Ehrenberg-Bass Institute **Working Paper**:

Where is the brand growth potential? An examination of buyer groups.

*This working paper, dated 24 May 2023, is forthcoming in the **Marketing Letters.***

Authors:

Assoc. Professor Giang Trinh - Ehrenberg-Bass Institute
Professor John Dawes - Ehrenberg-Bass Institute
Professor Byron Sharp - Ehrenberg-Bass Institute

University of South Australia | Ehrenberg-Bass Institute for Marketing Science

Where is the brand growth potential? An examination of buyer groups.

Assoc. Prof. Giang Tue Trinh

Ehrenberg-Bass Institute, UniSA Business

University of South Australia

Level 4, Yungondi Building

North Terrace

Adelaide, SA 5000

Australia

Telephone: +61 8 8302 0600

Email: Giang.Trinh@marketingscience.info

Prof. John Dawes

Ehrenberg-Bass Institute, UniSA Business

University of South Australia

Level 4, Yungondi Building

North Terrace

Adelaide, SA 5000

Australia

Telephone: +61 8 8302 0592

Email: John.Dawes@marketingscience.info

Prof. Byron Sharp

Ehrenberg-Bass Institute, UniSA Business

University of South Australia

Level 4, Yungondi Building

North Terrace

Adelaide, SA 5000

Australia

Telephone: +61 8 8302 0715

Email: Byron.Sharp@marketingscience.info

Abstract

Practitioners and academics have long discussed strategies for brand sales growth. A recent example is an industry debate in which different brand growth strategies were argued: https://www.mmaglobal.com/thegreatdebate (MMA Global & Neustarr, 2021). A central question in this arena is whether a brand should focus on its heavy, light, or non-buyers in its efforts to grow its sales. This study contributes to our knowledge about how sales growth can occur by investigating the potential contribution these three buyer groups can make to any sales gain. Using both a simulation study and an empirical study of purchases of approximately 12,400 households in the UK, across different brands and categories, we show that almost any brand's headroom growth potential lies mostly in light or non-buyers of that brand. Even for large brands with high penetration the growth potential of light brand buyers eclipses heavy brand buyers.

Keywords

Marketing strategy; Brand growth; Pareto share; Dirichlet model; Heavy buyers

Acknowledgments

The authors thank Kantar World Panel for providing data to enable this research.

1. Introduction

The degree to which marketing strategy should focus on heavy versus light buyers of a brand has generated much debate (e.g. Hallberg, 1995; Hallberg, 1999; Reilly & Rapacz Deb, 2009; Smith & Blair, 2018; Taussig, 2016). At the heart of this debate is the question of how much heavy buyers of a brand (the brand's heaviest 20% or top customers) contribute to its current sales, for example: is the contribution close to 80% (i.e., the Pareto rule). Kim et al. (2017) reported that heavy buyers contribute 73% of sales. Many authors suggest, therefore, that management attention should focus on these heavy buyers (Hallberg, 1999; Taussig, 2016). By contrast, Sharp (2010) reported that heavy buyers contribute around half the sales; therefore, concentrating on this segment should not be the only focus of marketing strategies aimed at brand growth. Similarly, Dawes et al. (2022) showed that extremely light buyers contribute 40% of total sales; as such, they should not be ignored by the brand's management. When considering these past results, it should be noted that estimates of the sales contribution of any buyer group vary depending on the length of time period, brands and product categories, as well as the selection of buyers included in the analysis (Schmittlein et al., 1993). Additionally, none of the previous studies on this topic has investigated the *growth potential* of heavy brand buyers versus light brand buyers, looking beyond the *current* sales importance of these buyer groups to the brand to appraise the potential contribution to future sales.

This research investigates the headroom potential for brand growth across different brand buyer groups by looking at the difference between brand purchases and category purchases of each buyer group. In this respect, headroom for growth incorporates a concept similar to share of category requirements (SCR) or share of wallet (Bhattacharya et al., 1996; Bowman & Narayandas, 2001; Du et al., 2007; Keiningham et al., 2011) which is the brand's share of category purchases among its own buyers. SCR, however, does not directly link to sales potential. This is because a brand's sales potential is a function not only of the proportion of its buyers' category sales that it wins, but also of how many such buyers there are, and how much they buy from the category.

We illustrate this idea with an example. For simplicity, we focus on current buyers of the brand; not non-buyers. We introduce non-buyers into the narrative later. Suppose that light and heavy buyers of a brand contribute 400 units and 600 units, respectively, to the brand's current sales. Furthermore, light buyers contribute 1,000 units to category sales, while heavy buyers contribute 1,200 units to category sales. The growth potential of the two groups is similar. Both light and heavy buyers have 600 units of headroom for the brand's growth; hence, it would make sense for marketers to focus on both groups to grow the brand. However, if light buyers contribute 600 units to the category while heavy buyers contribute 1,600 units, the headroom for the light buyers is only 200 units and the heavy buyers would be 1,000 units. Therefore, arguably the brand's management might be more inclined to focus on heavy brand buyers to grow the brand.

Of course, managers may pursue customer groups for reasons other than just sales potential, such as the cost to serve, or whether they are a good match to the company's capability to serve. This study provides knowledge about where the greatest sales potential resides. This is important information for managers to factor into strategy formulation.

A question that may immediately arise concerns the brand's non-buyers. There will usually be a much bigger number of category buyers who do not currently buy a brand, compared to those who do. Logically, this suggests the greatest sales potential *must* lie among non-buyers, simply because there are so many of them. However, again, the total number of buyers is not the sole criterion for assessing the growth potential of any buyer group; we also need to consider the rate at which that group purchases the category (Hess & Doe, 2013; Uncles et al., 1995). For example, a brand's non-buyers might be very numerous, but they may buy little of the category. It is also likely that while a brand's heavy buyers are a small proportion of its total buyer base (Barnard & Ehrenberg, 1997; Graham & Kennedy, 2022; Morrison & Schmittlein, 1988), they may buy a very large amount of the category. In this case the growth potential could be greater among the brand's current heavy buyers.

2. Background

Research on the sales contribution of heavy versus light buyers began more than half a century ago. Twedt (1964) examined light and heavy category purchasers in two consumer packaged goods (CPG) categories, based on a median split, and found that the top half of users contributed 87% and 86% of category purchases, respectively. Approximately 30 years later, Schmittlein et al. (1993) redefined the classification of heavy buyers according to the Pareto principle (the top 20% of customers) and investigated their sales contribution in four CPG categories. The study found that the heavy users' sales contribution varies across categories and time periods. For example, the heaviest 20% of category buyers accounted for

49% of sales in a month and 65% in a year in the Yoghurt category, but 32% and 50% in the Ketchup category, respectively. These two initial studies only looked at heavy and light buyers of a product category. They did not investigate heavy and light buyers of a brand, which is more relevant for a brand manager.

