

# Do higher spending households buy a greater variety of brands? : An Application of Repertoire Regression

*Melissa Banelis, Cam Rungie, Erica Riebe and Lars Meyer-Waarden*

In using panel data it is often difficult to determine the optimal number of purchases needed to determine the size of an individual's repertoire. Clearly, as more purchases are made from the category, the size of a buyer's repertoire becomes more apparent. Thus the number of purchases a buyer makes from the category should be correlated to the size of their repertoire. Heavier buyers of a category, spread their purchases amongst more brands. If it were possible to standardise for the number of purchases buyers make from the category, do they then have the same size repertoire? Is the category purchase rate the only factor that affects repertoire size? This paper uses a new modelling method for analysing repeated purchase patterns for a sample of buyers within a product category. The technique, Repertoire Regression (RR), fits covariates to repertoire size using Dirichlet Multinomial Distribution (DMD) regression. This paper reports on a specific study that investigates the relationship between standardized repertoire size and the covariate weekly household spend in supermarkets. The results show that for three product categories, store choice, detergents and shampoos, as weekly household spend in the category increases, standardized repertoire size remains approximately steady.

**Keywords:** Repertoire size, Dirichlet, loyalty.

## Introduction

Repertoire size is the number of brands a shopper purchases from a product category in a specified period. It is one way of measuring loyalty (Stern 2004). Shoppers with smaller repertoires in effect are more loyal since they buy from a smaller number of brands (Colombo & Jiang 2002).

Repertoire size is known to be a function of the category purchase rate (Banelis, Rungie & Colombo 2003; Colombo & Jiang 2002). In order to analyse variations in repertoire size relating to other characteristics of shoppers, the confounding effect of category purchase rate must be accounted for. Without removing the effect of category buying rate, it is difficult to determine whether any other factors have an impact on the size of a shopper's repertoire. It is important to consider whether there are other factors that affect the size of a shopper's repertoire as marketers may be able to manipulate these factors and therefore increase or decrease the overall rate of buyer loyalty for their brand. Knowing the particular variables that are correlated with repertoire size, and assuming some causal link between the two, managers may be able to manipulate these variables or target particular types of shoppers with their marketing efforts.

There are also clear implications for managers if the category purchase rate is the only variable that affects repertoire size. Marketers, particularly of small brands, may choose to promote the category, to ensure that repertoire sizes are large and their brand has a chance of being considered by the market.

To differentiate between the impact of category purchase rate and other variables on repertoire size, a model that involves standardization of the repertoire size must be used. A new technique for doing this is discussed below and is applied to four empirical data sets

capturing purchases in four distinct product categories. This paper addresses the research question “what happens to repertoire size if a household spends more per week on all supermarket items?”

## **The Repertoire Regression Model**

The Dirichlet Model has been used to describe and predict a variety of components of consumer behaviour (Ehrenberg 1988; Ehrenberg, Uncles & Goodhardt 2004; Goodhardt, Ehrenberg & Chatfield 1984). One of the most useful benefits of the model is that it can be used to generate theoretical brand performance measures (BPMs) for brands within a category, which can then be compared to the actual performance of the brands. Whilst it is a valuable descriptive model, the Dirichlet has attracted criticism for being limited in its ability to aid managerial decision-making. This criticism stems, in part, from the fact that marketing mix or consumer variables are not an integral part to the model. Managers may wish to explore hypothetical situations to determine where their efforts are best placed. For example, what happens to the size of a repertoire if it were possible to double advertising expenditure? In this paper the Repertoire Regression model is used as a starting point in addressing this criticism. By using such a model, we are aiming to provide some insight into the effect of shopper characteristics on the size of a repertoire.

The Repertoire Regression (RR) model uses the Dirichlet Multinomial Distribution (the “brand choice” part of the Dirichlet model) to construct a relationship between a given covariate and repertoire size. The temptation is to call these “explanatory variables” but we more correctly refer to them as covariates, as we do not purport any causal link. The covariates can be any descriptive variable recorded for each shopper in the data set. Some examples of covariates are age, income, size of household and weekly expenditure in supermarkets. This method is a special case of the Generalised Dirichlet model (Rungie & Laurent 2004). The Generalised Dirichlet model fits covariates to any brand performance measure the researcher may choose to examine. Thus, using this type of model researchers can see what the relationship is between the covariate and the brand performance measure of their choice. The RR model fits covariates to one brand performance measure, repertoire size, using the DMD and likelihood maximisation. This model allows the researcher to consider the relationship between any consumer attribute (covariate) and repertoire size. It is a special case of the Generalised Dirichlet model because it only looks at this one brand performance measure.

