further research across multiple brands is needed to ascertain the extent of this.

Nonetheless, the versatility of the Juster scale is well known, as the measure has been extended successfully to predict purchase rates for various time periods, ranging from 3 months to a year (Day et al., 1991). For non-durables, such as services and FMCGs, the scale can be applied to shorter 3-month periods (Clawson, 1971; Gan et al., 1986). Furthermore, the purchase probability scale has been administered through personal interviews, telephone surveys (Brennan, Hini & Esslemont, 1994; Wright et al., 2002) and self-completion surveys (Day, 1987).

More recently, Wright et al. (2002) demonstrated that Juster-based estimators performed very well as substitutes for panel data used in the Dirichlet model. Introduced by Goodhardt, Ehrenberg, and Chatfield (1984), the Dirichlet model is a stochastic model of purchase incidence and brand choice. The model describes how frequently bought branded consumer products are purchased when the market is stationary and unsegmented. The inputs required for the Dirichlet model can be estimated from category penetration, purchase frequency, and market share, which are derived from consumer panel data (Wright et al., 2002). It was discovered that there were high correlations between the Juster scale estimates of each input and measures obtained from consumer panel data (Wright et al., 2002). Panel data is not readily available for certain product categories and markets and if needed, can be expensive and time-consuming to acquire (Uncles & Lee, 2006). The findings by Wright et al. (2002) indicate that when panel data is not available or too expensive to obtain, the Dirichlet model can be satisfied through a Juster-based survey.

These Juster scale studies that have used purchase probabilities to predict actual purchase behaviour, have demonstrated the versatility of the measure. The Juster scale has been shown to predict purchase rates of a variety of products and service to varying degrees of success. Additionally, the measure is able to be administered through a variety of methods, over various time periods and can even replace panel data in the Dirichlet model. While some researchers have found that purchase probabilities over- and under-estimate actual purchase rates, they have been found to perform better then intention surveys. These strengths and weaknesses, among others, are discussed in more depth below.

4.5 Strengths of the Juster Scale

As discussed previously, an important strength of the Juster scale is its ability to provide reasonably accurate predictions of actual purchase rates across a variety of goods and services (Brennan & Esslemont, 1994; Brennan et al., 1995; Clawson, 1971; Day et al., 1991; Gan et al., 1986; Hamilton-Gibbs et al., 1992; Juster, 1966), across various time periods (Clawson, 1971; Day et al., 1991), and by using various administration methods (Brennan et al., 1994; Day, 1987). This indicates that the metric can be used in majority of industries to predict purchase rates. Additionally, the ability of the scale to be used in self-completion surveys allows for an easier and quicker collection of Juster data, as opposed to face-to-face interviewers.

The Juster scale has been found to outperform surveys of purchase intentions in predicting actual purchase behaviour (Clawson, 1971; Gan et al., 1986; Juster, 1966). To consolidate findings in this area, Wright and MacRae (2007) conducted a meta-analysis of studies that examined the accuracy of the Juster scale and purchase intentions in predicting actual purchase proportions from the same respondents. The purchase intention data from nine studies included 144,821 responses, while the purchase probability data from thirteen studies included 34,122 responses. This research concluded that purchase probabilities have smaller confidence intervals and provide a more accurate measure than purchase intention surveys in predicting purchase behaviour (Wright & MacRae, 2007).

Furthermore, the Juster scale removes an issue that has plagued intention surveys. Marketing academics acknowledge that a large proportion of purchases are made by consumers who report no intention to buy, contributing significantly to the poor predictive performances of intention scales (Juster, 1966, Wright & MacRae, 2007). This is because many non-intenders actually have a purchase probability above zero and subsequently make purchases. Theil and Kosobud (1968) confirmed this when analysing quarterly surveys of car purchases from the United States Bureau of the Census data. Respondents who had stated no buying intention made 70% of the car purchases, while less than 40% of the intenders actually purchased a car (Theil & Kosobud, 1968). However, the Juster scale has been found to outperform surveys of purchase intentions at discriminating between those who will buy and those who won't buy, therefore, providing more accurate predictions (Wright & MacRae, 2007). Supporting this, Juster (1966) discovered that 32 automobile purchases were made from 300 respondents who reported no intention to buy. Of these non-intenders, only 8 out of 180 automobile purchases were made by respondents who reported a zero purchase probability, as opposed to 24 out of 120 non-intenders who reported a non-zero probability (Juster, 1966). Therefore, this study demonstrated that purchase probabilities, obtained using the Juster scale, can more accurately distinguish between those that will and will not purchase.

In contrast to strong trimodal peaks in previous purchase probabilities studies, construct validity exists in the Juster scale as strong peaks at the 0, 5 and 10 points are minimized (Wright et al., 2002). This indicates that the qualitative and quantitative point descriptions on the 11-point scale allow respondents to more accurately specify their probability of purchasing a good or service.

As mentioned earlier, a further strength of the Juster scale is its ability to serve as a substitute to panel data used in the Dirichlet model (Wright et al., 2002). The Dirichlet model requires estimations of category penetration, purchase frequency, and market share, which can be calculated from probability data, obtained using the Juster scale. The ability to use the Juster scale as a substitute for panel data allows the Dirichlet model to be used in industries which panel data is unavailable. With the unavailability of panel data for certain products or industries, Juster data provides a cheaper and less time-consuming alternative (Uncles & Lee, 2006).

The Juster scale has demonstrated a strong predictive ability of actual purchase rates, providing a useful tool for academic and commercial market researchers. The scale improves on alternative measures of purchase intentions as purchase probabilities can more accurately differentiate between those that will and will not purchase. Additionally, in contrast to earlier measures of purchase probabilities, the Juster scale avoids a mid-point bias and has the ability to be used in the Dirichlet model. For these reasons, the Juster scale has been used in market research for nearly half a century. However, these strengths must be contrasted with its weaknesses before determining whether it is appropriate to apply the Juster scale to Net-Promoter data to calculate a likelihood mean.

4.6 Weaknesses of the Juster Scale

One of the biggest concerns of the Juster scale is that it frequently over- and under-estimates purchase rates, despite outperforming purchase intention surveys at estimating actual purchase rates. Some studies have witnessed an over estimations of actual purchase rates from Juster data (Brennan & Esslemont, 1994; Clawson, 1971; Gan et al., 1986; Hamilton-Gibbs et al., 1992; Juster, 1966), while under estimations have also been observed (Hamilton-Gibbs et al., 1992; Juster, 1966).

The overestimation of purchase rates has been observed for a variety of products and service across different time periods. For example, it was discovered that the purchase rates of two FMCGs, canned soup and yogurt, were overestimated over a

4-week period by 5 and 6% respectively (Brennan & Esslemont, 1994). Additionally, purchase rates were overestimated for each particular brand of canned soup and yogurt by between 4.8 and 13.6% (1994). Another study in the FMCG category found purchases were overestimated when the Constant Sum Method was used to account for multiple purchases by a consumer in a period (Hamilton-Gibbs et al., 1992).

Furthermore, Clawson (1971) found that purchase rates of various services over a 3-month period were overestimated. Overestimations in the purchase of services were also discovered by Gan et al. (1986). This study also found that purchases of two FMCG, an LP record and hard-cover book, were overestimated despite a pair of shoes being underestimated. Additionally, durable goods, including automobiles, were found to have overestimated purchase rates (Gan et al., 1986). However, the overestimation of automobile purchase was very slight, with 17 predicted and 16 actually brought.

The overestimation of goods and services, especially automobiles, contrasts with Juster's (1966) original finding of 6.7% more automobile purchases than was estimated by using purchase probabilities. For FMCGs, actual purchases were underestimated when the Multiple Question Method (Hamilton-Gibbs et al., 1992).

However, a meta-analysis of purchase probability data from 13 studies and 34,122 responses found that the under- and over-estimations of actual purchase rates is the result of random variations around an unbiased mean (Wright & MacRae, 2007). It was shown that the distribution around the mean is relatively small and follows a normal distribution. This research highlights that previous findings of under- and over-estimations in purchase rates are attributed to small sample sizes or studies incorporating new products (Wright & MacRae, 2007). Therefore, this study identifies that researchers can rely on the accuracy of the Juster scale to predict purchase rates.

Studies highlighting the over- and under-estimations of actual purchase present a significant threat to the accuracy of Juster scale estimates. Despite improvements in intention surveys, measuring purchase probabilities still requires respondents to forecast their behaviour over a certain time frame. Therefore, it is difficult for respondents to include in their purchase probabilities unexpected events, such as the breaking down of a washing machine forcing an unintended purchase. Additionally, for FMCGs in particular, it is difficult for respondents to estimate when they will be in a buying situation. However, Wright and MacRae (2007) have shown that the Juster scale can accurately predict purchase proportions for existing products with random variation existing around an unbiased mean. Despite varied results in the predictive ability of the Juster scale, it is shown to be more accurate compared with purchase intention scales, and the variation around the mean is smaller than previously thought. Therefore, it is an important measure for those concerned with consumer behaviour research.

4.7 Juster Scale Summary

The earlier studies in this chapter describe the development of the Juster scale and highlight the initial discovery that the Juster scale provides a better prediction of automobile purchases compared with purchase intention surveys (Juster, 1966). Since this discovery, despite some over- and under-estimations, the Juster scale is found to be a reasonably accurate estimator of actual purchase rates of a variety of goods and service, over various time periods and through different administration methods (Brennan & Esslemont, 1994; Brennan et al., 1995; Clawson, 1971; et al., 1991; Gan et al., 1986; Hamilton-Gibbs et al., 1992; Juster, 1966).

One of the biggest advantages of the Juster scale over purchase intention surveys is the scale's ability to differentiate between those who will and won't buy (Wright & MacRae, 2007). Its predictive ability across industries and the fact that it can be used as an alternative to panel data in the Dirichlet model are significant strengths of the metric (Wright et al., 2002).

Some studies have found that the Juster scale over-estimates actual purchase rates (Brennan & Esslemont, 1994; Clawson, 1971; Gan et al., 1986; Hamilton-Gibbs et al., 1992; Juster, 1966), while other studies have discovered under-estimates

(Hamilton-Gibbs et al., 1992; Juster, 1966). Still, the ability for measures to predict actual consumer behaviour is always going to result in discrepancies. Nonetheless, the Juster scale outperforms purchase intention surveys and for this reason this research seeks to apply the Juster mean calculation to NPS data to allow the calculation of a likelihood mean. Additionally, research has found that over- and under-estimations provided by the Juster scale are relatively small and occur randomly around an unbiased mean (Wright & MacRae, 2007). Therefore, researchers are able to rely on the accuracy of the Juster scale to predict actual purchase behaviour.

Importantly, the Juster scale has 11 scale points, exactly the same number as the Net-Promoter scale. This allows the scale probabilities in the Juster scale to be applied to each scale point in the Net-Promoter scale, thus allowing a likelihood mean to be calculated. This measure, along with the NPS, will be compared with the Polarization Index, which is discussed in the next chapter.

5.0 POLARIZATION INDEX

5.1 Introduction

The previous chapter discussed the 11-point Juster scale and the appropriateness of calculating a likelihood mean from 11-point Net-Promoter data, similar to the Juster mean calculation. This chapter discusses the Polarization Index, an alternative loyalty measure that will be applied to the 11-point Net-Promoter data in this research. The Polarization Index (φ) was first introduced by Sabavala and Morrison (1977) to describe television viewing behaviour, and later used to investigate variations in loyalty amongst a range of FMCGs.

The Polarization Index was developed to provide a more consistent measure of loyalty by excluding market share from its calculation (Corsi, Rungie & Casini, 2011). Researchers often use market share to determine loyalty as high market share brands tend to have more buyers with a greater purchase frequency (Ehrenberg et al., 1990). However, research has discovered that loyalty for “niche”, “change-of-pace” and high market share brands can differ from what is typically observed by the pattern of double jeopardy (Fader & Schmittlein, 1993; Kahn Kalwani & Morrison, 1988). The variations in loyalty not accounted for by the pattern of double jeopardy signalled the need for an alternative loyalty measure independent of market share. The Polarization Index has potential to fill the need for a more accurate loyalty measure for “niche”, “change-of-pace”, and high market share brands.

Since its introduction, the majority of research has focused on testing the ability of the Polarization Index to predict variations in loyalty in the context of the wine industry (Casini, Rungie & Corsi, 2009; Corsi et al., 2011; Jarvis & Goodman, 2005; Jarvis, Rungie & Lockshin, 2007; Krystallis & Chrysochou, 2010). Other research has used the Polarization Index to predict variations in loyalty for dairy products (Krystallis & Chrysochou, 2010) and cigarettes (Krystallis, 2013).

However, little research has applied the Polarization Index in the context of the services industry or compared it against Net-Promoter data. This research applies the Polarization Index in the context of the services industry and utilises Net-Promoter data to explore its utility as a loyalty measure. Therefore, before the Polarization Index is applied to Net-Promoter data in the service industry, an overview of the metric is needed.

This chapter provides an outline of how the Polarization Index is calculated and its development. This is followed by an evaluation of the research that has applied the Polarization Index to wine products and other FMCGs. Finally, the strengths and weaknesses of the Polarization index are identified and compared to determine the appropriateness of applying the Polarization Index to the Net-Promoter data.

5.2 How to calculate the Polarization Index (φ)

To recognise how the Polarization Index is applied to Net-Promoter data it is important to understand how it is calculated. To obtain φ, a beta binomial distribution (BBD) model is applied to purchase behaviour data acquired from panel data or a survey (Corsi et al., 2011; Sabavala & Morrison, 1977). A BBD assumes that a consumer has a purchase probability, *p*, between zero and one for a particular brand, with the alternative being not to purchase the brand, *1-p* (Kalwani & Morrison, 1980). The BBD applied in the calculation of φ has two parameters, alpha (α) and beta (β), which are used to calculate the S statistic (Kalwani & Morrison, 1980). The Polarization Index calculation is shown below (Sabavala & Morrison, 1977):

$$\varphi = \frac{1}{1 + S}$$

The Polarization Index can range from zero to one. If φ is close to zero, than little loyalty exists and the BBD is bell shaped, indicating significant brand switching is taking place. In contrast, if φ is close to one, strong loyalty and/or disloyalty exists and the BBD is U shaped, which indicates high repeat purchasing among consumers on the right side of the U. Ultimately, a U-shaped BBD indicates that the Polarization Index is a measure of both disloyalty and loyalty as respondents are distributed at both ends of the continuum (Sabavala & Morrison, 1977).