Adopting the same classification as Schmittlein et al. (1993) for brand buyers, Sharp (2010) reported that the heavy brand buyers in dozens of CPG categories in the US, Australia and South Africa on average contribute 39% in 3 months and 50% in a year to the total brand sales. Kim et al. (2017) used a six-year window, and reported that the heaviest 20% of brand buyers in 22 CPG categories in the US accounted for 73% of brand sales. More recently, McCarthy and Winer (2019) reported that the heaviest 20% of buyers of non-CPG brands in the US contributed 67% over two years to total company sales, while Dawes et al. (2022) reported that the heaviest 20% of brand buyers of 200 brands in 10 CPG categories in the UK contributed 60% of total brand sales over five years.

The differences in sales contribution in these previous studies are due to several known factors. Schmittlein et al. (1993) noted that frequently-bought categories and a longer time window for analysis will produce a higher Pareto ratio. Over time there will be extremely infrequent brand buyers who will occasionally buy the brand, increasing the total size of the brand's buyer base and making the sales proportion of the heaviest 20% appear higher. In summary, previous studies have established that heavy buyers are certainly important to current brand sales, but their contribution to sales is not as extreme as the Pareto 20:80 law (see Sharp et al., 2019).

These previous studies have not investigated the *potential* importance of those who are presently heavy or light buyers for brand *growth*. An increase in total sales can be achieved by winning more sales from a brand's heavy buyers, light buyers, or those who presently do not include the brand in their repertoire[1]. Although heavy buyers of a brand contribute substantially to brand sales, as showed in the previously cited studies, the question arises: are they the group with the greatest potential to contribute to brand growth?

[1] The answer to what constitutes a non-brand buyer is time dependent. Many buyers only buy a CPG brand very occasionally – for example once in five years (Dawes et al., 2022). We define a non-buyer here as a household that has not purchased the brand during the period of analysis.

Our research question can therefore be phrased as where (among which groups) does the greatest sales potential reside – non-buyers, light, or heavy brand buyers?

We assume that category sales are stable year after year. This is appropriate considering past research that indicates most consumer package-goods categories are quite stationary over a period of years (Dekimpe & Hanssens, 1995; Dunn et al., 2021). That said, if the category is growing or declining, we can adjust the headroom potential based on the category growth rate accordingly. We first use a simulation study to show the headroom potential for growth in different conditions, including brand sizes and category types. The use of a simulation allows us to pre-specify these conditions, and to ascertain the predictions about sources of brand growth potential made by a theoretical model with extensive real-world support – the NBD-Dirichlet. We then use an empirical study to validate the results from the simulation.

3. Simulation study

To address our research question, we use a special case of the NBD-Dirichlet model to simulate brand sales and category sales for brand heavy, light and non-buyers based on their frequency of buying the brand (e.g., non-buyers, one-time buyers, two-time buyers, …, n time buyers). The NBD-Dirichlet model is a comprehensive probability model developed by Goodhardt et al. (1984) that fits with a number of well documented scientific laws. The model is typically used to describe and appraise the competition among multiple brands within a category. It captures both product category purchase frequency and brand purchase frequency. It is recognised as one of the most well-established models in marketing (Driesener & Rungie, 2022; East & Hammond, 1996; Ehrenberg et al., 2004; Lam & Mizerski, 2009; Pare & Dawes, 2011; Rungie et al., 2013; Sharp et al., 2012; Stern & Hammond, 2004; Uncles & Lee, 2006; Uncles et al., 1995). The model can predict brand performance measures across a wide range of categories, countries and conditions with a high degree of accuracy (Casteran et al., 2019; Chowdhury et al., 2021; Driesener et al., 2022; Driesener et al., 2017; Riebe et al., 2014; Trinh et al., 2017; Trinh & Lam, 2016; Trinh et al., 2016). As we only look at purchases of one brand versus total purchases of the category, we use a special case of the Dirichlet model for a brand versus all other brands, which is the Beta Binomial Distribution (BBD). Appendix 1 shows details of the model.

Since we seek to investigate brand growth potential, we not only simulate current brand and product purchases, but also simulate future brand purchases and future product purchases of these buyer groups (Schmittlein et al., 1985). We run a simulation for four types of product categories using a mix of penetration (high / low) and purchase frequency (high / low). Penetration is the proportion of households who buy a brand (or category) at all in a time period such as a year. Purchase frequency is the average number of occasions a brand's (or category's) buyers buy it in a time period such as a year (Ehrenberg, 1988). We also use brands of two sizes: small (5% market share) and large (50% market share). We create a simulated sample of 10,000 shoppers. We use the model to simulate the sales potential of non, light and heavy buyers. As per past work we define light buyers as the lightest 80% and heavy buyers as the heaviest 20% (Kim et al., 2017; Martin et al., 2020; Schmittlein et al., 1993). We show an example of simulated data in Appendix 2 and the simulation scenarios in Appendix 3 with details of the model parameters. We next present the results for eight different scenarios.

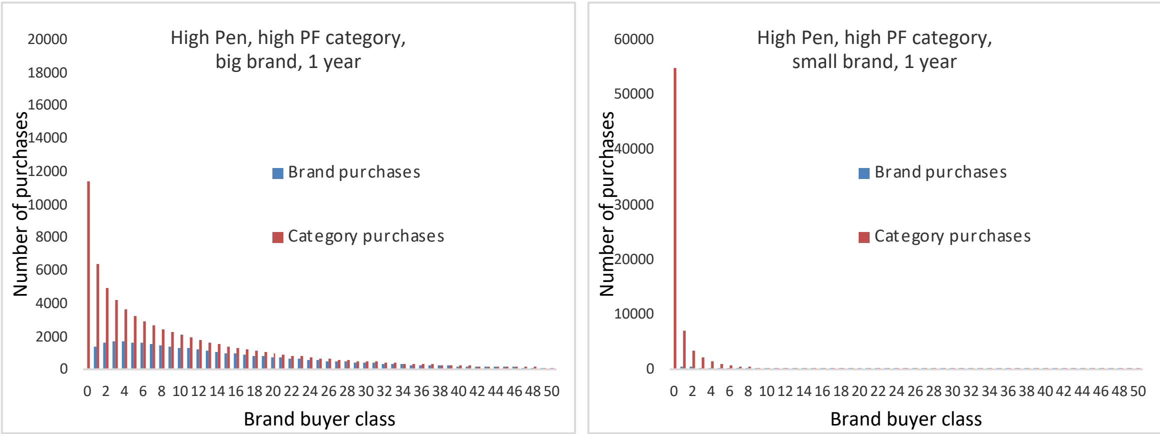
Figure 1 shows brand and category purchases of different brand buyer classes in a high penetration and high purchase frequency category to illustrate the results. For both large and small brands over a period of one year, the difference between total *category* purchases and current *brand* purchases (the difference between the red and blue columns) is mostly due to light buyers or non-brand buyers. In these graphs the y-axis represents the total number of purchases for brand and category (columns D and E in Appendix 2). The x-axis represents buyer frequency groups (column A in Appendix 2). For example, in the first graph we see that the buyer group that makes two brand purchases comprises 1,621 brand purchases and 4,975 category purchases. Therefore, the headroom for growth is 3,354 purchases. If we select a heavier buyer group in this graph, such as those who purchase the brand 10 times per year, they account for 1,347 brand sales and fewer category sales (2,087). In this instance, the headroom for growth is only 740 purchases.