In our quest to understand how repertoire size varies with weekly household spend in supermarkets, we are presented with a specific challenge. The challenge arises due to the large effect of category purchase rate on a household’s repertoire size. An established belief is that as the category purchase rate for a particular shopper increases (they buy more), so does their repertoire size (they buy from a larger repertoire of brands) (Colombo & Jiang 2002). While the number of category purchases made will naturally allow buyers to demonstrate the size of their repertoire (the more purchases you make, the greater the possible variation in the size of the repertoire), it is also sensible that shoppers will limit their repertoire. No shopper has complete information about the brands available to them; therefore it makes sense that they will only buy from a small number of brands in the category regardless of how heavy a category buyer they are. Thus, if we analyse the effect of spend on repertoire without accounting for category buying rate, we may simply see repertoire size increasing as spend increases because of the strong effect of the category purchase rate. It is

for this reason that we have used standardized repertoire size in this paper. Standardized repertoire size accounts for the category buying rate by giving all shoppers a category buying rate equal to the average for the sample. This allows the researcher to see the true relationship between repertoire size and the covariate of interest.

In order to explain this occurrence more clearly, consider the following illustrative example. Consider two shoppers and their repertoires of supermarkets. The product category is supermarket stores and the repertoire size is the number of unique stores visited over a year in a market where there are seven supermarkets. The results here reflect the data in which purchases were recorded at all seven supermarket stores in a small city. The seven stores are analysed as if there were seven brands in the category. The seven stores actually represent five chains. The data is discussed in further detail in subsequent sections of this paper. As an example, consider the following. Household A, spends a large amount, say 900 francs on average per week on supermarket purchases. Household B, on the other hand spends only a small amount, say 100 francs per week. In this case, because household A spends more, they also visit stores more often and so there is much more opportunity for their repertoire of stores to be larger. The average repertoire size for household A is 3.7 stores per year compared to 3.2 stores per year for household B. However, what we do not know from simply studying their repertoire over say a year, is how we would expect these households to differ in the size of their repertoires if we restricted our analysis to the same number of store visits (i.e. the same number of purchases from the category). On average, shoppers in this data made 73 store visits per year. If both households were limited to 73 store visits, we find that on average household A would have a portfolio of 3.0 stores per year compared with 3.5 stores per year for household B. So once we standardize for category buying rate, households with higher expenditure actually have a smaller standardized repertoire size. Therefore, in this particular example, heavier users of the category, have a smaller repertoire, once we account for their category purchase rate. In this case, it seems, they are more loyal.