However, φ commonly ranges between 0.2 and 0.6 for majority of product categories (Rungie, Brown, Laurent & Rudrapatna, 2005). As the typical Polarization Index range is relatively low, it indicates difficulty in achieving a customer base that is solely loyal to one brand. Instead, the range between 0.2 and 0.6 reveals that brand switching generally occurs between brands. This confirms that consumers most likely display divided loyalty and purchase from a repertoire of brands (Sharp et al., 2002).

Purchase probability data obtained from the 11-point Juster scale is commonly used for the inputs into the Polarization Index calculation (Krystallis, 2013; Krystallis & Chrysochou, 2010). Similarly, Net-Promoter data is acquired using an 11-point likelihood scale. Research has shown that there is little distinction between the predictive ability of 11-point likelihood and probability scales (Wright & MacRae, 2007). This similarity between 11-point likelihood and probability scales allows the Polarization Index to be applied to Net-Promoter data and compared against the NPS and likelihood mean.

5.3 The development of the Polarization Index

The Polarization Index was first introduced to describe television viewing behaviour as television ratings failed to measure viewer loyalty (Sabavala & Morrison, 1977). According to Sabavala and Morrison (1977), consumers have a probability of watching a particular television program, *p*, and a probability of not watching it, *1-p*. As consumers only have two available choices, watching or not watching a programme, the probability of both choices sum to one. Therefore, the stochastic process of the probability distributions follows a Bernoulli process, allowing a beta-binomial distribution to be applied (Sabavala & Morrison, 1977). By applying a beta distribution, a U, J, inverse-J or bell shape probability distribution is typically observed (Kalwani & Morrison, 1980; Sabavala & Morrison, 1977).

In the study by Sabavala and Morrison (1977), 504 completed responses were collected from United States households in 1971 and 1973. Respondents were asked to identify how many screenings of a variety of prime-time network television programmes they had watched (Sabavala & Morrison, 1977). This study calculated and compared three measures: the Loyalty Rate, the Average Rating, and the newly developed Polarization Index. The Loyalty Rate represents the repeat-purchase probability of the sample and the Average Rating indicates the proportion of households that viewed at least one screening of a programme averaged over the number of screenings (Sabavala & Morrison, 1977). Sabavala and Morrison (1977) discovered that a strong correlation (0.75) existed between the Polarization Index and Average Ratings for comedy programs. Therefore, higher rated comedy programmes generally had greater loyalty. However, this positive association between the Polarization Index and Average Ratings was not observed in other programme categories, indicating that ratings do not always impact on loyalty (Sabavala & Morrison, 1977). This indicated that the performance of the Polarization Index various significantly across program categories.

The Polarization Index is appealing to marketers as it provides a measure of loyalty that is independent of market share (Corsi et al., 2011; Sabavala & Morrison, 1977). To have a measure independent of market share is important as variations in loyalty exist aside from the regular pattern of double jeopardy (Fader & Schmittlein, 1993; Kahn et al., 1988). In particular, Kahn et al. (1988) discovered loyalty discrepancies in “niche” and “change-of-pace” brands. “Niche” brands are classified as speciality brands usually with low market share, while “change-of-pace” brands are those selected by consumers as they seek variety from a mature brand they usually purchase (Kahn et al., 1988). Research using four product categories – soft drinks, cereals, margarine, and sandwich bags – discovered that “niche” brands had a greater purchase frequency for their market share than expected (Kahn et al., 1988). Thus, “niche” brands have greater loyalty relative to their market share than is typically observed. Conversely, it was found that “change-of-pace” brands have a low purchase frequency relative to their market share (Kahn et al., 1988). The identification of “niche” and “change-of-pace” brands demonstrates that other patterns of loyalty occur that are not currently explained by the pattern of double jeopardy.

Loyalty patterns predicted by the Dirichlet model were also shown to have loyalty discrepancies. Fader and Schmittlein (1993) tested brands with high market share using a variety of supermarket goods across the United States and Japan. They found that some brands with high market share have even greater loyalty than what is predicted from the Dirichlet model. While the reason for this phenomenon is unclear, the findings of Fader and Schmittlein (1993) and Kahn et al. (1988) highlight that market share is sometimes inefficient at determining loyalty.

The variations discovered between loyalty and market share have generated academic interest in the Polarization Index because it is a loyalty measure independent of market share. This academic interest led to a significant body of research investigating the ability of the Polarization Index to predict differences in loyalty amongst brands and product attributes. These extensions of the Polarization Index are discussed in the following section.

5.4 Extensions of the Polarization Index

The Polarization Index was largely extended to the context of the wine industry because it consists of many “niche” and “change-of-pace” brands (Jarvis, Rungie & Lockshin, 2003). A significant body of research has applied the Polarization Index to the wine industry in Australia (Jarvis & Goodman, 2005; Jarvis et al., 2007), Italy (Casini et al., 2009; Corsi et al., 2011) and Greece (Krystallis & Chrysochou, 2010). In the wine category, the Polarization Index was demonstrated to provide an accurate measure of loyalty over time (Corsi et al., 2011). Additionally, the Polarization Index was extended to dairy products (Krystallis and Chrysochou, 2011) and cigarettes (Krystallis, 2013).

In one of the earliest applications in the wine industry, Jarvis and Goodman (2005) investigated whether loyalty varies between different price tiers of wine in Australia. Purchase data over a 1-year period from 4,000 wine shoppers were used in the research. To assess the loyalty among various prices, the wines were grouped into four price tiers measured in Australian dollars, under \$7.49, \$7.50 to \$12.49, \$12.50 to \$17.49, and greater than \$17.50 (Jarvis & Goodman, 2005). Wine priced under \$7.49 and above \$17.50 had the lowest market share of 8% and 27% respectively. Despite this, the greatest loyalty was observed among the price tiers of under \$7.49 (φ=0.44) and over \$17.50 (φ=0.45). Consequently, this research contradicts the usual double jeopardy pattern, as the price tiers with the largest market shares did not generate the greatest loyalty. This illustrates that the Polarization Index uncovered loyalty patterns that other behaviour performance and loyalty measures had missed.

Another study in the Australian wine industry was undertaken by Jarvis et al. (2007). This study tested whether variations in loyalty existed amongst four red wine attributes: brand size, price, region of origin, and variety (Jarvis et al., 2007). The final sample included 2,036 red wine shoppers who had used a loyalty card to purchase wine from a national wine shop chain. These shoppers had purchased a minimum of 10 bottles of wine over a 12-month period and had re-purchased a particular bottle of wine they had bought previously in the period. This research found that red wine shoppers had the greatest loyalty to price and variety, rather than to origin and brand. Similar to the findings of Jarvis and Goodman (2005), consumers were shown to have greater loyalty towards high- and low-priced red wines. Additionally, the Polarization Index differed across red wine varieties, with Shiraz (φ=0.35) and Cabernet (φ=0.38) generating the highest loyalty. The region of origin also impacted on the loyalty among shoppers. In particular, foreign brands had considerably high loyalty (φ=0.37), despite having the smallest market share, indicating that “niche” brands are found among wine imported into Australia (Jarvis et al., 2007). Therefore, this study highlights the usefulness of the Polarization Index as it uncovered new loyalty patterns in the Australian wine market.

The Polarization Index was also applied in the Greek wine industry. Krystallis and Chrysochou (2010) sought to determine whether frequent red and white wine consumers in Greece had varied loyalty towards different product attributes. This study discovered that loyalty was greater for certified white wines than for certified red wines. Additionally, the corporate size of the winemaker generated higher loyalty towards white wines. This demonstrates that the Polarization Index is able to detect which wine attributes generate the greatest loyalty among respondents in the Greek wine industry.

Casini et al. (2009) used the Polarization Index to investigate loyalty among wine attributes in the Italian wine industry. The sample used in this research was obtained from a consumer panel of 5,299 households. Only households who had purchased over ten bottles of wine and had bought wine in two different purchase situations between 2003 and 2005 were included in this sample (Casini et al., 2009). This study investigated the effect price, quality classification, and format (size and type of packaging) had on loyalty to wine in the Italian market. It was found that the format of the wine (φ=0.49) generated the most loyalty, followed by quality classification (φ=0.37) and price (φ=0.29). In particular, wine sizes over 1.5 litres (φ=0.58), equal to 0.75 litres (φ=0.50), table wines (φ=0.46), and wines priced under €3 (φ=0.42) were shown by the Polarization Index as having the highest loyalty (Casini et al., 2009). This study confirms previous research that the Polarization Index is able to determine which wine attributes generate the greatest loyalty amongst consumers in the sample.

Research has also investigated whether the Polarization Index provides a valid longitudinal measure of loyalty in the Italian wine industry (Corsi et al., 2011). This research included a final sample of regular wine consumers from two separate data sets, 3,858 households collected between 2003 and 2005, and 4,643 families collected between 2006 and 2008. Similar to the study by Casini et al. (2009), wines were categorised based on three attributes: price, quality classification, and format. This study confirmed previous findings that format had the greatest loyalty among Italian wine consumers, followed by quality specifications and price (Corsi et al., 2011). Furthermore, it was found that the Polarization Index provides an effective measure for assessing longitudinal changes in loyalty (Corsi et al., 2011). This is an important finding as it demonstrates that the Polarization Index can be applied to longitudinal as well as cross-sectional data. This supports the objective of the present research to apply the Polarization analysis to Net-Promoter data to assess whether it provides a more accurate measure of loyalty over time in comparison to the likelihood mean and NPS.

The Polarization Index was used to investigate loyalty patterns in two dairy product categories, milk and yogurt (Krystallis & Chrysochou, 2011). For this research, purchase probabilities were collected for 20 milk brands and 16 yogurt brands over a 4-week period using the Juster scale. The final sample was obtained in Athens during 2007 through a convenience sample that consisted of 177 respondents for the milk and 167 for the yogurt category. Despite findings of greater loyalty for the milk category (φ=0.23) compared with the yogurt category (φ=0.07), these Polarization Index values are relatively low, indicating low brand loyalty and limited repurchase behaviour (Krystallis & Chrysochou, 2011). Furthermore, by comparing three product health claim attributes, various levels of loyalty were discovered. In particular, it was found that milk consumers have greater loyalty to fat content and ways of processing health claims than to enrichment health claims. Similarly, yogurt consumers have greater loyalty to fat content claims than to enrichment health claims (Krystallis & Chrysochou, 2011). These findings highlight that the Polarization Index is able to distinguish which product attributes generate the greatest loyalty among milk and yogurt consumers in the sample. The major advantage of this research is that it provides a successful application of the Polarization Index outside the wine industry and suggests that the Polarization Index is also useful in FMCG categories.

Another study used the Polarization Index to investigate category and attribute loyalty levels for different cigarette brands (Krystallis, 2013). This research was carried out in Iceland and included a final sample of 155 current smokers who had attempted to quit in the past. Respondents were asked to indicate their probability of purchasing 15 of the highest selling cigarette brands in Iceland using the Juster scale. It was discovered that overall brand loyalty was high (φ=0.72), signifying that respondents have a high repeat purchase rate. Additionally, it was found that the greatest loyalty existed for light cigarette brands, long cigarette length, regular cigarette flavour, and slim cigarettes (Krystallis, 2013).

In summary, these studies have shown that the Polarization Index is a consistent measure of loyalty across countries and the FMCG category. The usefulness of the measure to predict brand loyalty in industries with a high prominence of "niche" and "change-of-pace" brands is demonstrated. Additionally, these studies have highlighted that the Polarization Index is able to measure loyalty levels among various product attributes. These strengths, along with others, are discussed in more depth below.

5.5 Strengths of the Polarization Index

The core strength of the Polarization Index is that it provides a measure of loyalty independent of market share. This is important, as loyalty for "niche", "change-of-pace", and high market share brands differs from what is typically observed in the pattern of double jeopardy (Kahn et al., 1988; Fader & Schmittlein, 1993). Therefore, the Polarization Index provides a measure for these brands that more accurately quantifies their brand loyalty (Kahn et al., 1988).

The Polarization Index is also able to determine loyalty levels among different product attributes (Casini et al., 2009; Corsi et al., 2011; Jarvis & Goodman, 2005; Jarvis et al., 2007; Krystallis & Chrysochou, 2010, 2011; Krystallis, 2013). The ability for the Polarization Index to determine different levels of loyalty among product attributes allows companies to focus on producing and marketing products with attributes likely to attract higher loyalty and purchase frequency.

Another advantage is that the Polarization Index is very easy to understand and interpret as it ranges between zero and one (Krystallis & Chrysochou, 2009). Values close to one indicate high loyalty and that repeat purchasing is likely to occur among consumers. In contrast, values close to zero indicate low loyalty and that regular brand switching is likely to occur among consumers.

The inputs required for the Polarization Index calculation are typically obtained through panel data. However, a Juster scale survey is an appropriate approach to use because research has discovered that purchase probabilities obtained via the Juster scale provide an accurate substitute for panel data (Wright et al., 2002). Since this finding, studies have successfully used the Juster scale estimates as inputs into the Polarization Index calculation to identify variations in loyalty across brands and attributes (Krystallis, 2013; Krystallis & Chrysochou, 2009). The ability of the Polarization Index to be satisfied by Juster scale estimates is a significant advantage as panel data is difficult to obtain in some product categories and countries (Wright et al., 2002), and is typically expensive and time-consuming to obtain (Uncles & Lee, 2006).

A significant strength of the Polarization Index metric is that it provides a reliable assessment of changes in loyalty over time, as discussed in section 5.4 (Corsi et al., 2011). This is important as most companies look to gauge whether their brand loyalty is increasing, decreasing or remaining stable over time. Therefore, companies are able to use this measure to evaluate whether marketing efforts are successful in generating more loyal customers.

Overall, the Polarization Index is independent of market share and able to detect variations in brand loyalty among “niche”, “change-of-pace”, and high market share brands that other measures have not observed. Additionally, the Polarization Index is easy to interpret, can be satisfied by Juster scale purchase probability estimates, is able to determine which product attributes generate the greatest loyalty, and can provide an accurate measure of longitudinal loyalty. However, the Polarization Index has limitations and these need careful consideration when adopting the metric.