Appendix 2 (columns G, H, I) and Appendix 4 show the future brand purchases, future category purchases and future sales potential of these brand buyer groups. Despite the difference between simulated current purchases and future purchases, simulated current sales potential and future sales potential are the same. This is also proven by equation (A11) in Schmittlein et al. (1985).

Considering the empirical evidence these charts portray, when we compare brands of different sizes, there is more growth potential from non-buyers than light buyers for small brands, while the opposite holds for big (large share) brands.

Figure 1. Brand and category purchases of different brand buyer classes in a high penetration and high purchase frequency category

Part a Part b



To add more context to these results, Table 1 quantifies the importance across different brand buyer groups by calculating the percentage contribution of each buyer group to total growth potential in all eight simulation scenarios. For example, for a large brand, in a one-year period and in a high-penetration, high-frequency category, 31% of its sales potential exists in its current non-buyers, 52% in current light buyers, and 17% in current heavy buyers.

Overall, as seen in the bottom three rows of the table, these additional results show that the greatest sales percentage contribution originates, in the first place, from non-brand buyers (62% on average) and, in the second instance, current light buyers of the brand (28%). The smallest increment to growth comes from current heavy brand buyers (10%). Indeed, for small brands, the growth potential of current heavy buyers is negligible at under 10%.

Table 1 Growth potential of brand non, light and heavy buyers: Simulation results

		Growth Potential = Total category purchases – Total brand purchases				
		Large brand		Small brand		**Overall average**
Category	**Brand buyer class**	**Growth Potential**		**Growth Potential**		
		N	%	N	%	
High Pen, High PF	Non	11,416	**31**	54,829	**77**	54
	Light	19,575	**52**	11,142	**16**	34
	Heavy	6,338	**17**	4,940	**7**	12
High Pen, Low PF	Non	11305	**38**	52459	**92**	65
	Light	16539	**55**	3990	**7**	31
	Heavy	2156	**7**	550	**1**	4
Low Pen, High PF	Non	4,598	**22**	30,028	**75**	49
	Light	10,187	**49**	6,299	**16**	33
	Heavy	6,172	**29**	3,492	**9**	19
Low Pen, Low PF	Non	3,758	**64**	10,694	**96**	80
	Light	1,556	**27**	348	**3**	15
	Heavy	536	**9**	73	**1**	5
Average	**Non**		**39**		**85**	62
	Light		**46**		**11**	28
	Heavy		**16**		**5**	10

The results in the graphs and Table 1 show that for a small brand, the largest source of growth potential by far is current non-buyers, followed by current light buyers. For a (very) large brand, current light buyers followed by non-buyers are the largest source of growth potential, with heavy buyers representing only 16% of growth potential. The reason why there is more sales potential for a big brand among its light brand buyers is simply that a large brand *has* many customers – therefore, there are fewer *non*-buyers as a source of potential sales (for example, if a brand had 90% of the population as customers, non-customers would represent a small source of potential sales).

We also checked the potential effect of higher brand loyalty on growth potential by simulating using a higher polarisation index (Rungie & Laurent, 2012; Schmittlein et al., 1985; Trinh & Dawes, 2020). Results are shown in Appendix 5. The results show that for a given brand, the higher the polarisation index (i.e higher loyalty), the higher the sales potential of the non-buyers and the lower the sales potential of the light and heavy buyers.

We now determine if this simulated outcome based on the NBD-BBD model holds when we use real brand purchasing data.

4. Empirical study

To check if the results of the simulation study reflect what occurs in the real world, we used panel consumer packaged goods purchasing data provided by Kantar. The Kantar panel consists of 12,407 households in the UK who continuously report their purchasing behaviour. For this study, we concentrate on data covering two consecutive years: from 2009 to 2011. We report both the actual growth potential and the estimated growth potential from the NBD-BBD model for different brands and categories. By 'actual' growth potential, we mean that we calculated the category purchases and brand purchases for brand non, light and heavy buyers from actual (observed) data. Appendix 6 shows details of the brands and categories used in our analysis and the parameters estimated by the NBD-BBD model. Figure 2 illustrates the results for the toothpaste category where Colgate is the market leader with a 40% market share and Arm+Hammer is a small brand with a 5% market share. Results for all categories are showed in Table 2 with the growth potential of different brand buyer groups in percentages.

Figure 2. Growth potential of different brand buyer classes across different brands in the toothpaste category: Model-Estimated vs. Actual

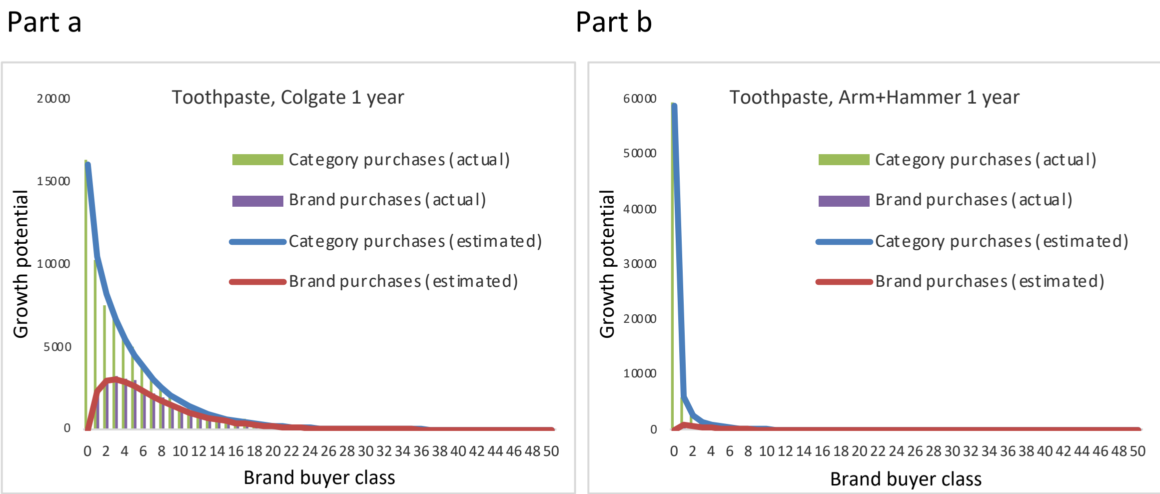


Table 2. Growth potential of brand non, light and heavy buyers: Model-Estimated vs. Actual