## The Data

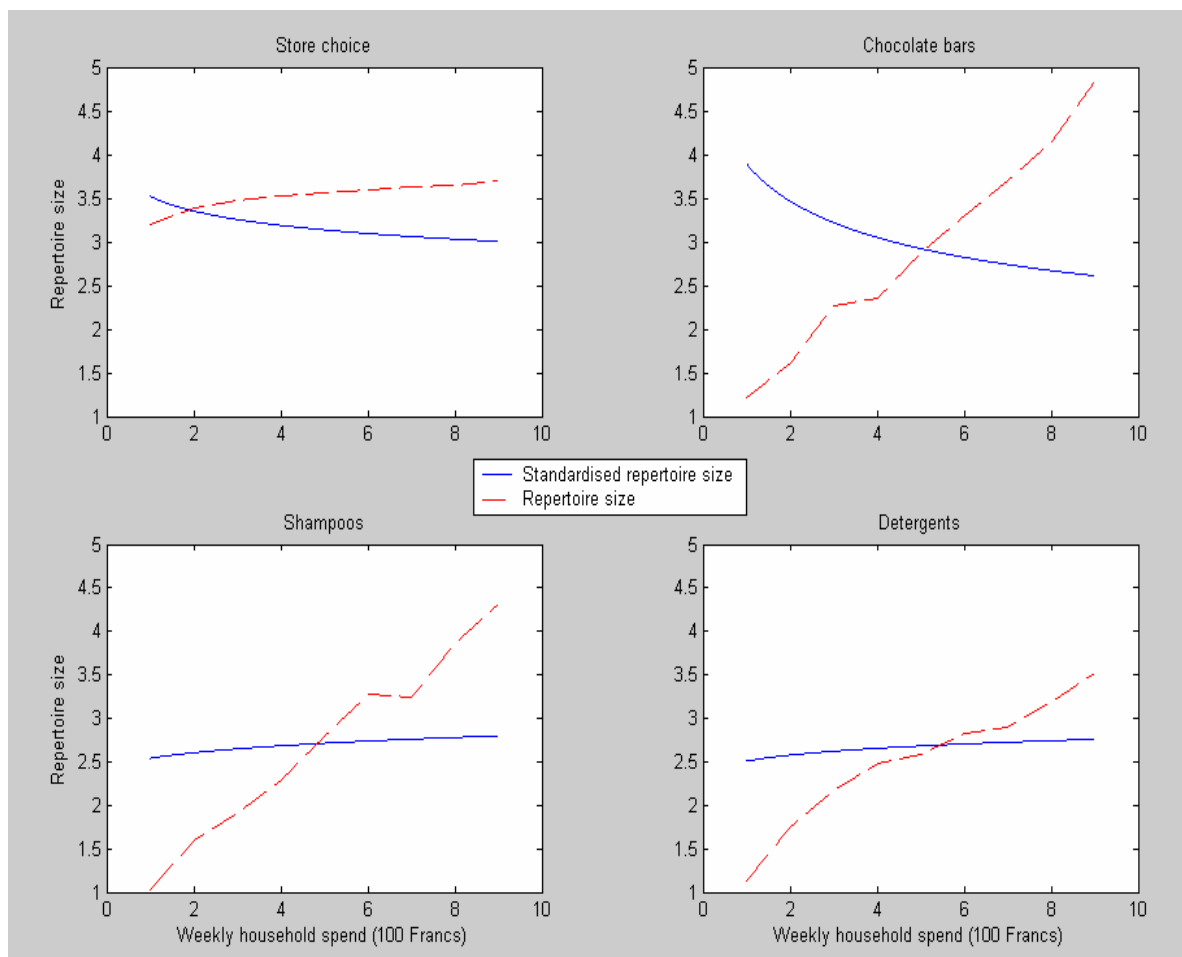
Marketing Scan provided purchase data for one year for 3,459 households in Angers, France. In one of the product categories there were 21,979 purchases in the period in which the data was collected. Purchases of all brands in three product categories, as well as store choice data was recorded over a one-year period in the late 1990s. The panellists also recorded information on a number of covariates. For the purposes of this paper, information for only one covariate was used. This covariate was average weekly spend by the household, in supermarkets per week. Descriptive statistics for the categories are given in Table 1. The purchase rate is the number of purchases from the category for one year averaged over all households in the sample including those households not purchasing from the category. The penetration is the percentage of households in the sample who purchased from the category over the one year.

## Results

Figure 1 shows repertoire size graphed as a function of household spend across all four product categories. As expected, in all plots, before standardisation, repertoire sizes increases as household spend increases. The standardization for each category is based in each case on the average repertoire size as reported in Table 1. Once the effect of the category purchase rate is removed, the effect of household spend on standardized repertoire size is reduced and is not always a positive, linear relationship.