5.6 Weaknesses of the Polarization Index

One of the biggest weaknesses of the Polarization Index is the lack of replication across a variety of industries and product categories. Since the metric was first introduced for television viewer choice by Sabavala and Morrison (1977), the majority of research has occurred in the wine industry (Casini et al., 2009; Corsi et al., 2011; Jarvis & Goodman, 2005; Jarvis et al., 2007; Krystallis & Chrysochou, 2010), while some has occurred for other FMCGs such as dairy products (Krystallis & Chrysochou, 2011) and cigarettes (Krystallis, 2013). Despite research in the FMCG product category, no significant research has explored the ability of the Polarization Index to measure loyalty for durable goods and services. This is a major concern of the metric, as without replication, it cannot be assumed that the Polarization Index is an effective measure of loyalty in these industries.

Another weakness of the Polarization Index is that it is not extensively adopted in commercial market research. In comparison, NPS has received significant commercial adoption despite its recent introduction. The lack of commercial use of the Polarization Index is a concern as it does not allow for comparisons to be made across industries, companies, and products.

The Polarization Index accurately detects variations in loyalty among brands with low market share (Kahn et al., 1988). However, for accurate calculations of the Polarization Index, large sample sizes are often needed (Rungie et al., 2005; Kalwani & Morrison, 1980). In particular, research has a discovered that to gain a reliable estimate of the standard error, a sample size of over 1,000 respondents should be used for brands with close to 1% market share (Rungie et al., 2005). Therefore, if panel data is unavailable, data will need to be obtained from a survey, which can be expensive or time consuming for a large sample.

5.7 Polarization Index Summary

This chapter has discussed the development of the Polarization Index as a loyalty measure (Sabavala & Morrison, 1977). Since its introduction, research has highlighted that variations in loyalty exist among "niche", "change-of-pace", and high market share brands from what is typically predicted by market share (Fader & Schmittlein, 1993; Kahn et al., 1988). One of the biggest advantages of using the Polarization Index is that it has the ability to detect variations in brand loyalty because the measure is independent of market share.

In the last two decades, the majority of research investigated the ability of the Polarization Index to detect variations in loyalty in the context of the wine industry in Australia (Jarvis & Goodman, 2005; Jarvis et al., 2007), Italy (Casini et al., 2009; Corsi et al., 2011), and Greece (Krystallis & Chrysochou, 2010). This is because the wine industry consists of many "niche" and "change-of-pace" brands (Jarvis et al., 2003). However, the Polarization Index was also successfully applied to other FMCGs such as dairy products (Krystallis & Chrysochou, 2011) and cigarettes (Krystallis, 2013). In particular, these studies identified that the Polarization Index is able to detect variations in loyalty among product attributes across countries and the FMCG category. It was also discovered that the Polarization Index provides a reliable measure of loyalty over time (Corsi et al., 2011).

Despite these advantages, academic application of the Polarization Index to measure loyalty for durable goods or services is limited. Another weakness of the Polarization Index is that large sample sizes are required to generate accurate loyalty predictions for brands with low market share. Additionally, the Polarization Index is used sparingly in commercial market research and, therefore, data is not available for comparisons to be made across industries, companies, and products.

Notably, the Polarization Index is relatively simple and easy to interpret. Additionally, purchase probability estimates obtained from the 11-point Juster scale can be used as inputs into the model. Likewise, the Net-Promoter data is obtained from an 11-point likelihood scale. As there are few differences between the accuracy of 11-point likelihood and probability scales (Wright & MacRae, 2007), the Polarization Index is able to be applied to the 11-point likelihood scale used in the collection of Net-Promoter data. Given these findings, this research compares the predictive ability of the Polarization Index with the NPS and likelihood mean calculation.

6.0 RESEARCH OBJECTIVES AND HYPOTHESES

6.1 Objectives of this research

This research intends to explore whether the application of a likelihood mean and Polarization Index to 11-point Net-Promoter data provides a closer association with overall company revenue and spend made by a sample of customers than the NPS calculation proposed by Reichheld (2003). The motivation of this research is the significant academic criticism of the NPS calculation and the limited support for its ability to predict company growth rates.

The first aim of this research, therefore, is to determine whether the three loyalty measures, NPS, likelihood mean and Polarization Index, discriminate or replicate one another when applied to 11-point likelihood data obtained from the Net-Promoter question. This phase of the research corresponds to hypotheses one to three.

The second aim of this research is to determine which of the three loyalty measures is more closely associated with past (t-1), present (t), and future (t+1) spend of the sampled customers and company revenue. As Net-Promoter is a WOM metric, aggregate NPSs are traditionally examined against company revenue. This is because WOM is believed to attract new customers and therefore grow the overall company revenue through increased market penetration. However, it is also beneficial to examine whether customers who provide a high NPS spend more than customers who provide a low NPS, as this may go some way to explaining why NPS is also claimed to be correlated with past revenue growth. Therefore, this research also examines the relationship between the three loyalty measures and past (t-1), present (t), and future (t+1) spend by the sampled customers.

This research compares the three measures with past spend and revenue as Reichheld (2003) actually correlated NPS with past sales data (Keiningham et al., 2007a; Sharp, 2008). Additionally, Keiningham et al. (2007) provided the first cross-industry, longitudinal assessment of NPS against current sales growth. Therefore, this research partially replicates Keiningham et al. (2007) by investigating the ability of the NPS, likelihood mean, and Polarization Index to predict current spend and revenue. Comparisons are also made with future spend and revenue as Reichheld (2003) claims that NPS provides the most accurate indicator of future company growth, and research has correlated NPS with future performance indicators (Van Doorn et al., 2013). This phase of the research corresponds to hypotheses four to six and eight to ten.

6.2 Research questions and hypotheses

The specific research questions and null hypotheses are:

RQ1: Do the NPS, likelihood mean, and Polarization Index measure different aspects of loyalty?

- **H1**: NPS and likelihood mean are positively correlated.
- **H2**: NPS and Polarization Index are positively correlated.
- **H3**: Likelihood mean and Polarization Index are positively correlated.

RQ2: At the aggregate level, which loyalty metric best predicts each of past, present, and future spend of the sampled customers?

- **H4:** NPS is more positively correlated with past spend (t-1) of the sampled customers than either the likelihood mean or Polarization Index.
- **H5:** NPS is more positively correlated with current spend (t) of the sampled customers than either the likelihood mean or Polarization Index.
- **H6:** NPS is more positively correlated with future spend (t+1) of the sampled customers than either the likelihood mean or Polarization Index.

RQ3: At the individual level, does the NPS category score predict spending growth?

- **H7**: The NPS category score predicts future spend (t+1) of the sampled customers.

<u>RQ4:</u> At the aggregate level, which loyalty metric best predicts each of past, present, and future company revenue?

- **H8:** NPS is more positively correlated with past company revenue (t-1) than either the likelihood mean or Polarization Index.
- **H9:** NPS is more positively correlated with current company revenue (t) than either the likelihood mean or Polarization Index.
- **H10:** NPS is more positively correlated with future company revenue (t+1) than either the likelihood mean or Polarization Index.

7.0 METHODOLOGY

7.1 Introduction

The primary objective of this research is to investigate whether applying alternative methods of analysis to Net-Promoter likelihood data reveal a closer association with average spend of the sampled customers and company revenue than an NPS calculation. The measurement methods used in this research are the calculation of a likelihood mean similar to the mean calculated from the Juster scale (Juster, 1966) and the Polarization Index (Sabavala & Morrison, 1977). The previous chapters have demonstrated that the NPS is extensively adopted in commercial market research, despite significant criticisms of its calculation and limited academic support of its relationship with company growth rates. This highlights the importance of finding an alternative method of analysis to apply to 11-point likelihood data. Given their ability to be applied to 11-point likelihood data and the substantial academic support these metrics have received, this research uses the likelihood mean and Polarization Index as alternative methods.

The calculations used in this research were developed and tested in previous studies (see Table 1). Known formulas were transformed for ease of use into spreadsheet format by Chickamenahali (1999). The Dirichlet S statistic, used in the spreadsheet to compute the Polarization Index is calculated by adding α and β parameters (Kalwani & Morrison, 1980). The likelihood mean calculation presented below was used by Wright et al. (2002); however, the equation does vary slightly depending on how it is operationalised.

Table 1: Calculations of the three loyalty measures

Loyalty metric	Calculation	Foundation studies
NPS	NPS = % of Promoters - % of Detractors	Reichheld (2003)
Likelihood mean	$b_j = (\sum_i p_{ij}/n)$ x 100	Juster (1966)
φ	$\varphi = \frac{1}{1 + S}$	Sabavala and Morrison (1977)

This chapter proceeds with a discussion of the sample obtained in the agricultural and real estate sectors, and an evaluation of the research design that has taken place. This is followed by an outline of the analytical procedure carried out to compare the three loyalty measures in both sectors. Finally, the analytical procedure used to assess the relationship each measure has with average spend of the sampled customers and company revenue in the agricultural sector is discussed.

7.2 Research design and Implementation

This research incorporates two samples, longitudinal data in the agricultural sector and cross-sectional data in the real estate sector. In both data sets, respondents provided individual Net-Promoter data. To keep confidentiality, the companies from which data were collected are referred to as an 'agricultural company' and a 'real estate service company'.

For both sectors, individual Net-Promoter data were collected. This allowed aggregated frequencies for each selected value on the 11-point likelihood scale to be calculated. Yearly frequencies were calculated between 2011 and 2015 for the agricultural company and for all eleven brands studied in the real estate sector. The

Net-Promoter frequencies allowed the NPS, likelihood mean, and Polarization Index to be calculated in both sectors.

The research on behalf of the real estate service company did not include sales data. Consequently, the ability to predict spend and revenue using the three measures is not made with the real estate data. However, the findings in the real estate sector have been included as they supplement the longitudinal comparison of the three measures made in the agricultural sector by providing a cross-sectional comparison. The research comparing the NPS, Likelihood mean, and Polarization Index in the real estate sector is published in an ANZMAC conference proceeding (Mecredy, Feetham & Wright, 2015).

Therefore, only the longitudinal data collected in the agricultural sector are used for the exploratory investigation of whether the likelihood mean, Polarization Index or NPS has the closest associations with customer spend and company revenue growth rates. The company in the agricultural sector operates as a near monopoly. When the Net-Promoter metric was introduced, it was unable to predict growth rates successfully in sectors where a near monopoly company existed (Reichheld, 2003). However, re-examining the predictive ability of NPS for a near monopoly is important for two reasons. First, significant criticisms exist about Reichheld's methodology due to the NPS correlation with past growth rates (Keiningham et al., 2007). Comparing NPS with current and future spend and revenue data may produce alternative findings to that of Reichheld (2003). Second, alternative methods, such as the Polarization Index and the likelihood mean, may provide a more accurate prediction of company growth rates for near monopoly companies. The following sub-sections describe the research design in the agricultural and real estate sectors.

7.2.1 Real estate sector

This data was collected by a commercial market research provider that carried out research on behalf of a real estate service company in New Zealand. The data was obtained through an on-line survey using a commercial panel provider. The on-line survey was built through the Qualtrics survey platform and recruited participants across five regions in New Zealand: Northland, Auckland, Central North Island, Lower North Island, and the South Island.

While there are some concerns that samples obtained from on-line commercial panels contain potential biases and may not represent the overall population, recruitment bias is minimised as the commercial panel consists of over 75,000 members. Likewise, coverage bias is likely to be minimal as internet access in New Zealand is over 90% of the population (Gibson, Miller, Smith, Bell & Crothers, 2013). To further ensure the sample was representative of the overall New Zealand population, region, age, and gender quotas were applied by the panel provider. This ensured the sample consisted of mixed demographics similar to the overall New Zealand population.

Another concern with on-line panels is that they may contain respondents who only complete the survey for the associated reward and do not answer the survey questions properly. However, research found that the provision of rewards for participation does not affect response quality and survey outcomes (Göritz, 2004).

Five on-line surveys, one for each region, were undertaken in July 2014. The South Island survey was released first as a method of pre-testing the survey. When no issues were discovered with the survey responses and implementation, all other regions were released.

The survey began by showing selected participants an introduction page that explained the survey and procedure, and provided assurance of anonymity. Following this, participants were asked whether they or a member of their close family work in the real estate sector. If participants selected yes, they were thanked for their participation and did not follow through into the rest of the survey. This ensured that the respondent bias in the sample was reduced. The next two filter questions asked participants to select their age and gender category and the region in which they lived. For example, for the Central North Island survey, participants were asked whether they lived in Gisborne, Whanganui, Manawatu, Wairarapa, Taranaki or other region. Once quotas were reached for the age and gender category and for each region, additional respondents were thanked for their participation and did not follow through to the next section of the survey.

The main body of the survey was split into seven sections. Relevant to this research, section four asked participants their likelihood to recommend six different real estate brands. The six real estate brands tested differed in each region. Normally, to calculate a company's NPS, the "would recommend" question is only asked of current customers (Reichheld, 2003). This is because current customers have direct experiences with a company and subsequently are able to form opinions as to whether they would or would not recommend the company to a friend or colleague. Conversely, including non-customers could significantly bias results as they are unfamiliar with the particular company and cannot form an opinion on their likelihood to recommend. However, the six brands tested were well-known real estate brands regularly advertised across multiple forms of media in the region where the participants live. Additionally, the majority of the sample had at least one interaction with a real estate company over the past 5 years. Therefore, given the prominent nature of the brands and the extensive real estate experiences of the sample, respondents are likely to have sufficient knowledge of each brand to determine their recommendation likelihood.

Participants viewed the six brand logos in a randomised order and were asked, "On a scale of 0–10, how likely are you to recommend 'brand X' to a friend or colleague?" Participants were asked to indicate their response on a 0–10 numerical scale that included verbal descriptors at 0 "not likely at all" and 10 "extremely likely". The brand logos shown to the participants were placed on a background image of a house. The house image was the same for each brand and the logo size was similar for each brand shown. In total, eleven brands were assessed across New Zealand with sample sizes ranging from 146 to 1,818.

In the final section of the survey respondents were presented with three additional demographic questions, the ethnicity of the participant, their household income before tax, and the year they were born. This section, combined with the demographics requested in the screening questions, allows for an evaluation of whether the sample represented the overall New Zealand population.

7.2.2 Agricultural sector

The Net-Promoter question was included in a satisfaction survey completed by current customers of an agricultural company over a 5-year period from 2011 to 2015. The survey was conducted in-house by the company's market research team. Within the survey, customers were asked to indicate their likelihood of recommending the company to a friend or colleague on the 11-point Net-Promoter scale.