			Growth Potential (%)				
			Large brand		**Small** brand		
Category	Category penetration and purchase frequency	Brand buyer class	Model-Estimated	Actual	Model-Estimated	Actual	Average (actual)
Pasta sauce		Non	49	45	68	70	58
High Pen	Pen=74%	Light	37	42	23	22	32
High PF	Freq=10.4	Heavy	14	14	9	8	11
Toothpaste		Non	37	38	85	86	62
High Pen	Pen=91%	Light	50	49	11	10	30
Low PF	Freq=6.4	Heavy	14	13	4	3	8
Beer		Non	60	55	70	72	64
Low Pen	Pen=56%	Light	28	29	19	18	24
High PF	Freq=7.2	Heavy	12	16	11	10	13
Razor Blades		Non	68	70	96	95	83
Low Pen	Pen=49%	Light	24	22	3	4	13
Low PF	Freq=2.4	Heavy	8	8	1	1	5
Average		**Non**	**53**	**52**	**80**	**81**	66
		Light	**35**	**36**	**14**	**14**	25
		Heavy	**12**	**13**	**6**	**5**	9

Consistent with the simulation study, growth potential in all conditions comes mostly from non-buyers or light buyers of the brand (66% and 25% of sales potential, respectively). In relation to brand size, we see that the growth potential from current light and heavy buyers is more significant for larger brands. However, growth potential from current non-buyers is still dominant. That said, for a very large brand such as Colgate, growth potential among current light buyers is comparatively greater than among non-buyers.

To validate the model, we calculate the observed year 2 growth potential of those who are non, light and heavy buyers of the brand in year 1. We then compare these observed results with the future growth potential predicted by the NBD-BBD model. Appendices 7 and 8 show the results. The model accurately predicts future growth potential, providing confidence that using the NBD-BBD model is sound.

5. Conclusions, limitations, and directions for future research

Previous studies have investigated the importance of heavy and light buyers to current brand sales. Specifically, a brand's heaviest 20% of buyers account for 50 to 73% of its sales. While not as extreme as the 20:80 rule, there is no doubt that to maintain the current level of brand sales, heavy buyers play a significant role. However, when a brand looks for sources of growth, this study shows that the sales potential for a brand does not reside among current heavy brand buyers. Both the simulation and empirical studies in this paper support that there is vastly more 'headroom for growth' among light brand buyers than heavies for big brands; and by far the most headroom for growth for small brands resides among current non-brand buyers. Current heavy buyers represent a very minor source of sales potential for either big or small brands, due to two simple facts: (a) they are a minority of the brand's customer base; and (b) the brand is already representing a large proportion of their category purchasing, so they cannot buy the brand much more than they do currently.

For marketers looking to grow their brand, this study provides evidence for the need to market beyond the brand's heavy buyers. Brands should aim to obtain additional share of category purchasing from the large pool of households who buy the brand occasionally, as well as also recruiting consumers who are currently non-buyers.

This study examined the sales potential of different buyer groups. The method was straightforward, albeit to the best of the author's knowledge has not been utilised in prior work, namely to investigate the extent of *category* buying among various *brand* buyer groups. No examination of actual brand growth over time was conducted. However, a future direction for this research is to identify brands that grew in sales over a time period, and ascertain the extent to which their sales growth was obtained from the buyer groups in the manner predicted by this study. Future research could also investigate the sales potential of different buyer groups for new brands that ultimately succeed or fail.

Another important consideration in relation to brand growth concerns the time period over which growth occurs. Managers should be interested in the management of brands over the long term (Ataman et al., 2010; Dawes et al., 2020; Golder, 2000; Golder & Tellis, 1993; Mitra & Golder, 2006; Trinh & Anesbury, 2015). Therefore, future research could examine the sales potential of current non-buyers as well as light and heavy buyers, to brand growth over long term (i.e five years or more).

The context of this study is consumer packaged goods, yet the broad approach could be used for quite different markets such as banking or insurance. In these markets, consumers often deal with more than one brand (e.g. Dawes, 2014; Mundt et al., 2006). Therefore, as per approaches outlined in Du et al (2007) it would be possible for managers in service categories to obtain information via surveys, pertaining to the brand and category purchasing of its customers and non-customers, and calculate the sales potential of its light and heavy buyers as well as current non-buyers. The resultant information would allow more informed growth and targeting strategies.

Our focus has been to describe the 'lay of the land' in respect of the sales potential of non-, light, and heavy buyers. We find that empirically, the greatest sales potential resides with current non-or light brand buyers. We examine the context of a stationary category, but given that there are predictable patterns of category growth (Dunn et al., 2021; Nenycz-Thiel et al., 2018), it is likely that our findings will hold for growing or declining categories also. Managers can use this information, and take additional information into account, such as company capabilities, the cost of servicing or reaching a buyer segment in formulating their strategy. Lastly, the focus of this study is sales. The question of the profit potential of buyer groups has not been examined and is a direction for future research.

References

Anesbury, Z. W., Bennett, D., & Kennedy, R. (2021). How persistent are duplication of purchase partitions? *Journal of Consumer Behaviour*, *Ahead-of-print*(Ahead-of-print). https://doi.org/https://doi.org/10.1002/cb.1985

Anesbury, Z. W., Davies, C., Driesener, C., Page, B., Greenacre, L., Yang, S., & Bruwer, J. (2022). Death by 1000 'true fans': Do marketing laws apply to music listening? *Journal of Consumer Behaviour*, 1-16. https://doi.org/https://doi.org/10.1002/cb.2114

Ataman, M. B., Van Heerde, H. J., & Mela, C. F. (2010). The long-term effect of marketing strategy on brand sales. *Journal of Marketing Research*, *47*(5), 866-882.

Baker, B. J., McDonald, H., & Funk, D. C. (2016). The uniqueness of sport: Testing against marketing's empirical laws. *Sport Management Review*, *19*(4), 378-390. https://doi.org/https://doi.org/10.1016/j.smr.2016.02.002

Barnard, N., & Ehrenberg, A. (1997). Advertising: Strongly Persuasive or Nudging? *Journal of Advertising Research*, *37*, 21-28.

Bhattacharya, C. B., Fader, P. S., Lodish, L. M., & Desarbo, W. S. (1996). The Relationship Between the Marketing Mix and Share of Category Requirements. *Marketing Letters*, *7*(1), 5-18.

Bowman, D., & Narayandas, D. (2001). Managing customer-initiated contacts with manufacturers: The impact on share of category requirements and word-of-mouth behavior. *Journal of Marketing Research*, *38*(3), 281-297.

Casteran, G., Chrysochou, P., & Meyer-Waarden, L. (2019). Brand loyalty evolution and the impact of category characteristics. *Marketing Letters*, *30*(1), 57-73. https://doi.org/10.1007/s11002-019-09484-w

Chowdhury, S., Barker, A., Trinh, G., & Lockshin, L. (2021). Using the Dirichlet model to predict how fashion brands compete and grow on eBay. *Journal of Consumer Behaviour*, *3*(4), 357-373.

Dawes, J. (2014). Patterns in competitive structure among retail financial services brands. *European Journal of Marketing*, *48*(5-6), 7.

Dawes, J., Graham, C., & Trinh, G. (2020). The long-term erosion of repeat-purchase loyalty. *European Journal of Marketing*, 1-27. https://doi.org/10.1108/EJM-01-2018-0042

Dawes, J., Graham, C., Trinh, G., & Sharp, B. (2022). The Unbearable Lightness of Buying. *Journal of Marketing Management*, *38*(7-8), 683-708. https://doi.org/https://doi.org/10.1080/0267257X.2021.1963308

Dekimpe, M. G., & Hanssens, D. M. (1995). Empirical Generalizations About Market Evolution and Stationarity. *Marketing Science*, *14*(No. 3, Part 2), G109-G121.

Driesener, C., & Rungie, C. (2022). The Dirichlet model in marketing. *Journal of Consumer Behaviour*, *21*(1), 7-18. https://doi.org/https://doi.org/10.1002/cb.1975

Driesener, C., Rungie, C., & Wright, M. (2022). Dirichlet implications for portfolio management. *Journal of Consumer Behaviour*, *21*(1), 49-62. https://doi.org/https://doi.org/10.1002/cb.1996

Driesener, C., Rungie, C. E., & Banelis, M. (2017). If the Model Fits, Use It: Methods and Benchmarks for Evaluating NBD-Dirichlet Goodness-of-Fit. *Australasian marketing journal*, *25*(4), 288-293. https://doi.org/https://doi.org/10.1016/j.ausmj.2017.10.003

Du, R. Y., Wagner, K. A., & Mela, C. F. (2007). Size and share of customer wallet. *Journal of Marketing*, *71*(April), 94-113.

Dunn, S., Graham, C., Nenycz-Thiel, M., & Tanusondjaja, A. (2021). Investigating Undercurrents of Stationarity and Growth With Long-Term Panel Data. *International Journal of Market Research*, *63*(6), 786-809. https://doi.org/https://doi.org/10.1177/14707853211039189

East, R., & Hammond, K. (1996). The Erosion of Repeat-Purchase Loyalty. *Marketing Letters*, *7*(2), 163-171.

Ehrenberg, A. (1988). *Repeat-buying: Facts, theory and applications*. Oxford University Press.

Ehrenberg, A., Uncles, M. D., & Goodhardt, G. J. (2004). Understanding brand performance measures: using dirichlet benchmarks. *Journal of Business Research*, *57*(12), 1307-1325. https://doi.org/10.1016/j.jbusres.2002.11.001

Golder, P. N. (2000). Historical Method in Marketing Research with New Evidence on Long-Term Market Share Stability. *Journal of Marketing Research*, *37*(May), 156-172, Article 2.

Golder, P. N., & Tellis, G. J. (1993). Pioneer Advantage: Marketing Logic or Marketing Legend? *Journal of Marketing Research*, *30*(May), 158-170.

Goodhardt, G. J., Ehrenberg, A., & Chatfield, C. (1984). The Dirichlet: A comprehensive model of buying behaviour. *Journal of the Royal Statistical Society*, *147*(5), 621-643. https://doi.org/10.2307/2981696

Graham, C., & Kennedy, R. (2022). Quantifying the target market for advertisers. *Journal of Consumer Behaviour*, *21*(1), 33-48. https://doi.org/https://doi.org/10.1002/cb.1986

Grasby, A., Corsi, A., Dawes, J., Driesener, C., & Sharp, B. (2022). How loyalty extends across product categories. *Journal of Consumer Behaviour*, *21*(1), 153-163. https://doi.org/https://doi.org/10.1002/cb.1981

Hallberg, G. (1995). *All consumers are not created equal*. John Wiley.

Hallberg, G. (1999). The Pareto Imperative. *DMA Insider*(Winter 1999).

Hess, M., & Doe, P. (2013). The Marketer's Dilemma: Focusing on a Target or a Demographic? The Utility of Data- Integration Techniques. *Journal of Advertising Research*, *53*(2), 231-236.

Keiningham, L., Buoye, A., & Cooil, B. (2011). Customer Loyalty Isn't Enough. Grow Your Share of Wallet. *Harvard Business Review*, *89*(10), 29-31.

Kim, B. J., Singh, V., & Winer, R. S. (2017). The Pareto rule for frequently purchased packaged goods: an empirical generalization. *Marketing Letters*, *28*(4), 1-17. https://doi.org/https://link.springer.com/article/10.1007/s11002-017-9442-5

Lam, D., & Mizerski, R. (2009). An investigation into gambling purchases using the NBD and NBD-Dirichlet models. *Marketing Letters*, *20*(3), 263-276.

Lam, D., & Ozorio, B. (2013). Duplication of Purchase Law in the gaming entertainment industry—A transnational investigation. *International Journal of Hospitality Management*, *33*, 203-207.

Lynn, M. (2013). Lessons from Duplication of Purchase Data. *Cornell Hospitality Report*, *13*(3), 4-16.

Martin, J., Nenycz-Thiel, M., Dawes, J., Tanusondjaja, A., Cohen, J., McColl, B., & Trinh, G. (2020). Fundamental basket size patterns and their relation to retailer performance. *Journal of Retailing and Consumer Services*, *54*, 102032. https://doi.org/10.1016/j.jretconser.2020.102032

McCarthy, D. M., & Winer, R. S. (2019). The Pareto rule in marketing revisited: is it 80/20 or 70/20? *Marketing Letters*, 1-12. https://doi.org/10.1007/s11002-019-09490-y

Meyer-Waarden, L., & Benavent, C. (2006). The impact of loyalty programmes on repeat purchase behaviour. *Journal of Marketing Management*, *22*(1/2), 61-88. https://doi.org/10.1362/026725706776022308

Mitra, D., & Golder, P. N. (2006). How Does Objective Quality Affect Perceived Quality? Short-Term Effects, Long-Term Effects, and Asymmetries. *Marketing Science*, *25*(3), 230-247. http://ezlibproxy.unisa.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=ufh&AN=21515554&site=ehost-live

Morrison, D. G., & Schmittlein, D. C. (1988). Generalizing the NBD model for customer purchases: What are the implications and is it worth the effort? *Journal of Business & Economic Statistics*, *6*(2), 145-159. https://doi.org/https://doi.org/10.2307/1391551

Mundt, K., Dawes, J., & Sharp, B. (2006). Can a brand outperform competitors on cross-category loyalty? An examination of cross-selling metrics in two financial services markets. *Journal of Consumer Marketing*, *23*(7), 465-469. https://doi.org/http://dx.doi.org/10.1108/07363760610713019

Nenycz-Thiel, M., McColl, B., Dawes, J., Trinh, G., & Graham, C. (2018). *Predicting Category Growth from Quarterly Penetration* ANZMAC 2018, Adelaide, Australia

Pare, V., & Dawes, J. (2011). The persistence of excess brand loyalty over multiple years. *Marketing Letters*, *21*(2), 163-175.