The data provided by the agricultural company included individual customers' NPSs, along with the corresponding year that the score was collected. Additionally, subject to availability, an indexed yearly spend by each customer was provided between the period 2009 to 2014. No demographic information was provided by the company in the agricultural sector. However, it is assumed that the sample closely resembles the overall customer base of the company.

The data obtained from the company in the agricultural sector was cleaned to ensure all respondents with missing NPSs were removed. Additionally, respondents who gave their NPS but had no corresponding spend data were also removed. Therefore, providing the customer spend was recorded for at least 1 year and the participant had provided a NPS in one of the 5 years data was collected, they were included in the final sample for at least one of the analyses. The final sample comprised of 2,785 respondents with yearly samples ranging from 317 to 965 respondents.

7.3 Phase 1: Comparison between measures

To address the first research question of whether the NPS, likelihood mean, and Polarization Index measure different aspects of loyalty, the three measures were compared against one another using a bivariate correlation. This allowed for the strength of the relationship between the three measures to be examined. The bivariate correlation was computed using aggregate NPS, likelihood mean, and Polarization Index between 2011 and 2015 for the agricultural company and for eleven brands studied in the real estate sector. Therefore, the three measures were compared using cross-sectional Net-Promoter data in the real estate sector and longitudinal Net-Promoter data in the agricultural sector.

7.4 Phase 2: Relationship with past, present and future economic indicators

This phase of the research attempts to answer research questions two to four and determine which of the three loyalty measures are closely associated with and best predicts past, present, and future average spend by the sampled customers and company revenue. To address these questions, the NPS score was correlated against the past, present, and future average spend of the sampled customers and company revenue. This analysis was repeated for both the likelihood mean and Polarization Index.

Yearly spend by each customer was provided in index form between 2009 and 2014 for the company in the agricultural sector. This allowed for a past (t-1), present (t), and future (t+1) average spend by the sampled customers to be calculated and used in the aggregate analysis for each year the Net-Promoter data were collected. To further assess the relationship between NPS and future spend of the sampled customers, multiple regression analysis was carried out at an individual level. Additionally, the company revenue data used were obtained from the agricultural company's annual reports and used at an aggregate level. The results from both phases of this research are presented in the next chapter.

8.0 RESULTS

8.1 Introduction

This section discusses the findings of the analysis carried out in the agricultural and real estate sectors. First, the demographic characteristics were assessed. In the agricultural sector, no demographic characteristics were provided by the company. However, it is assumed that the sample characteristics closely align with their overall customer base.

The demographic characteristics of the sample in the real estate sector were relatively mixed (see Appendix A – Sample demographics). In the sample there was a slight over representation of females (66%) compared with males (34%). Respondents were distributed among age groups from 18 to 34 years (28%), 35 to 54 years (40%), and 55 years and older (32%). The sample did not include respondents under the age of 18 as they are unlikely to have had any previous experiences with buying, selling or renting a house. There was a sufficient variation in both region and annual household incomes among the respondents. The ethnicities contained in the sample were widespread, despite a small under representation of New Zealand Māori.

The sample in the real estate sector also included respondents in a wide variety of living situations (see Appendix B – Sample living situations). In particular, there was a mixture of couples and singles with children living at home, children no longer living at home, and no children. In the sample, 62% of respondents lived in their own home and 30% rented. This indicates that 92% of the sample have had direct contact with one or more real estate agencies in the past and, therefore, are likely to have formed opinions about various real estate brands.

Furthermore, the majority of the sample has had interactions with at least one real estate company in the past 5 years (see Appendix C – Sample real estate experiences). In the last 5 years, 51% of the sample personally signed a real estate rental agreement, or purchase agreement as a buyer or seller. Furthermore, at the time of the survey, 14% of the sample currently owned residential investment properties. This highlights a high familiarity with brands in the real estate sector among the respondents.

After establishing the sample demographics in both data sets, the NPS, likelihood mean, and Polarization Index were compared against one another and their ability to predict past, present, and future customer spend and overall revenue for the agricultural company was explored. The subsequent results are discussed below.

8.2 Comparison of loyalty measures

This section addresses the following research question:

RQ1: Do the NPS, likelihood mean, and Polarization Index measure different aspects of loyalty?

8.2.1 Real estate sector

Table 2 shows the NPS, likelihood mean, and Polarization Index for eleven real estate brands in New Zealand. The sample size for each brand varied between 146 and 1,818 respondents. Among all brands, the number of “promoters” was significantly less than the number of “detractors”, indicated by the NPS ranging from -90% to -50%. The likelihood mean varied from 0.40 to 0.59 and the Polarization varied from 0.16 to 0.28, demonstrating that only slight differences in loyalty occur across real estate brands in New Zealand. Additionally, the likelihood mean appears to show no significant discrimination beyond that offered by NPS, as brands with a higher NPS tended to also have a higher likelihood mean (Mecredy et al., 2015[1]). However, the Polarization Index appeared to show an alternative pattern of loyalty compared with

[1] The results discussed in section 8.2.1 have previously been presented at an ANZMAC conference and will soon be published in the 2015 ANZMAC Conference Proceedings by the author and his supervisors. Therefore, there are similarities between this section and the ANZMAC Conference paper, see Appendix D.

the NPS and the likelihood mean. In particular, the Polarization Index showed that Brands B (φ=0.21), E (φ=0.28), and G (φ=0.24) have the greatest loyalty despite having a weaker NPS and likelihood mean compared with other brands (Mecredy et al., 2015[1]). All three of these brands have a sample size under 250, were only tested in one of the five regions, and have a relatively low market share. This finding is similar to other research that discovered higher than expected levels of loyalty among "niche" brands with low market share (Kahn et al., 1988). However, to gain accurate calculations of the Polarization Index, larger sample sizes are required for brands with low market share (Kalwani and Morrison, 1980; Rungie et al. 2005).

All three measures demonstrate that low brand loyalty exists among the real estate sector in New Zealand. In particular, the low Polarization Index across the brands indicates that respondents are undecided as to whether they will recommend a real estate company or not. This illustrates that significant brand switching is likely to occur among real estate brands. However, this could be a result of the large mid-point bias discovered among the sample. For all eleven real estate brands, between 23% and 32% of respondents indicated their likelihood to recommend the brand to a friend or colleague as a 5, indicated by "neutral" (Mecredy et al., 2015[1]). The Polarization Index associates 50/50 probabilities with brand switching and limited loyalty towards a particular brand. Therefore, the mid-point bias has a substantial influence on the Polarization Index, consequently reducing the loyalty calculated for each brand.

Table 2: Comparison of loyalty measures across brands

	Sample Size, n	NPS	Likelihood mean	φ
Brand A	849	-50%	0.59	0.18
Brand B	233	-62%	0.53	0.21
Brand C	1818	-69%	0.53	0.17
Brand D	1818	-71%	0.52	0.16
Brand E	244	-74%	0.43	0.28
Brand F	736	-75%	0.50	0.17
Brand G	146	-75%	0.43	0.24
Brand H	1428	-79%	0.47	0.18
Brand I	1818	-82%	0.46	0.16
Brand J	1672	-88%	0.40	0.19
Brand K	146	-90%	0.40	0.17

The large mid-point bias in the sample also contributes to the negative NPS witnessed among all brands as respondents who select the mid-point value are classified as “detractors”. Additionally, the mid-point bias contributes to all brands having a likelihood mean relatively close to the mid-point.

Table 3 shows the correlations between the NPS, likelihood mean, and Polarization Index across brands in the real estate sectors. A strong correlation (0.90) is observed between the likelihood mean and NPS, indicating that brands with a high NPS also had a high likelihood mean, and vice versa (Mecredy et al., 2015[1]). This demonstrates that both measures observed a similar pattern in loyalty and for that reason there is no significant difference between which of the two measures is used to evaluate loyalty in the real estate sector within New Zealand. Alternatively, the Polarization Index appears to provide an alternative assessment of loyalty as the Polarization Index had a negative correlation with the likelihood mean (-0.30) and a weak correlation with NPS (0.10). As stated earlier, this could be attributed to the large mid-point bias, as the Polarization Index is more likely to be influenced by 50/50 probabilities. Therefore, the cross-sectional findings in the real estate sector support the hypothesis that the NPS and likelihood mean are positively correlated. Conversely, the hypotheses that the NPS and Polarization Index, and the likelihood mean and Polarization Index are positively correlated are rejected.

Table 3: Correlation coefficients across brands

Correlation	NPS	Likelihood mean	φ
NPS	1		
Likelihood mean	0.90**	1	
φ	0.10	-0.30	1

Note: ** Correlation is significant at the 0.01 level (2-tailed).

8.2.2 Agricultural sector

Table 4 shows the NPS, likelihood mean, and Polarization Index for the agricultural company over a 5-year period. In contrast to the real estate sector, no mid-point bias was identified in the data. NPS ranged from -9% to 19% indicating that the ratio of “promoters” to “detractors” varied across the period. In 2011 the NPS was -9%, highlighting that “detractors” outweighed the number of “promoters”. However, the NPS was positive in the subsequent periods, with a yearly rise observed from 2011 to 2014, before falling in 2015. The likelihood mean ranged from 0.71 to 0.79 and the Polarization Index ranged from 0.15 to 0.25 indicating only slight variations across the 5-year period. Similar to the NPS, the likelihood mean increased over the first 4 years, before decreasing in 2015. This indicates that the likelihood mean and NPS appear to show no significant differences in the discrimination of loyalty, as periods with a higher NPS typically had a higher likelihood mean.

However, the Polarization Index rose over the first year, before diminishing in each subsequent year, indicating that a different pattern of loyalty was observed. Therefore, periods that observed a high NPS and likelihood mean, typically observed a low Polarization Index, and vice versa. All the Polarization Index values are under 0.3, demonstrating that loyalty to the agricultural company is limited and that respondents were generally undecided on whether they would or would not recommend the company to a friend or colleague.

Table 4: Longitudinal comparison of loyalty measures

	Sample Size, n	NPS	Likelihood mean	φ
2011	317	-9%	0.71	0.23
2012	965	11%	0.75	0.25
2013	725	11%	0.76	0.22
2014	377	19%	0.79	0.16
2015	401	12%	0.77	0.15

The correlations between the three loyalty measures for the agricultural company are shown in Table 5. The likelihood mean and NPS have a significantly strong correlation (0.98**), which indicates a linear relationship exists between the NPS and likelihood mean. The similarity between the two measures highlights that there is little difference between each measure in determining loyalty for the agricultural company. On the other hand, the Polarization Index has a negative correlation with

NPS (-0.54) and the likelihood mean (-0.68). This indicates the Polarization Index has a considerably different evaluation of loyalty compared with NPS and the likelihood mean for the agricultural company. Therefore, as with the findings in the real estate sector, these findings support hypothesis one that the NPS and likelihood mean are positively correlated and do not support hypothesis two and three that the Polarization Index is positively correlated with the NPS and likelihood mean respectively.

Table 5: Correlation of loyalty measures across time

Correlation	NPS	Likelihood mean	φ
NPS	1		
Likelihood Mean	0.98**	1	
φ	-0.54	-0.68	1

Note: ** Correlation is significant at the 0.01 level (2-tailed).

8.3 Relationship with sampled customers past, present and future spend

This section addresses the following research question:

RQ2: At the aggregate level, which loyalty metric best predicts each of past, present and future spend of the sampled customers?

8.3.1 Comparison between loyalty measures and sampled customers spend

For each year between 2009 and 2014, the amount spent at the company in the agricultural sector was collected from all respondents in the sample. Using this data, the average past (t-1), present (t), and future (t+1) spend per customer was calculated for each year that Net-Promoter data was collected (see Table 6). The average spend per customer is computed to an index figure to ensure anonymity of the agricultural company is protected.

Table 6: Longitudinal comparison between loyalty measures and average spend

Year	n	NPS	LM	φ	Avg.spend (t-1)	Avg.spend (t)	Avg.spend (t+1)
2011	317	-9%	0.71	0.23	35.80	64.86	41.90
2012	965	11%	0.75	0.25	39.20	43.58	40.96
2013	725	11%	0.76	0.22	41.28	39.10	38.39
2014	377	19%	0.79	0.16	41.17	39.29	-
2015	401	12%	0.77	0.15	39.83	-	-

8.3.2 Relationship with past spend (t-1)

For the agricultural company, the NPS, likelihood mean, and Polarization Index was correlated with average spend from the previous year (t-1). As shown in Table 7, it was found that the previous year's spend had a significantly strong correlation with the NPS (0.93*) and the likelihood mean (0.91*). This indicates that as average spend increases, the NPS and likelihood mean in the following year typically increases. Alternatively, a negative correlation was found between the previous year's spend and the Polarization Index (-0.44), indicating that as the average spend increases, the Polarization Index in the following year typically decreases. Only a slight difference between the NPS and likehood mean is observed, so this analysis provides mixed supports for hypothesis four: that the NPS is not more positively correlated with past spend (t-1) of the sampled customers than the likelihood mean, but it is more positively correlated with past spend (t-1) than the Polarization Index.

8.3.3 Relationship with current spend (t)

A comparison with the current year's spend found that the NPS (-0.96*) and likelihood mean (-0.93) had a strongly negative correlation with the average spend for the current year (see Table 7). This highlights that for current spend to increase, the NPS needs to decrease, demonstrating that "promoters" spend less in the current year at the agricultural company than "detractors". Growth in current spend among the sampled customers therefore appears to come through an increase in "detractors", not "promoters". In contrast, despite not being statistically significant, a moderately positive relationship (0.40) was found between average spend for the current year and the Polarization Index. This indicates that as the Polarization Index increases, the average spend for the current year also increases. Therefore, it was identified that the NPS and likelihood mean provide a strong inverse indicator of current spend (t), compared with the Polarization Index, which appears to be a respectable positive indicator for the agricultural company. Consequently, hypothesis five that NPS is more positively correlated with current spend (t) is rejected.

8.3.4 Relationship with future spend (t+1)

The ability for the NPS, likelihood mean, and Polarization Index to provide an indicator of future spend (t+1) was also assessed (see Table 7). As in the relationship with current spend, the NPS (-0.71) and likelihood mean (-0.76) both had strongly negative correlation with average spend in the following year (t+1). Again, this highlights that reductions in the NPS and growth in the number of "detractors" result in improvements in future spend (t+1) among the sampled customers for the agricultural company. Alternatively, the Polarization Index has a strong relationship with future spend (0.62), indicating that increases in the Polarization Index will likely cause an increase in next year's average spend. Therefore, despite no statistically significant correlations due to the small sample sized used, the Polarization Index appears to be the best positive indicator of future spend (t+1) for the agricultural company. Therefore, hypothesis six that NPS is more positively correlated with future spend (t+1) is rejected.

Table 7: Correlation coefficients with past, present and future spend

Correlations	NPS	LM	φ
Avg. spend (t-1)	0.93*	0.91*	-0.44
Sig. (2-tailed)	0.02	0.03	0.46
N	5	5	5
Avg. spend (t)	-0.96*	-0.93	0.40
Sig. (2-tailed)	0.04	0.07	0.60
N	4	4	4
Avg. spend (t+1)	-0.71	-0.76	0.62
Sig. (2-tailed)	0.50	0.45	0.58
N	3	3	3

Note: * Correlation is significant at the 0.05 level (2-tailed).

Sample size for average spend varied due to missing data.

8.4 Multiple regressions for future spend (t+1)

This section addresses the following research question:

RQ3: At the individual level, does the NPS category score predict spending growth?

To assess further the relationship between NPS and future spend (t+1) by the sampled customers, a multiple regression model was created. This allows multiple variables to predict the average spend in the following year. The variables used in the regression include current spend (t) and the NPS categories of "promoters", "passively satisfied" and "detractors". These variables were used to create three multiple regressions in order to predict the spend in 2012, 2013, and 2014, as shown in Table 8. Unlike the previous analyses that were conducted at the aggregate level with limited data points, this analysis is conducted at the individual level with many more data points.

The variation in the future spend (t+1) explained by the models is relatively consistent across the three regressions, as between 23.6% and 35.5% of variation in the next year's spend are explained by the models. For all three regressions, the F statistic was statistically significant at the 1% level, indicating that the differences observed in future spend are unlikely to be due to random sampling error.

There are only slight variations among the beta coefficient for current spend among all three regressions, ranging from 0.465 to 0.532. Similarly, the beta coefficient for NPS $_{(category)}$ ranged between 0.052 and 0.151, indicating little variation across regressions. For all three regressions, the beta coefficient is higher for the current spend than the current NPS $_{(category)}$, indicating that the current spend has a greater impact on next year's spend than the current NPS $_{(category)}$. However, despite various levels of statistical significance due to the small sample sized used, the current NPS $_{(category)}$ has a positive effect on the future spend (t+1).

Table 8: Multiple regression statistics

					Beta		Sig.	
Spend$_{(t+1)}$	n	Adj. R^2	F	Sig. F	T	NPS $_{(category)}$	T	NPS $_{(category)}$
2012	294	0.236	45.057	0.000	0.465	0.074	0.000	0.161
2013	883	0.260	156.050	0.000	0.498	0.052	0.000	0.081
2014	675	0.355	186.123	0.000	0.532	0.151	0.000	0.000

8.5 Relationship with past, present and future company revenue

This section addresses the following research questions and hypotheses:

RQ4: At the aggregate level, which loyalty metric best predicts each of past, present, and future company revenue?

8.5.1 Comparison between loyalty measures and company revenue

The company revenue for each year between 2010 and 2015 was collected from annual reports for the agricultural company. The revenue has been computed to an index figure to ensure the anonymity of the agricultural company is protected. Between 2010 and 2015, the company revenue increased across every period (see Table 9).

Table 9: Longitudinal comparison between loyalty measures and revenue

Year	n	NPS	LM	φ	Rev. Index (t-1)	Rev. Index (t)	Rev. Index (t+1)
2009	-	-	-	-	-	-	100
2010	-	-	-	-	-	100	121
2011	317	-9%	0.71	0.23	100	121	130
2012	965	11%	0.75	0.25	121	130	142
2013	725	11%	0.76	0.22	130	142	152
2014	377	19%	0.79	0.16	142	152	167
2015	401	12%	0.77	0.15	152	167	-

8.5.2 Relationship with past revenue (t-1)

As shown in Table 10, strong correlations between company revenue from the previous year (t-1) and NPS (0.85) and likelihood mean (0.92*) were discovered for the agricultural company. This indicates that an increase in previous year's revenue will likely result in an increase in the current NPS and likelihood mean. In comparison, the Polarization Index had a strongly negative relationship (-0.84) with past revenue (t-1), demonstrating that increases in the previous year's revenue will likely result in a decrease in the current Polarization Index. As the strongest and statistically significant relationship is found between the likelihood mean and the previous year's revenue, hypothesis eight that NPS is more positively correlated with the previous year's revenue is rejected. However, the NPS still provides a very accurate indicator of past revenue (t-1) for the agricultural company.

8.5.3 Relationship with current revenue (t)

Similar to the relationship with past revenue, strong correlations between company revenue from the current year (t) and NPS (0.70) and the likelihood mean (0.80) were found (see Table 10). This indicates that increasing the number of customers who are likely to recommend the agricultural company to a friend or colleague will increase the current company revenue. Therefore, an increase in current revenue for the agricultural company comes from increasing the number of "promoters" and decreasing the number of "detractors". Additionally, the Polarization Index is a poor predictor of current revenue for the agricultural company as it was found to have a significantly strong negative correlation (-0.91*) with company revenue in the current year (t). The null hypothesis that the NPS is more positively correlated with current company revenue compared with the likelihood mean and Polarization Index is rejected as the strongest relationship is found for the likelihood mean. Despite the rejection of the null hypothesis, the NPS still provides a very accurate predictor of current revenue (t) for the agricultural company.

8.5.4 Relationship with future revenue (t+1)

The correlations between next year's revenue and the three loyalty measures – the NPS, likelihood mean, and Polarization Index – were also evaluated (see Table 10). Likewise, a strong relationship between future revenue (t+1) and the NPS (0.91) and the likelihood mean (0.95*) were found. Again, this highlights that growth in next year's revenue comes from increasing the number of "promoters" and decreasing the number of "detractors". In contrast, a strongly negative relationship between future revenue (t+1) and the Polarization Index (-0.84) was detected for the agricultural company, highlighting that the Polarization Index is a very poor predictor of future revenue (t+1) for the agricultural company. Similar to the past (t-1) and current revenue (t), the likelihood mean is a slightly better predictor of next year's revenue than the NPS. Therefore, hypothesis ten, that NPS is more positively correlated with future company revenue (t+1) than either the likelihood mean or Polarization Index is rejected.

Table 10: Correlation coefficients with past, present and future revenue

Correlations	NPS	LM	φ
Revenue (t-1)	0.85	0.92*	-0.84
Sig. (2-tailed)	0.07	0.03	0.08
N	5	5	5
Revenue (t)	0.70	0.80	-0.91*
Sig. (2-tailed)	0.19	0.10	0.31
N	5	5	5
Revenue (t+1)	0.91	0.95*	-0.84
Sig. (2-tailed)	0.09	0.05	0.16
N	4	4	4

Note: * Correlation is significant at the 0.05 level (2-tailed).

9.0 DISCUSSION AND CONCLUSIONS

9.1 Summary and discussion

Customer loyalty measures are used extensively in commercial and academic market research as companies and academics seek to better understand consumers and their purchasing behaviour. Loyalty metrics are used in commercial market research because there are claims that increases in loyalty relate to increases in company revenue. However, several academic researchers have questioned the accuracy of these loyalty measures.

NPS is one of the most commonly used loyalty metrics in commercial market research because it is simple to use and easy to understand. Numerous companies across the globe, including multi-national corporations such as GE, American Express, and Microsoft, have adopted NPS because of the claims that NPS is the only number needed to measure company growth (Reichheld, 2003) and that a 12-point increase in NPS relates to a doubling of a company's growth rates (Reichheld, 2006b). These claims have even resulted in companies reporting the NPS to investors and using it to determine pay for senior staff (Creamer, 2006).

Despite extensive commercial adoption of the NPS, there is limited academic support for the metric. Only Marsden et al. (2005) found support for the NPS, arguing that a 7-point increase in NPS results in a one percent increase in brand growth. Other cross-sectional and longitudinal studies investigating the ability for the NPS to predict company growth rates have found no evidence to validate the superiority of the Net-Promoter metric (Keiningham et al., 2007a; Morgan & Rego, 2006; Van Doorn et al., 2013).

Additionally, NPS is heavily criticised by marketing academics (East et al., 2011; Grisaffe, 2007; Keiningham et al., 2007a, 2007b, 2008a; Kristensen & Westlund, 2004; Morgan & Rego, 2006; Pingitore et al., 2007; Rego & Morgan, 2004; Sharp, 2008). The revenue growth claims of Reichheld (2003) and Marsden et al. (2005) are disputed as these studies correlated NPS with past growth rates as opposed to current or future growth rates (Keiningham et al., 2007a). Grisaffe (2007) also raised concerns over the clustering of the Net-Promoter scale and the inclusion of only "promoters" and "detractors" in the NPS calculation. Furthermore, Grisaffe (2007) criticises the inclusion of likelihood scores of 5 and 6 in the "detractors" category as these respondents indicated they have a "neutral" or slightly above likelihood of giving a recommendation. These academic criticisms, among others, highlight that the extensive commercial adoption of NPS is a significant issue, given companies are making business decisions based on a metric that may not accurately predict company growth.

The purpose of this research, therefore, is to determine whether applying the likelihood mean (Juster, 1966) and Polarization Index (Sabavala & Morrison, 1977) to 11-point Net-Promoter predicts company growth better than the NPS calculation. To achieve this objective, longitudinal data obtained in the agricultural sector and cross-sectional data obtained in the real estate sector was analysed. The first step of the analysis compared the NPS, likelihood mean, and Polarization Index against each other at an aggregate level to detect whether they measure different aspects of loyalty. Second, the ability of the three measures to predict the past (t-1), current (t), and future (t+1) spend of the sampled customers and company revenue was assessed.

The comparison between the three loyalty measures using bivariate correlations discovered that the NPS and the likelihood mean have a very strong correlation across brands in the real estate sector and across time in the agricultural sector. The relationship between the NPS and likelihood mean was statistically significant in both industries. Therefore, the strength of the relationship demonstrates there is little discrimination between the aspects of loyalty measured by the two metrics.

The comparison of the three loyalty measures also found that the Polarization Index has a very small correlation with NPS and a negative correlation with the likelihood mean in the real estate sector, and has a moderately negative correlation with the NPS and likelihood mean in the agricultural sector. These results indicate that the

Polarization Index measures different aspects of loyalty compared with the NPS and likelihood mean. While a mid-point bias may have slightly affected the results across brands in the real estate sector, similar results are observed across time in the agricultural sector. Therefore, these findings are robust as the discoveries are stable across both cross-sectional and longitudinal data sets and across different sectors.

While it is traditional to test the NPS predictive ability with company revenue at an aggregate level, this research has also compared the NPS, likelihood mean, and Polarization Index with the past (t-1), present (t), and future (t+1) average spend of the sampled customers for the agricultural company. This allows an evaluation of whether loyal customers classified as “promoters” spend more than disloyal customers classified as “detractors”. A correlation with the past (t-1) spend of the sampled customers revealed that statistically strong correlations existed with both the NPS and likelihood mean, whereas a negative correlation existed with the Polarization Index. This indicates that increases in average spend in the past year result in an increase in the NPS and likelihood mean in the current year. This phenomenon is very unusual, as a relationship between the loyalty measure and current and future spend is more typical. This may indicate that as customers increase their expenditure, they become more likely to give recommendations.

In spite of the findings with past spend by the sampled customers, the NPS and likelihood mean both have strong negative correlations with current (t) and future spend (t+1). This demonstrates that customers who indicate a high likelihood to recommend the agricultural company to a friend or colleague reduce their spend in the future compared to customers who are unlikely to recommend the agricultural company. This highlights that loyal customers, classified as “promoters”, become relatively less valuable customers over time. In fact, “detractors” tend to have more growth in spend than “promoters” in the current and following year. This result contradicts research by Reichheld and Sasser (1990), which highlights the financial importance of retaining loyal customers. In particular, this finding contradicts claims that loyal customers pay higher prices and purchase more products from a company and thus spend more (Reichheld & Sasser, 1990; Zeithaml, 2000). In contrast, this research supports research claims that long-term loyal customers do not always spend more (East et al., 2006), they are not a source of growth, and that a focus on acquiring new customers is more beneficial than customer retention (Riebe et al. 2014).

Furthermore, unlike the NPS and likelihood mean, the Polarization Index has moderately positive correlations with current (t) and future (t+1) spend of the sampled customers. Therefore, this reinforces the findings that disloyal customers, those that provide a low NPS, spend more as the Polarization Index responds favourably to consumers at both ends of the 11-point NPS continuum. The relationship the three measures have with current (t) and future (t+1) spend highlight that to increase current and future customer spending, the agricultural company needs to focus on those who have lower recommendation probabilities. These are most likely lighter buyers, and this is consistent with the research of Sharp (2010) that points out the importance of such buyers.

To further assess the ability of the NPS to predict future (t+1) spend by the sampled customers, three multiple regressions were computed at an individual level, as limited data points did not allow this analysis at an aggregate level. The independent variables used to calculate next year's spend were current spend (t) and NPS $_{(category)}$. The results indicated that current spend has a greater impact on next year's spend than the current NPS $_{(category)}$. However, this research discovered that the NPS $_{(category)}$ had a slightly positive impact on future spend (t+1) in all three regressions. This result differs from the aggregate level comparison, which found NPS has a negative relationship with future spend (t+1). At an individual level, therefore, current NPS $_{(category)}$ had a positive effect on next year's spend by the sampled customers, while at an aggregate level the NPS had a negative effect on next year's spend by the sampled customers.

A likely factor influencing these findings is that for the aggregate comparisons only "promoters" and "detractors" are used to calculate NPS, while at an individual level, the NPS $_{(category)}$ includes these two categories along with "passively satisfied" category for the multiple regressions. Thus, the proportion of "passively satisfied" customers may also play an important role in growth in spend for the sampled customers. This is supported by a simple correlation between the individual NPSs and current spend (t) of the sampled customers (see Table 11). It was found that a significantly positive relationship existed between individual NPS and current customer spend for 3 out of the 4 years compared, demonstrating the association between current spend and recommendation, and reinforcing the potential importance of the "passively satisfied" customers for company growth.