Reilly, M., & Rapacz Deb. (2009). Heavy buyers: are they even more important than generally thought? *Journal of Integrated Marketing Communications*, 1-8.

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.

Rungie, C., & Laurent, G. (2012). Brand loyalty vs. loyalty to product attributes. In A. Diamantopoulos (Ed.), *Quantitative Marketing and Marketing Management* (pp. 423-444). Springer.

Rungie, C. M., Uncles, M. D., & Laurent, G. (2013). Integrating consumer characteristics into the stochastic modelling of purchase loyalty. *European Journal of Marketing*, *47*(10), 1667-1690.

Schmittlein, D. C., Bemmaor, A. C., & Morrison, D. G. (1985). Why does the NBD model work? robustness in representing product purchases, brand purchases and imperfectly recorded purchases. *Marketing Science*, *4*(No. 3, Summer), 255-266.

Schmittlein, D. C., Cooper, L. G., & Morrison, D. G. (1993). Truth in Concentration in the Land of (80/20) Laws. *Marketing Science*, *12*(2), 167-183. https://doi.org/https://doi.org/10.1287/mksc.12.2.167

Sharp, B. (2010). How brands grow. In B. Sharp (Ed.), *How Brands Grow* (pp. 16-27). Oxford University Press.

Sharp, B., Romaniuk, J., & Graham, C. (2019). *Marketing's 60/20 Pareto Law* (Available at SSRN 3498097, Issue.

Sharp, B., Wright, M., Dawes, J., Driesener, C., Meyer-Waarden, L., Stocchi, L., & Stern, P. (2012). It's a Dirichlet world: Modeling individuals' loyalties reveals how brands compete, grow, and decline. *Journal of Advertising Research*, *52*(2), 203-213.

Smith, A., & Blair, N. (2018, 22 January). *Dunnhumby on why 'heavy' buyers matter more than Byron Sharp says*. Campaign. Retrieved 16 October from

https://www.campaignlive.co.uk/article/dunnhumby-why-heavy-buyers-matter-byron-sharp-says/1455175
Stern, P., & Hammond, K. (2004). The Relationship Between Customer Loyalty and Purchase Incidence. *Marketing Letters*, *15*(1), 5-19.
Taussig, A. (2016). To solve your marketing woes, focus on the heavy buyers. *Medium*. https://medium.com/lightspeed-venture-partners/to-solve-your-marketing-woes-focus-on-the-heavy-buyers-dd6dd76f6894
Trinh, G., Anesbury, Z., & Driesener, C. (2017). Has behavioural loyalty to online supermarkets declined? *Australasian Marketing Journal (AMJ)*, *25*(4), 326-333. https://doi.org/https://doi.org/10.1016/j.ausmj.2017.10.005
Trinh, G., Corsi, A. M., & Lockshin, L. (2019). How country of origins of food products compete and grow. *Journal of Retailing and Consumer Services*, *49*, 231-241.
Trinh, G., & Dawes, J. (2020). A comparison of brand loyalty between on the go and take-home consumption purchases. *Journal of Retailing and Consumer Services*, *53*, 101968. https://doi.org/https://doi.org/10.1016/j.jretconser.2019.101968
Trinh, G., & Lam, D. (2016). Understanding the attendance at cultural venues and events with stochastic preference models. *Journal of Business Research*, *69*(9), 3538-3544. https://doi.org/https://doi.org/10.1016/j.jbusres.2016.01.033
Trinh, G., Romaniuk, J., & Tanusondjaja, A. (2016). Benchmarking buyer behavior towards new brands. *Marketing Letters*, *27*(4), 743-752. https://doi.org/10.1007/s11002-015-9376-8
Trinh, G. T., & Anesbury, Z. W. (2015). An investigation of variation in brand growth and decline across categories. *International Journal of Market Research*, *57*(3), 347-356.
Twedt, D. W. (1964). How Important to Marketing Strategy is the "Heavy User"? *Journal of Marketing*, *28*(1), 71-72. https://doi.org/ https://doi.org/10.2307/1249232
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.
Uncles, M. D., Ehrenberg, A., & Hammond, K. (1995). Patterns of buyer behavior: Regularities, models, and extensions. *Marketing Science*, *14*(3), G71-G78. https://doi.org/10.1287/mksc.14.3.G71

Online supplemental Appendices

Appendix 1. NBD-Dirichlet and NBD-BBD

The NBD-Dirichlet assumes that purchasing rate *n* of a product category of a given consumer in successive time periods follows a Poisson distribution with parameter λ

$$f(n) = \frac{\exp{(-\lambda)}\lambda^n}{n!} \tag{1}$$

with mean

$$E[n] = \lambda$$

and the mean frequencies of purchasing a product category λ of different consumers in the long run differ and their distribution is a gamma distribution.

$$f(\lambda; k, a) = \lambda^{k-1}\frac{\exp{(-\frac{\lambda}{a})}}{a^k\Gamma(k)} \tag{2}$$

where *k* and *a* are the shape and scale parameters of the gamma distribution, respectively.

Combining (2) and (1), the probability density function of purchase frequency *n* is

$$f(n) = \frac{(1+a)^{-k}\Gamma(n+k)}{n!\,\Gamma(k)}(\frac{a}{1+a})^n \tag{3}$$

This is the Negative Binomial Distribution (the NBD part of the model).

The model also assumes that the choice among different brands in the product category follows the Dirichlet multinomial distribution (the Dirichlet part of the model).

The probability density function of the Dirichlet is

$$f_{\alpha_1,\alpha_2,\ldots,\alpha_m}(x_1, x_2, \ldots x_m | x_1 + x_2 + \cdots + x_m = n) = \frac{\frac{\Gamma(s)n!}{\Gamma(s+n)}\Pi_{i=1}^{m}\Gamma(\alpha_i + x_i)}{x_i!\,\Gamma(\alpha_i)} \tag{4}$$

where *m* is the number of brands, *x* is brand purchase frequency, α_i is parameter of the DMD distribution, and *s* is the sum of α_i.