Table 11: Correlation coefficients between NPS and current spend (t) at an individual level

Correlations	**NPS (2011)**	**NPS (2012)**	**NPS (2013)**	**NPS (2014)**
Current spend (t)	-0.05	0.15**	0.32**	0.22**
Sig. (2-tailed)	0.43	0.00	0.00	0.00
N	314	929	704	370

Note: * Correlation is significant at the 0.05 level (2-tailed).

Most of the studies exploring the ability of the NPS to predict company growth rates have correlated NPS against aggregated financial indicators such as company revenue or sales growth (Keiningham et al., 2007a; Marsden et al., 2005; Morgan & Rego, 2006; Reichheld, 2003; Van Doorn et al. 2013). This is because Net-Promoter is a WOM mechanism, so growth usually occurs through attracting new customers rather than increasing the amount spent by the current customers. This research has therefore correlated the three loyalty measures with past (t-1), current (t), and future (t+1) company revenue. This analysis discovered a strong relationship between company revenue and the NPS and likelihood mean. Conversely, a negative relationship between company revenue and the Polarization Index was found. In particular, the strong relationship between the NPS and past (t-1) company revenue confirms the results of Reichheld (2003) and Marsden et al. (2005) studies that correlated NPSs with past growth rates.

These results also indicate that the NPS and likelihood mean are strongly correlated with current (t) and future (t+1) company revenue. This produces alternative findings to Keiningham et al. (2007a) and Van Doorn et al., (2013) which did not find support for the use of the Net-Promoter metric over alternative loyalty measures. While this research uses longitudinal data rather than cross-sectional data, it does support Reichheld's (2003) findings that the NPS is an accurate predictor of future company growth.

Despite the positive relationships between past (t-1), current (t), and future (t+1) company revenue and the NPS and likelihood mean, the Polarization Index had a negative correlation with past (t-1), current (t), and future (t+1) company revenue.

Even though the NPS provided an accurate predictor of past (t-1), current (t), and future (t+1) company revenue for the agricultural company, the likelihood mean provided a slightly better prediction. This is likely due to the fact that the NPS calculation categorises individual NPSs into "promoters", "passively satisfied", and "detractors", but then does not incorporate "passively satisfied" in the calculation. This means that individual respondents who give a score of 6 on the Net-Promoter scale are treated the same as those who give a score of 0; while the likelihood mean treats all likelihood scores as individual probabilities that are included in the analysis without grouping them together. These results confirm that while the Net-Promoter question can be argued to predict company growth through measuring the likelihood to give WOM, the analysis can be improved by using a likelihood mean that incorporates and separately treats all data points on the 11-point Net-Promoter scale.

When the Net-Promoter metric was first introduced by Reichheld (2003), he claimed that the NPS could not predict growth rates in industries dominated by a near monopoly. However, the agricultural company used in this analysis is a near monopoly and therefore these results contradict Reichheld's claim. The differences between these results and Reichheld's (2003) could be because this research assessed the performance of NPS across time for one near monopoly company, while Reichheld (2003) tested the performance across brands within a monopoly sector.

Overall, the results indicate that loyal customers are less valuable customers for the agricultural company because they provide less scope of growth in spend for the current (t) and future (t+1) years. However, loyal customers may promote the company to friends and colleagues. Therefore, "promoters" may provide a useful marketing tool to increase company revenue, but this increase in revenue will likely occur through an increase in market penetration by attracting more new light buyers. The mechanism for doing so is presumably the WOM from existing customers; however, this can only be inferred from Reichheld's claims as the WOM is not directly observed in the type of data used in this study.

9.2 Conclusions

This research has confirmed that the NPS and likelihood mean measure similar aspects of loyalty when applied to 11-point Net-Promoter data. However, compared with the NPS and likelihood mean, the Polarization Index appears to measures different aspects of loyalty. These results are well supported as they were stable across both cross-sectional and longitudinal data sets and across different sectors.

The research also suggests that at an aggregate level the NPS and likelihood mean are poor predictors of the current (t) and future (t+1) spend by the sampled customers, compared with the Polarization Index, which provides a more accurate prediction. Therefore, for the agricultural company, "promoters" are less valued customers as they exhibit less spending growth than "detractors" over the current (t) and future (t+1) years.

By using longitudinal data for an agricultural company at aggregate level, this research provides some support for the role of NPS as a predictor of past (t-1), current (t), and future (t) company revenue. The findings of the strong relationship with past (t-1) company revenue support the findings of Reichheld (2003) and Marsden et al. (2005), as these studies correlated NPSs with past growth rates. The strong relationship found with current (t) and future (t+1) company revenue supports Reichheld's (2003) claim that NPS is an accurate predictor of future company revenue. In contrast, these results are inconsistent with previous research by Keiningham et al. (2007a) and Van Doorn et al. (2013), both of whom found the NPS was not an accurate predictor of current and future growth rates.

This research also found that when applied to 11-point Net-Promoter data the Polarization Index is a poor predictor of past (t-1), current (t), and future (t) company revenue, while the likelihood mean is a slightly more accurate measure compared with the NPS. These findings suggest that a likelihood mean may provide a more accurate predictor of revenue growth than the NPS. This is likely due to the fact that the likelihood mean does not cluster the scale points into categories like the NPS does. However, the evidence in support of this point is weak, as the small sample size does not guarantee confidence that these are real differences rather than simple sampling variation.

This research suggests that although loyal customers increase their spending less than disloyal customers, they do contribute to overall growth in company revenue, presumably through WOM advertisement. Essentially, loyal customers of the agricultural company appear to grow company revenue by attracting new customers. This indicates that growth comes through both penetration growth and increasing the spend of the least loyal customers, rather than through increasing the purchase frequency of all customers. The most loyal customers have likely reached saturation point with the company's products, making them unlikely to increase their purchase rate.

9.3 Limitations and Future Research

This research explored new applications of 11-point Net-Promoter data and therefore some limitations and potential areas for future research are identified. The aggregate comparisons between the three loyalty measures and past (t-1), current (t), and future (t-1) spend of the sampled customers and company revenue was only carried out using longitudinal data for a near monopoly company in the agricultural sector. Therefore, the assumption that the findings are consistent across companies and sectors cannot be made.

Additionally, the aggregate comparisons in this research included limited data points as NPSs were only collected between 2011 and 2015. This resulted in only a maximum of 5 data points being included in the longitudinal assessment of the three measures, which is likely to have impacted on the statistical significance of the relationships observed.

The comparison of the NPS with average spend of the sampled customers is also relatively rare as traditionally it is compared against company revenue. The collection of the amount spent by the customers in the sample allowed aggregated and individual analysis to be carried out. At an aggregate level, the findings that suggest "promoters" increase their spending less than "detractors" because they have likely reached saturation point with the company's products is confined to the agricultural company in the study. Similarly, the individual level findings are confined to the agricultural company as limited studies have analysed individual NPSs.

Overall, this research has uncovered some useful findings when comparing the NPS, likelihood mean and Polarization Index. Specifically, this research suggests that loyal customers may not be the most valuable customers in terms of growth in their individual spending, however they do contribute to company revenue growth, presumably by attracting new customers through WOM. Future research is needed to help support the initial findings of this study by assessing the performance of the three loyalty measures to predict company revenue using longitudinal or cross-sectional data in different sectors that are not dominated by a near monopoly.

Future research should ensure NPSs are collected and correlated against financial indicators over a longer time period with more data points to improve the accuracy of the results. Additionally, to confirm the findings of this research further study in different sectors can be used to assess the ability of NPS to predict the amount spent by the sampled customers at an individual level.

10.0 REFERENCES

Anderson, E. W. (1998). Customer satisfaction and word of mouth. *Journal of service research*, *1*(1), 5-17.

Anderson, E. W., Fornell, C., & Lehmann, D. R. (1994). Customer satisfaction, market share, and profitability: findings from Sweden. *The Journal of Marketing*, 58 (3), 53-66.

Anderson, E. W., & Sullivan, M. W. (1993). The antecedents and consequences of customer satisfaction for firms. *Marketing science*, *12*(2), 125-143.

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of marketing Research*, 4(3), 291-295.

Barwise, T. P. (1986). Repeat-viewing of prime-time TV series. *Journal of Advertising Research*, *26*(4), 9-14.

Barwise, T. P., & Ehrenberg, A. S. (1987). The liking and viewing of regular TV series. *Journal of Consumer Research*, 14(1), 63-70.

Bennett, D., & Graham, C. (2010). Is loyalty driving growth for the brand in front? A two-purchase analysis of car category dynamics in Thailand. *Journal of Strategic Marketing*, *18*(7), 573-585.

Bernhardt, K. L., Donthu, N., & Kennett, P. A. (2000). A longitudinal analysis of satisfaction and profitability. *Journal of Business Research*, *47*(2), 161-171.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing science*, *17*(1), 45-65.

Bowen, J. T., & Chen, S. L. (2001). The relationship between customer loyalty and customer satisfaction. *International journal of contemporary hospitality management*, *13*(5), 213-217.

Brennan, M., & Esslemont, D. (1994). The accuracy of the Juster scale for predicting purchase rates of branded, fast-moving consumer goods. *Marketing Bulletin*, *5*(5), 47-52.

Brennan, M., Esslemont, D., & Clarence, U. (1995). Using the Juster Scale to Estimate the Demand-Price Relationship. *Asia-Australia Marketing Journal*, *3*(1), 27-37.

Brennan, M., Hini, D., & Esslemont, D. (1994). Obtaining purchase probability data via telephone surveys: A preliminary test of two techniques. *Marketing Bulletin*, *5*, 64-70.

Byrnes, J. (1964, September). An experiment in the measure of consumer intentions to purchase. In *Proceedings of the Business and Economic Statistics Section* (pp. 658-696).

Casini, L., Rungie, C., & Corsi, A. M. (2009). How loyal are Italian consumers to wine attributes?. *Journal of Wine Research*, *20*(2), 125-142.

Charlett, D., Garland, R., & Marr, N. (1995). How damaging is negative word of mouth. *Marketing Bulletin*, *6*(1), 42-50.

Chickamenahali, S. (1999). *Suma's Great Juster Project: Based on marginal method of moments*. Unpublished methodology sheet: South Australia.

Clawson, C. J. (1971). How useful are 90-day purchase probabilities?. *The Journal of Marketing*, 35(4), 43-47.

Corsi, A. M., Rungie, C., & Casini, L. (2011). Is the polarization index a valid measure of loyalty for evaluating changes over time?. *Journal of Product & Brand Management*, *20*(2), 111-120.

Creamer, M. (2006). Do You Know Your Score: Your Business Rides on How Consumers Answer the One Crucial Question. *Advertising Age*, *77*(1), 24.

Cunningham, R. M. (1956). Brand loyalty-what, where, how much. *Harvard Business Review*, *34*(1), 116-128.

Dawkins, P., & Reichheld, F. (1990). Customer retention as a competitive weapon. *Directors and Boards*, *14*(4), 42-47.

Day, D. (1987). *An examination of the accuracy of two versions of the Juster Scale for predicting consumer purchase behaviour using self-completion questionnaires* (Doctoral dissertation, Massey University).

Day, D., Gan, B., Gendall, P., & Esslemont, D. (1991). Predicting purchase behaviour. *Marketing Bulletin*, *2*(3), 18-30.

Dick, A. S., & Basu, K. (1994). Customer loyalty: toward an integrated conceptual framework. *Journal of the academy of marketing science*, *22*(2), 99-113.

Donthu, N. (1994). Double jeopardy in television program choice. *Journal of the academy of marketing science*, *22*(2), 180-185.

East, R., Gendall, P., Hammond, K., & Lomax, W. (2005). Consumer loyalty: singular, additive or interactive?. *Australasian Marketing Journal (AMJ)*, *13*(2), 10-26.

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing*, *25*(3), 215-224.

East, R., Hammond, K., & Gendall, P. (2006). Fact and fallacy in retention marketing. *Journal of Marketing Management*, *22*(1-2), 5-23.

East, R., Hammond, K., & Wright, M. (2007). The relative incidence of positive and negative word of mouth: A multi-category study. *International Journal of Research in Marketing*, *24*(2), 175-184.

East, R., Romaniuk, J., & Lomax, W. (2011). A critique and an alternative metric. *International Journal of Market Research*, *53*(3), 327-346.

Ehrenberg, A. S., Goodhardt, G. J., & Barwise, T. P. (1990). Double jeopardy revisited. *The Journal of Marketing*, 54(3), 82-91.

Ehrenberg, A., & Goodhardt, G. (2002). Double jeopardy revisited, again. *Marketing Research*, *14*(1), 40-42.

Ehrenberg, A. S., Uncles, M. D., & Goodhardt, G. J. (2004). Understanding brand performance measures: using Dirichlet benchmarks. *Journal of Business Research*, *57*(12), 1307-1325.

Engel, J. F., Blackwell, R. D., & Kegerreis, R. J. (1969). How information is used to adopt an innovation. *Journal of Advertising Research*, *9*(4), 3-8.

Fader, P. S., & Schmittlein, D. C. (1993). Excess behavioral loyalty for high-share brands: Deviations from the Dirichlet model for repeat purchasing. *Journal of Marketing research*, 30(4), 478-493.

Ferber, R., & Piskie, R. A. (1965). Subjective probabilities and buying intentions. *The Review of Economics and Statistics*, 47(3), 322-325.

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *the Journal of Marketing*, 56(1), 6-21.

Furse, D. H., Punj, G. N., & Stewart, D. W. (1984). A typology of individual search strategies among purchasers of new automobiles. *Journal of consumer research*, 10(4), 417-431.

Fornell, C., & Wernerfelt, B. (1987). Defensive marketing strategy by customer complaint management: a theoretical analysis. *Journal of Marketing research*, 24(4), 337-346.

Gan, B. C., Esslemont, D. H. B., & Gendall, P. J. (1986). *A Test on the Accuracy of the Juster Scale as a Predictor of Purchase Behaviour*. Market Research Centre, Massey University.

Gibson, A., Miller, M., Smith, P., Bell, A., Crothers, C. (2013). *The Internet in New Zealand 2013*. Auckland, New Zealand: Institute of Culture, Discourse & Communication, AUT University.

Goodhardt, G. J., Ehrenberg, A. S., & Chatfield, C. (1984). The Dirichlet: A comprehensive model of buying behaviour. *Journal of the Royal Statistical Society. Series A (General)*, 147(5), 621-655.

Goritz, A. (2004). The impact of material incentives on response quantity, response quality, sample composition, survey outcome and cost in online access panels. *International Journal of Market Research*, *46*, 327-346.

Green, P. E., & Rao, V. R. (1970). Rating scales and information recovery. How many scales and response categories to use?. *The Journal of Marketing*, 34(3), 33-39.

Grisaffe, D. B. (2007). Questions about the ultimate question: conceptual considerations in evaluating Reichheld's net promoter score (NPS). *Journal of Consumer Satisfaction Dissatisfaction and Complaining Behavior*, *20*, 36-53.

Gronholdt, L., Martensen, A., & Kristensen, K. (2000). The relationship between customer satisfaction and loyalty: cross-industry differences. *Total quality management*, *11*(4-6), 509-514.

Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International journal of service industry management*, *7*(4), 27-42.

Hamilton-Gibbs, D., Esslemont, D., & McGuinness, D. (1992). Predicting the demand for frequently purchased items. *Marketing Bulletin*, *3*, 18-23.

Heskett, J. L., Jones, T. O., Loveman, G. W., Sasser, W. E., & Schlesinger, L. A. (1994). Putting the service-profit chain to work. *Harvard business review*, *72*(2), 164-174.

Ittner, C. D., & Larcker, D. F. (1998). Are nonfinancial measures leading indicators of financial performance? An analysis of customer satisfaction. *Journal of accounting research*, 36, 1-35.

Ittner, C. D., & Larcker, D. F. (2003). Coming up short on nonfinancial performance measurement. *Harvard business review*, *81*(11), 88-95.

Ittner, C. D., Larcker, D. F., & Rajan, M. V. (1997). The choice of performance measures in annual bonus contracts. *Accounting Review*, 72(2), 231-255.

Jacoby, J., & Kyner, D. B. (1973). Brand loyalty vs. repeat purchasing behavior. *Journal of Marketing research*, 10(1), 1-9.

Jarvis, W., & Goodman, S. (2005). Effective marketing of small brands: niche positions, attribute loyalty and direct marketing. *Journal of Product & Brand Management*, *14*(5), 292-299.

Jarvis, W., Rungie, C., & Lockshin, L. (2003). *Analysing wine behavioural loyalty* (Doctoral dissertation, University of South Australia).

Jarvis, W., Rungie, C., & Lockshin, L. (2007). Revealed preference analysis of red wine attributes using polarisation. *International Journal of Wine Business Research*, *19*(2), 127-138.

Juster, F. T. (1966). Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association*, *61*(315), 658-696.

Kahn, B. E., Kalwani, M. U., & Morrison, D. G. (1986). Measuring variety-seeking and reinforcement behaviors using panel data. *Journal of Marketing Research*, 23(2), 89-100.

Kahn, B. E., Kalwani, M. U., & Morrison, D. G. (1988). Niching versus change-of-pace brands: Using purchase frequencies and penetration rates to infer brand positionings. *Journal of Marketing Research*, 25(4), 384-390.

Kalwani, M. U., & Morrison, D. G. (1980). Sample size requirements for zero-order models. *Journal of Marketing Research*, 17(2), 221-227.

Keiningham, T. L., Cooil, B., Andreassen, T. W., & Aksoy, L. (2007a). A longitudinal examination of net promoter and firm revenue growth. *Journal of Marketing*, *71*(3), 39-51.

Keiningham, T. L., Cooil, B., Aksoy, L., Andreassen, T. W., & Weiner, J. (2007b). The value of different customer satisfaction and loyalty metrics in predicting customer retention, recommendation, and share-of-wallet. *Managing Service Quality: An International Journal*, *17*(4), 361-384.

Keiningham, T. L., Aksoy, L., Cooil, B., Andreassen, T. W., & Williams, L. (2008a). A holistic examination of Net Promoter. *Journal of Database Marketing & Customer Strategy Management*, *15*(2), 79-90.

Keiningham, T. L., Aksoy, L., Cooil, B., & Andreassen, T. W. (2008b). Linking customer loyalty to growth. *MIT Sloan Management Review*, *49*(4), 50-57.

Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *The Journal of Marketing*, 59(2), 71-82.

Kristensen, K., & Westlund, A. H. (2004). The one number you need to grow:(comment on Reichheld). *Harvard Business Review*, *82*(4), 138-139.

Krystallis, A. (2013). Uncovering attribute-based determinants of loyalty in cigarette brands. *Journal of Product & Brand Management*, *22*(2), 104-117.

Krystallis, A., & Chrysochou, P. (2010). An exploration of loyalty determinants in Greek wine varieties. *EuroMed Journal of Business*, *5*(2), 124-137.

Krystallis, A., & Chrysochou, P. (2011). Health claims as communication tools that enhance brand loyalty: The case of low-fat claims within the dairy food category. *Journal of Marketing Communications*, *17*(3), 213-228.

Kuehn, A. A. (1962). Consumer brand choice as a learning process. *Journal of Advertising Research*, *2*(4), 10-17.

Kumar, V., Petersen, J. A., & Leone, R. P. (2007). How valuable is word of mouth? *Harvard business review*, *85*(10), 139.

LaBarbera, P. A., & Mazursky, D. (1983). A longitudinal assessment of consumer satisfaction/dissatisfaction: the dynamic aspect of the cognitive process. *Journal of marketing research*, 20(4), 393-404.

Lee, M. S., Elango, B., & Schnaars, S. P. (1997). The accuracy of the Conference Board's buying plans index: A comparison of judgmental vs. extrapolation forecasting methods. *International Journal of Forecasting*, *13*(1), 127-135.

Lees, G., & Wright, M. (2013). Does the duplication of viewing law apply to radio listening?. *European journal of Marketing*, *47*(3/4), 674-685.

Lipstein, B. (1959, September). The dynamics of brand loyalty and brand switching. In *Proceedings of the fifth annual conference of the advertising research foundation* (pp. 101-108). New York, NY: Advertising Research Foundation.

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, *4*(2), 73-79.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business horizons*, *52*(4), 357-365.

Marsden, P., Samson, A., & Upton, N. (2005). Advocacy drives growth. *Brand strategy*, (198), 45-47.

Massey, W., Montgomery, D. B., & Morrison, D. G. (1970). Stochastic Models of Buying Behavior.

McDowell, W. S., & Dick, S. J. (2005). Revealing a double jeopardy effect in radio station audience behavior. *Journal of Media Economics*, *18*(4), 271-284.

Mecredy, P., Feetham, P., & Wright, M. (2015). *Can we get more out of Net-Promoter data?* Paper presented at the ANZMAC Conference, Sydney, Australia.

Moore, W. L., & Lehmann, D. R. (1980). Individual differences in search behavior for a nondurable. *Journal of consumer research*, 7(3), 296-307.

Morgan, N. A., & Rego, L. L. (2006). The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science*, *25*(5), 426-439.

Pickering, J. F., & Isherwood, B. C. (1974). Purchase probabilities and consumer durable buying behaviour. *Journal of the Market Research Society*, *16*(3), 203-226.

Pingitore, G., Morgan, N. A., Rego, L. L., Gigliotti, A., & Meyers, J. (2007). The Single-Question Trap. *Marketing Research*, *19*(2), 9-13.

Preston, C. C., & Colman, A. M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica, 104,* 1-15.

Rego, L. L., & Morgan, N. A. (2004). The one number you need to grow. *Harvard Business Review*, *82*(4), 134-136.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, *81*(12), 46-55.

Reichheld, F. (2006a). The microeconomics of customer relationships. *MIT Sloan Management Review*, *47*(2), 73-78.

Reichheld, F. F. (2006b). *The ultimate question: Driving good profits and true growth*. Harvard Business Press.

Reichheld, F. P., & Sasser, W. E. (1990). Zero defections: Quality comes to services. *Harvard business review*, *68*(5), 105-111.

Reinartz, W. J., & Kumar, V. (2000). On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. *Journal of marketing*, *64*(4), 17-35.

Reinartz, W., & Kumar, V. (2002). The mismanagement of customer loyalty. *Harvard business review*, *80*(7), 86-95.

Riebe, E., Wright, M., Stern, P., & Sharp, B. (2014). How to grow a brand: Retain or acquire customers?. *Journal of Business Research*, *67*(5), 990-997.

Roberts, J. (1989). A grounded model of consideration set size and composition. *Advances in Consumer Research*, *16*(1), 749-757.

Romaniuk, J., Nguyen, C., & East, R. (2011). The accuracy of self-reported probabilities of giving recommendations. *International Journal of Market Research*, *53*(4), 507-522.

Rungie, C., Brown, B., Laurent, G., & Rundrapatna, S. (2005). *A standard error estimator for the polarization index: assessing the measurement error in one approach to the analysis of loyalty* (Doctoral dissertation, Massey University, Department of Marketing).

Rust, R. T., & Zahorik, A. J. (1993). Customer satisfaction, customer retention, and market share. *Journal of retailing*, *69*(2), 193-215.

Sabavala, D. J., & Morrison, D. G. (1977). A model of TV show loyalty. *Journal of Advertising Research*, 17(6), 35-43.

Sharp, B. (2008). Net promoter score fails the test. *Marketing research*, *20*(4), 28-30.

Sharp, B. (2010). *How brands grow: what marketers don't know*. Melbourne, Australia: Oxford University Press.

Sharp, B., Wright, M., & Goodhardt, G. (2002). Purchase loyalty is polarised into either repertoire or subscription patterns. *Australasian Marketing Journal (AMJ)*, *10*(3), 7-20.

Theil, H., & Kosobud, R. F. (1968). How informative are consumer buying intentions surveys?. *The Review of Economics and Statistics*, 50(1), 50-59.

Uncles, M. D., East, R., & Lomax, W. (2010). Market share is correlated with word-of-mouth volume. *Australasian Marketing Journal (AMJ)*, *18*(3), 145-150.

Uncles, M., Ehrenberg, A., & Hammond, K. (1995). Patterns of buyer behavior: Regularities, models, and extensions. *Marketing Science*, *14*(3_supplement), G71-G78.

Uncles, M., & Lee, D. (2006). Brand purchasing by older consumers: an investigation using the Juster scale and the Dirichlet model. *Marketing Letters*, *17*(1), 17-29.

Van Doorn, J., Leeflang, P. S., & Tijs, M. (2013). Satisfaction as a predictor of future performance: A replication. *International journal of research in marketing*, *30*(3), 314-318.

Westbrook, R. A. (1987). Product/consumption-based affective responses and postpurchase processes. *Journal of marketing research*, 24(3), 258-270.

Wright, M., & MacRae, M. (2007). Bias and variability in purchase intention scales. *Journal of the Academy of Marketing Science*, *35*(4), 617-624.

Wright, M., Sharp, A., & Sharp, B. (1998). Are Australasian brands different?. *Journal of Product & Brand Management*, *7*(6), 465-480.

Wright, M., Sharp, A., & Sharp, B. (2002). Market statistics for the Dirichlet model: using the Juster scale to replace panel data. *International Journal of Research in Marketing*, *19*(1), 81-90.

Yeung, M. C., Ging, L. C., & Ennew, C. T. (2002). Customer satisfaction and profitability: A reappraisal of the nature of the relationship. *Journal of Targeting, Measurement and Analysis for Marketing*, *11*(1), 24-33.

Young, M. R., DeSarbo, W. S., & Morwitz, V. G. (1998). The stochastic modeling of purchase intentions and behavior. *Management Science*, *44*(2), 188-202.

Zeithaml, V. A. (2000). Service quality, profitability, and the economic worth of customers: what we know and what we need to learn. *Journal of the academy of marketing science*, *28*(1), 67-85.

11.0 APPENDICES

Appendix A: Sample demographics (real estate industry)

Demographic Variable		Percent
	Male 18 - 34 years	5
	Male 35 – 54 years	13
	Male 55 years or older	16
	Female 18 - 34 years	22
Gender and age	Female 35 – 54 years	27
	Female 55 years or older	16
	Northland	8
	Auckland	47
	Central North Island	19
Region	Lower North Island	13
	South Island	13
	Less than $10,000	4
	$10,001 - $20,000	6
	$20,001 - $40,000	20

	$40,001 - $60,000	18
	$60,001 - $80,000	15
	$80,001 - $100,000	15
Household Income before tax (annual)	$100,001 - $120,000	9
	$120,001 - $140,000	6
	More than $140,000	7

	New Zealander of European descent	75
	New Zealand Māori	6
	Pacific peoples	3
	Chinese	4
	Indian	4
	Other Asian	4
Ethnicity	Born overseas of European decent	8
	Other	3

Appendix B: Sample living situations (real estate industry)

Living Variable		Percent
	Couple with child(ren) living at home	33
	Couple with child(ren) no longer living at home	19
	Couple without children	16
	Single person without children	18
Family situation	Single parent with child(ren) living at home	7
	Single parent with child(ren) no longer living at home	8
	Boarding, housesitting, or living with parents	7
	Renting in a shared house or apartment	8
	Renting a home just for you, or just for your family	22
	Living in a home you own, with a mortgage	36
Housing situation	Living in a home you own, without a mortgage	26
	Other	1

Appendix C: Sample real estate experiences (real estate industry)

Experience Variable		Percent
Housing experience (past 5 years)	Personally signed a real estate rental agreement	27
	Personally signed a real estate purchase agreement as a seller	18
	Personally signed a real estate purchase agreement as a buyer	28
	Personally signed three or more real estate agreements for either sale or purchase	3
	None of the above	49
Investment property ownership	Yes	14
	No	86

Appendix D: ANZMAC Conference Paper

Can we get more out of Net-Promoter data?

Philip Mecredy, Massey University, pmecredy63@hotmail.com*
Pamela Feetham, Massey University, p.m.feetham@massey.ac.nz
Prof Malcolm Wright, Massey University, malcolm.wright@marketingscience.info

Abstract

Net-Promoter Score (NPS), a loyalty measure, is used extensively in commercial market research due to its simplicity of use and ease of understanding, despite criticism of the metric. Given the widespread use of NPS commercially, it is important to understand whether applying alternative loyalty measures has any advantages over Net-Promoter. This paper aims to demonstrate whether a likelihood mean and Polarization Index, *φ*, provide different results to Net-Promoter. These three measures were applied to data collected from an on-line survey of 1,818 participants who evaluated brands in a service industry. The findings show that all three measures provided similar variations in loyalty across brands and regions. The likelihood mean and NPS are strongly correlated, indicating that no one measure is more superior to the other at measuring loyalty within a service industry in New Zealand. However, the Polarization Index appears to assess loyalty differently to the likelihood mean and NPS.