Combining (3) and (4) the probability density function of the NBD-Dirichlet model is

$$f_{k,a,\alpha_1,\alpha_2,\ldots,\alpha_m}(x_1, x_2, \ldots x_m) = \frac{(1+a)^{-k}\Gamma(n+k)}{n!\,\Gamma(k)}(\frac{a}{1+a})^n\frac{\frac{\Gamma(s)n!}{\Gamma(s+n)}\Pi_{i=1}^{m}\Gamma(\alpha_i + x_i)}{x_i!\,\Gamma(\alpha_i)}$$

(5)

Using the parameters k, a, s and α_i, we can simulate brand purchases and category purchases for different classes of brand buyers. For the case of one brand versus total purchases of the category, the Beta Binomial Distribution (BBD) has only two parameters α (brand) and β (all other brands), where $s = \alpha + \beta$. The Dirichlet embodies certain assumptions, including that a buyer's choice of brand is independent of which other brands they buy (Goodhardt et al., 1984). These assumptions result in a predictable pattern called the 'duplication of purchase law' for which there is considerable empirical support (Anesbury et al., 2021; Anesbury et al., 2022; Baker et al., 2016; Grasby et al., 2022; Lam & Ozorio, 2013; Lynn, 2013; Meyer-Waarden & Benavent, 2006; Trinh et al., 2019). An important consequence of the duplication of purchase law is that it implies that any brand sales growth will come proportionally from all other brands in the product category.

Appendix 2. Sales potential of buyer groups (Simulation)

A	B	C	D	E	F	G	H	I
Buyer group (frequency of purchasing the brand per year currently)	Number of buyers (or non-buyers in the zero case)	Average frequency of purchasing the **category** per year	Current brand purchases **(Column A times Column B)**	Current category purchases **(Column B times Column C)**	Current sales potential **(Column E minus Column D)**	Future brand purchases (next year)	Future category purchases (next year)	Future sales potential **(Column H minus Column G)**
0 (non brand buyers)	8696.2	6.3	0.0	54829.3	**54829.3**	578.4	55407.7	**54829.3**
1	636.9	11.1	636.9	7037.9	**6401.0**	493.5	6894.5	**6401.0**
2	246.7	14.1	493.5	3470.9	**2977.4**	391.0	3368.4	**2977.4**
3	130.3	16.5	391.0	2154.2	**1763.1**	316.7	2079.9	**1763.1**
4	79.2	18.7	316.7	1479.5	**1162.8**	260.8	1423.6	**1162.8**
5	52.1	20.7	260.7	1077.3	**816.5**	217.4	1033.9	**816.5**
6	36.2	22.5	217.4	815.4	**598.0**	183.0	781.0	**598.0**
7	26.1	24.2	183.0	633.9	**450.9**	155.3	606.2	**450.9**
8	19.4	25.9	155.3	502.9	**347.6**	132.7	480.3	**347.6**
9	14.7	27.5	132.7	405.5	**272.8**	114.0	386.8	**272.8**
10	11.4	29.1	114.0	331.3	**217.3**	98.5	315.8	**217.3**
11	8.9	30.6	98.4	273.4	**175.0**	85.4	260.4	**175.0**
12	7.1	32.0	85.3	227.7	**142.3**	74.3	216.6	**142.3**
13	5.7	33.4	74.3	191.0	**116.8**	64.9	181.6	**116.8**
14	4.6	34.8	64.8	161.3	**96.5**	56.8	153.3	**96.5**
15	3.8	36.2	56.8	137.0	**80.2**	49.9	130.1	**80.2**
16	3.1	37.5	49.8	116.9	**67.1**	43.9	111.0	**67.1**
17	2.6	38.8	43.9	100.2	**56.4**	38.7	95.1	**56.4**
18	2.1	40.1	38.7	86.3	**47.6**	34.2	81.8	**47.6**
19	1.8	41.4	34.2	74.5	**40.3**	30.3	70.6	**40.3**
20+	10.7	50.1	284.2	535.9	**251.7**	254.6	506.3	**251.7**

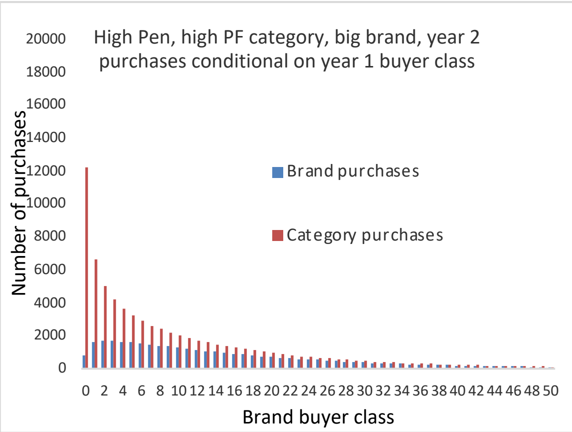
Appendix 2 shows an example of the simulated data. We use the example of a small brand in a frequently bought category to illustrate the scope of the simulation. The brand's market share is 5%, the category purchase frequency is 10, the category penetration is 75%. As Appendix 2 shows, the sales potential of each buyer group depends on (1) how many buyers are in that group (2) how frequently they buy the brand, and (3) how frequently they buy the category. The headroom for growth is the group's total category purchasing, minus their brand purchasing. For example, there is sales potential of 54,829 purchases per year among those who do not buy the brand currently, but buy the product category. However, while the brand's 20+ buyers buy the category very frequently (50 occasions per year), there are only 11 of them, and they already make 284 purchases of the brand as part of 536 category purchases, leaving little growth potential. We also simulate future brand and category purchases of these buyer groups and find the same results in terms of growth potential. We see that the sales potential is by far the greatest among current non-brand buyers, principally because there are so many of them.

Appendix 3 Simulation scenarios

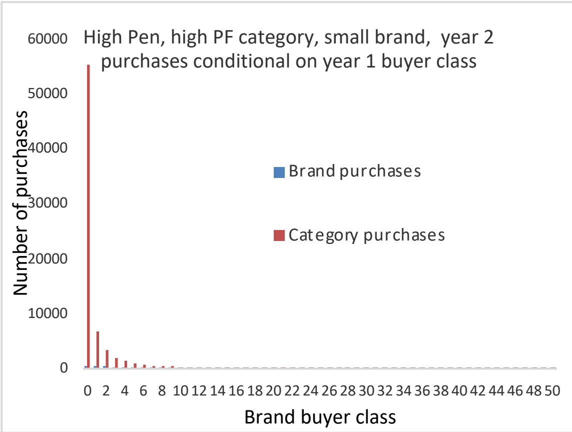
	Category penetration	Category purchase frequency	Brand size	Model parameters (k, a, s, α_i)
1	High (75%)	High (10)	Big (50% share)	0.5, 15, 1, 0.5
2	High (75%)	High (10)	Small (5% share)	0.5, 15, 3, 0.15
3	High (97%)	Low (6.2)	Big (50% share)	4, 1.5, 1.2, 0.6
4	High (97%)	Low (6.2)	Small (5% share)	4, 1.5, 1, 0.05
5	Low (56%)	High (7.6)	Big (50% share)	0.3, 14, 1.8, 0.9
6	Low (56%)	High (7.6)	Small (5% share)	0.3, 14, 4.4, 0.22
7	Low (49%)	Low (2.4)	Big (50% share)	0.65, 1.8, 1, 0.5
8	Low (49%)	Low (2.4)	Small (5% share)	0.65, 1.8, 0.8, 0.04