Keywords: Net-Promoter, Polarization Index, brand loyalty metrics

Track: Marketing Research Methods

1.0 Introduction

Net-Promoter was first introduced by Reichheld (2003) as an alternative loyalty metric to predict brand growth. Reichheld (2003) proposes that by asking a single "would recommend" question, loyalty is determined and growth predicted. To determine the Net-Promoter Score (NPS), customers are asked, "On a scale of 0-10, how likely is it that you would recommend [company X] to a friend or colleague?" Scores of 10 indicate "extremely likely" to recommend, 0 indicates "not at all likely" and 5 indicates "neutral". Those scoring 9-10 are classified as "promoters", 7-8 as "passively satisfied" and 0-6 as "detractors". The NPS is calculated by subtracting the percentage of "detractors" from the percentage of "promoters".

Within the majority of the industries examined, Reichheld (2003) found that there was a strong correlation between NPS and company growth. Reichheld also claimed that a twelve-point increase in NPS corresponds to a doubling of a company's growth rate (Reichheld, 2006). Other research has found that a seven-point increase in NPS produces a one-percent increase in brand growth (Marsden, Samson and Upton, 2005). However, the findings of Reichheld (2003) and Marsden et al., (2005) are considered flawed because their NPS's were correlated with past growth rates (Keiningham, Cooil, Andreassen and Askoy 2007).

To address the concerns about Net-Promoter, Keiningham et al., (2007) examined the relationship between NPS and company growth rates in a cross-industry longitudinal study. They replicated Reichheld's study, but instead of using past growth rates they correlated NPS with company growth rates from identical time periods. They found no support for Reichheld's claim that Net-Promoter is the only question required to measure growth in customer surveys. By comparing correlations of both NPS and the American Customer Satisfaction Index (ACSI) with company growth rates, they determined that Net-Promoter performance is not superior to the ACSI.

Other researchers have identified further concerns with Reichheld's methodology as he broke the 11-point Net-Promoter scale into three categories and excluded the "passively satisfied" category from his calculations (Grisaffe, 2007). Grisaffe raises the concern that Reichheld's clustering of the scale results in different scenarios requiring diverse managerial actions being seen as similar (2007). For example, Company X may have 20 "promoters", 0 "passively satisfied" and 20 "detractors". Alternatively, Company Y may have 0 "promoters", 40 "passively satisfied" and 0 "detractors". Despite significantly different scenarios requiring different managerial actions, both companies have a NPS of zero. In this case, relying on NPS is likely to mislead marketing decisions as it is unclear which customer group requires more focus.

Further work by East, Hammond and Lomax (2008) revealed that Net-Promoter fails to directly measure negative word-of-mouth (NWOM). Instead, in Net-Promoter Score, NWOM is inferred from respondents who indicate low scores on the willingness to engage in positive word-of-mouth (PWOM), classified by "detractors" (East et al., 2008). Later work discovered that NPS is poor at capturing NWOM because “detractors” were shown to engage in both PWOM and NWOM (East, Romaniuk and Lomax, 2011).

Despite these criticism and weaknesses, Net-Promoter has been extensively adopted in commercial market research because of its simplicity of use and ease of understanding. NPS is being used in multi-national corporations including eBay, American Express and Apple (Reichheld, 2006). NPS is even reported to shareholders and used to determine pay in employment contracts (Creamer, 2006).

The lack of academic support, methodological concerns and extensive commercial adoption of NPS present an increasing need to establish whether calculating the NPS from “promoters” minus “detractors” is superior to the application of additional measures to the Net-Promoter question. This paper presents the first stage of research intended to test the performance of NPS. It applies two other loyalty measures; a likelihood mean and Polarization Index, φ, to the Net-Promoter question. The three measures are compared across a variety of service brands commonly known in New Zealand; however, the data used in this research did not allow correlations with company growth rates.

2.0 Methodology

Data for this research is drawn from a survey of service brands in one category in New Zealand during 2014. Participants were recruited by a commercial panel provider across five regions in New Zealand; Northland, Auckland, Central North Island, Lower North Island and South Island. The final sample contained 1,818 respondents of mixed demographics, close to those of New Zealand’s census data.

In a randomized order, participants viewed the logos of six service brands tested and were asked, “On a scale of 0-10, how likely are you to recommend ‘brand X’ to a friend or colleague?” A scale from 0-10 was provided with numerical descriptors at each scale point and verbal descriptors at 0 “not at all likely” and 10 “extremely likely”. The six service brands tested varied in each region depending on brand presence. In total, eleven brands were evaluated across New Zealand with mixed sample sizes.

Using the eleven brands assessed, observed frequencies were generated for each Net-Promoter scale point. These frequencies allowed for a NPS, likelihood mean and Polarization Index to be calculated for each brand. The calculation of the likelihood mean is similar to that employed in assessing the variability in purchase intention scales (Wright and MacRae, 2007). The formula used for the calculation of the Polarization Index was consistent with that used by Corsi, Rungie and Casini (2011). It has been calculated as follows:

$$\varphi = \frac{1}{1+S}$$

Once each measure was calculated, correlation coefficients were produced to assess whether the measures had any variations in assessing loyalty across brands.

Potential variations in each measure were analysed across regions using the four brands that were surveyed in at least four of the five regions.

Biases common in survey data were minimized by presenting the brands in a randomized order. Internet coverage bias is unlikely as in New Zealand over 90% of the population have internet access (Gibson, Miller, Smith, Bell and Crothers, 2013). Recruitment bias in the on-line panel is unlikely as the panel size is substantial (n = 75,000).

3.0 Findings

Table 1 shows the NPS, likelihood mean and Polarization Index for each brand. NPS varied from -90% to -50% indicating that "detractors" significantly outweighed the "promoters" amongst all brands in New Zealand. The likelihood mean ranged from 0.40 to 0.59 and the Polarization Index ranged from 0.16 to 0.28, indicating that variation across brands was small. Additionally, the likelihood mean appears to show no significant discrimination beyond that offered by NPS, as brands with a higher NPS tended to also have a higher likelihood mean. The Polarization Index indicated that Brand B (φ=0.21), E (φ=0.28) and G (φ=0.24) have the greatest loyalty despite having a weaker NPS and likelihood mean compared to other brands. Each of these brands has a low market share in just one of the five regions analysed, and thus has a very small sample size. Larger sample sizes are required for accurate calculations of the Polarization Index, especially for brands with low market share (Kalwani and Morrison, 1980).

Table 1: Comparison of measures across brands

	Sample Size, n	NPS	Likelihood mean	φ
Brand A	849	-50%	0.59	0.18
Brand B	233	-62%	0.53	0.21
Brand C	1818	-69%	0.53	0.17
Brand D	1818	-71%	0.52	0.16
Brand E	244	-74%	0.43	0.28
Brand F	736	-75%	0.50	0.17
Brand G	146	-75%	0.43	0.24
Brand H	1428	-79%	0.47	0.18
Brand I	1818	-82%	0.46	0.16
Brand J	1672	-88%	0.40	0.19
Brand K	146	-90%	0.40	0.17

Table 2 shows the correlations between the three measures across brands. The likelihood mean and NPS have a strong correlation (0.904), which demonstrates that brands with a high NPS also had a high likelihood mean. This suggests there is no significant difference between which of the two measures is used to evaluate loyalty in this service industry within New Zealand. However, the Polarization Index has a low correlation with NPS (0.100) and a negative correlation with the likelihood mean (-0.304). This indicates that the Polarization Index provides an alternative evaluation of loyalty in contrast to the two other measures. A potential reason for this could be the large mid-point bias identified in the analysis. For all eleven brands, between 23% and 32% of respondents indicated their likelihood to recommend the brand to a friend as a 5, indicated by "neutral". The Polarization index is more likely to be influenced by the size of the mid-point bias, as the uncertainty associated with 50/50 probabilities has a substantial effect on the value of the index. While the mid-point bias may be affecting these results, this preliminary examination requires further testing across countries, industries and brands before such a conclusion can be drawn.

Table 2: Correlation coefficients across brands

Correlation	***NPS***	***Likelihood Mean***	***φ***
NPS	1		
Likelihood mean	0.904	1	
φ	0.100	-0.304	1

The three measures were also analysed to assess variation across regions. Four prominent brands, surveyed in at least four out of the five regions, were selected for the analysis. Figure 1-3 displays the variation across regions for the NPS, likelihood mean and Polarization Index respectively. As shown below, the variation in all three measures is relatively minimal across regions for each brand. However, visual inspection suggests that the Polarization Index shows less variation and so may be a more stable measure. However, the lack of variation could be due to other reasons. The next stage of this research will test other data sets to determine whether the Polarization Index is a more stable measure.

Figure 1: NPS across regions

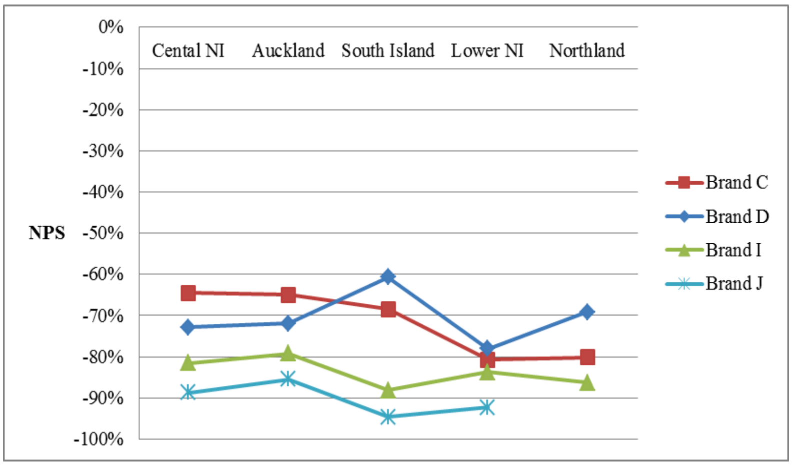

Figure 2: Likelihood mean across regions

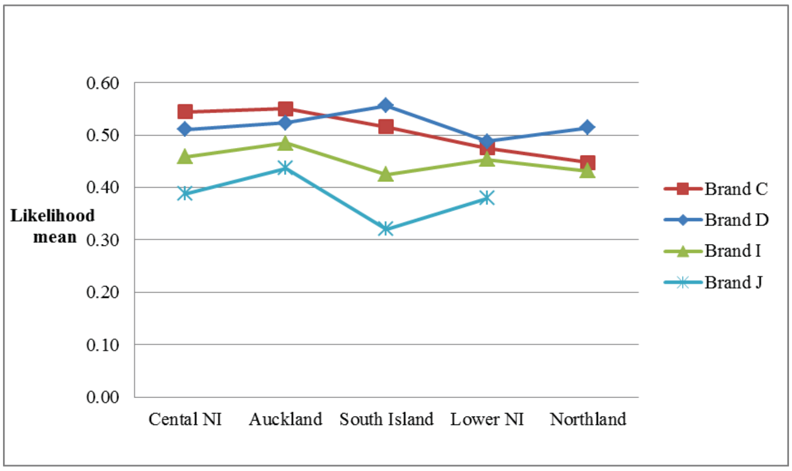

Figure 3: Polarization Index across regions

4.0 Conclusion

This research determined that when comparing brands across a service industry in New Zealand, no measure applied to the Net-Promoter question is superior at identifying variability in loyalty across brands. It also found that there is little variation in each measure across regions, with visual inspections of the Polarization Index showing slightly more stability. Additionally, the NPS and likelihood mean are highly correlated. This indicates that when conducting a comparative analysis of service brands, there is little difference in reported loyalty levels between the two measures used to evaluate the Net-Promoter question. However, the Polarization Index appears to assess loyalty differently to the NPS and the likelihood mean. As growth rates were not collected, this research was unable to correlate the three measures with growth to assess whether the contrasting Polarization Index is superior to NPS and a likelihood mean.

Further research will consider assessing the measures in alternative industries and countries. The additional research will investigate which measure is more accurate at predicting brand growth. Calculating correlations between the three measures and drivers of growth will establish which measure is superior when applied to the Net-Promoter question.

References

Corsi, A. M., Rungie, C., & Casini, L. (2011). Is the polarization index a valid measure of loyalty for evaluating changes over time?. *Journal of Product & Brand Management*, *20*(2), 111-120.

Creamer, M. (2006). Do you know your score: your business rides on how consumers answer the one crucial question. *Advertising Age*, *77*(1), 24.

East, R., Hammond, K., & Lomax, W. (2008). Measuring the impact of positive and negative word of mouth on brand purchase probability. *International Journal of Research in Marketing*, *25*(3), 215-224.

East, R., Romaniuk, J., & Lomax, W. (2011). The NPS and the ACSI: a critique and an alternative metric. *International Journal of Market Research*, *53*(3), 327-346.

Gibson, A., Miller, M., Smith, P., Bell, A., Crothers, C. (2013). The Internet in New Zealand 2013. Auckland, New Zealand: Institute of Culture, Discourse & Communication, AUT University. Retreived from: http://www.aut.ac.nz/__data/assets/pdf_file/0007/424816/wipnz2013final.pdf

Grisaffe, D. B. (2007). Questions about the ultimate question: conceptual considerations in evaluating Reichheld's net promoter score (NPS). *Journal of Consumer Satisfaction Dissatisfaction and Complaining Behavior, 20*, 36.

Kalwani, M. U., & Morrison, D. G. (1980). Sample size requirements for zero-order models. *Journal of Marketing Research*, 221-227.

Keiningham, T. L., Cooil, B., Andreassen, T. W., & Aksoy, L. (2007). A longitudinal examination of net promoter and firm revenue growth. *Journal of Marketing*, *71*(3), 39-51.

Reichheld, F. F. (2003). The one number you need to grow. *Harvard business review*, *81*(12), 46-55.

Reichheld, F. F. (2006). *The ultimate question: Driving good profits and true growth*. Harvard Business Press.

Wright, M., & MacRae, M. (2007). Bias and variability in purchase intention scales. *Journal of the Academy of Marketing Science*, *35*(4), 617-624.