Appendix 4. Future brand and category purchases of different brand buyer classes in a high penetration and high purchase frequency category

Part a



Part b



Appendix 5. Growth potential of brand among non, light and heavy buyers: Simulation results using a higher polarization index

		Growth Potential = Total category purchases – Total brand purchases				
Category	Brand buyer class	Large brand		Small brand		Overall average
		N	%	N	%	%
High Pen, High PF	Non	18711	**50**	60242	**85**	68
	Light	14943	**40**	7631	**11**	26
	Heavy	3678	**10**	3037	**4**	7
High Pen, Low PF	Non	16675	**56**	54260	**95**	76
	Light	11845	**39**	2455	**4**	22
	Heavy	1479	**5**	285	**1**	3
Low Pen, High PF	Non	7829	**37**	32949	**83**	60
	Light	9315	**44**	4820	**12**	28
	Heavy	3814	**18**	2049	**5**	12
Low Pen, Low PF	Non	4436	**76**	10860	**98**	87
	Light	1064	**18**	204	**2**	10
	Heavy	350	**6**	51	**0**	3
Average	**Non**		**55**		**90**	73
	Light		**35**		**7**	21
	Heavy		**10**		**3**	6

This appendix shows the simulation results when the polarization index (p) of the Beta distribution increases. As p = 1/(1+s), to increase p we reduce both parameters s and α_i in Appendix 3 by half.

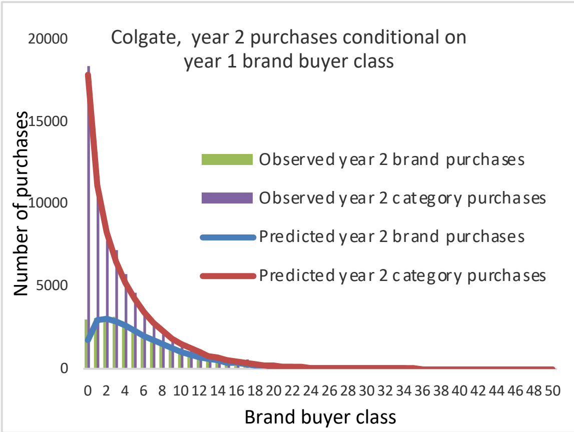
Appendix 6. Empirical scenarios

Category	Category penetration (%)	Category purchase frequency	Brand	Market Share (%)	Brand penetration (%)	Brand purchase frequency	Model estimated parameters (k, a, s, α_i)
Pasta sauces	74	10.4	Dolmio*	27	38	5.5	0.47, 16.26, 1.12, 0.30
Pasta sauces	74	10.4	Ragu	8	19	3.2	0.47, 16.26, 2.77, 0.21
Toothpaste	91	6.4	Colgate*	40	62	3.7	1.48, 3.93 1.64, 0.65
Toothpaste	91	6.4	A+H	5	13	2.1	1.48, 3.93 2.88, 0.13
Beer	56	7.2	Stella Artois*	17	20	3.4	0.31, 12.97 1.75, 0.29
Beer	56	7.2	Budweiser	6	12	2.2	0.31, 12.97 4.53, 0.29
Razor Blades	49	2.4	Gillette*	43	26	1.9	0.64, 1.83 1.03, 0.44
Razor Blades	49	2.4	Tesco	5	3	1.7	0.64, 1.83 0.85, 0.04

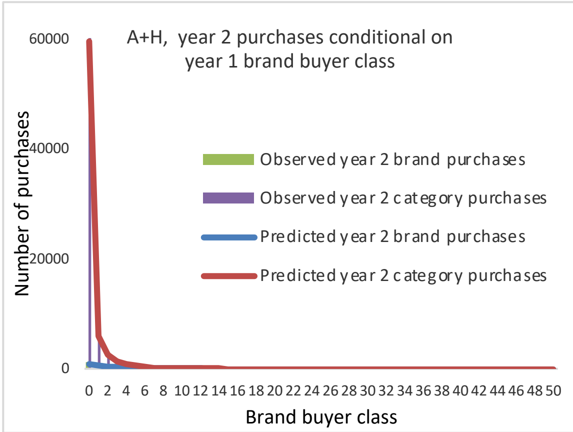
(*) These brands are market leaders in their categories

Appendix 7. Predicted vs. Observed Future Purchases

Part a



Part b



Appendix 8. Future growth potential of brand among non, light and heavy buyers: Model Predicted v.s Observed

		Future growth potential = Y2 category purchases conditional on Y1 brand buyer class – Y2 brand purchases conditional on Y1 brand buyer class								
		Large brand				Small brand				**Overall average**
Category	**Brand buyer class**	**Predicted**	**Observed**	**Predicted**	**Observed**	**Predicted**	**Observed**	**Predicted**	**Observed**	**Observed**
Pasta sauce		N	N	%	%	N	N	%	%	%
High Pen, High PF	Non	33827	30637	**49**	**47**	59657	60540	**68**	**69**	**58**
	Light	25860	25945	**37**	**40**	20296	19622	**23**	**22**	**31**
	Heavy	9584	9057	**14**	**14**	7669	7646	**9**	**9**	**11**
Toothpaste										
High Pen, Low PF	Non	16065	15392	**37**	**36**	58826	58899	**85**	**85**	**60**
	Light	21549	21489	**50**	**50**	7349	7496	**11**	**11**	**30**
	Heavy	5899	6421	**14**	**15**	2703	2620	**4**	**4**	**9**
Beer										
Low Pen, High PF	Non	25165	22375	**60**	**55**	33311	32326	**70**	**73**	**64**
	Light	11964	11362	**28**	**28**	8976	7637	**19**	**17**	**23**
	Heavy	5096	6662	**12**	**16**	5100	4350	**11**	**10**	**13**
Razor Blades										
Low Pen, Low PF	Non	5645	5429	**68**	**69**	13412	11931	**96**	**94**	**82**
	Light	2043	1753	**24**	**22**	435	607	**3**	**5**	**14**
	Heavy	661	674	**8**	**9**	91	102	**1**	**1**	**5**
Average	**Non**			**53**	**52**			**80**	**80**	**66**
	Light			**35**	**35**			**14**	**14**	**24**
	Heavy			**12**	**13**			**6**	**6**	**10**