**Table 1. Descriptive statistics for the data.**

| <b>Product Category</b>                       | <b>Chocolate bars</b> | <b>Shampoos</b> | <b>Detergents</b> | <b>Store choice</b> |
|---|-----------------------|-----------------|-------------------|---------------------|
| Number of 'brands' in the category            | 35                    | 26              | 22                | 7                   |
| Market Share of leading brand (%)             | 17                    | 11              | 15                | 18                  |
| Category purchase rate per household per year | 6.9                   | 5.4             | 6.4               | 72.9                |
| Category penetration per year                 | 0.68                  | 0.78            | 0.91              | 1.0                 |



**Figure 1. Repertoire size versus household spend for all four product categories**

*For store choice, shampoos and detergents, as spend increases standardized repertoire size remains approximately the same. For chocolate bars, as spend increases standardized repertoire size decreases slightly.*

For store choice, the overall change in standardized repertoire size is around 0.5 of a store, dropping from 3.5 stores for the households with low spend, down to 3.0 stores as household

spend increases. There is a similar result in the second plot with the standardized repertoire size for chocolate bars changing from 4 brands down to 2.5. For the other two categories, shampoos and detergents, as household spend increased, standardized repertoire size remained stable. While there are slight differences between the markets in the relationship between household spend and repertoire size, generally we can say that once category buying rate is accounted for, households with low spend in supermarkets and high spend in supermarkets have approximately the same sized repertoire. Category buying rate alone accounts for the size of their repertoire. We see only one real exception to this result in this data. That is, for chocolate bars, as household spend increases, standardised repertoire size decreases (heavier shoppers buy from a smaller repertoire).

## Discussion

As households spend more on all supermarket items, their repertoire size increases but this is not because of some aspects of their behaviour related to their spending level and their loyalty towards brands in particular categories rather it is simply because the increased number of purchases allows heavier category buyers to fully demonstrate the size of their repertoire. The direct relationship between number of purchases observed and repertoire size can lead to misinterpretation of the level of brand loyalty in a category. In order to understand the implications of such a misinterpretation, consider the situation without standardization. Had the analysis in this paper been conducted without standardisation it would show that as household spend increases, repertoire size increases. This implies that households with larger spend are less loyal than those with a smaller spend. For brand managers, this may result in them focussing less on these higher spending households and more on the lower spending households. Once we standardise for category purchase incidence, we actually see that such a strategy may be incorrect. Our analysis would suggest that while heavy supermarket buyers spread their purchases out amongst a range of brands, they either do this as much as we would expect them to given their category buying rate, or less than we might expect in the case of chocolate bars. By using standardized repertoire in the repertoire regression model, the chance of such misinterpretation of buyer loyalty to their repertoire is eliminated because it shows the relationship between repertoire size and household spend without the effect of category purchase rate. The model that we used in this paper therefore has implications for managers as well as for marketing researchers.

The results of the analysis also have implications for brand managers. We have found that, with the exception of chocolate bars, weekly spend in supermarkets is not a determinate in the size of a buyer's repertoire. The reasons for such a result may be many and varied. One such reason may be that these categories are what most researchers would refer to as low involvement categories. Choices made in these categories are not made with a great deal of thought. For store choice, detergents and shampoos, as consumers spend more in supermarkets, the size of their repertoires stay approximately the same. Chocolate bars, the one exception to our general finding, may be considered to be different in terms of the level of involvement that buyers have with the category. Maybe shoppers are more involved in this category than they are in other categories? Even consumers with small supermarket spend may be more interested in learning about attributes of chocolate bars than they are to learn about attributes of detergents. Consequently the standardized repertoire for low spending households may be higher.

Another possible explanation for the negative relationship between standardised repertoire size and household spend in supermarkets may relate to how the product is sold in this

category. Households with a larger spend are able to buy chocolate in bulk packets, therefore reducing the chances of other brands entering the buyer's repertoire. Households with higher spend are buying more in the supermarket which leads to the purchase of packets of chocolate bars. This in turn leads to a smaller repertoire for these consumers because they are purchasing what they need from a small number of brands.

A final potential reason for the different result for the chocolate bar category is that there are far more brands available in this category than in the other categories analysed. As shown above, there are 35 brands in this category. With so many brands, buyers may limit the size of their repertoire. We do note however, that if the number of brands in a category was linearly related to the size of the repertoire, we would have expected a similar (if not so extreme) result in the other categories. The detergent and shampoo categories had 22 and 26 brands respectively, while the store choice was limited to 7 stores. Had there been a linear relationship between the number of brands and the size of the standardised repertoire, we would have seen a result similar to that for chocolate bars in the detergent and shampoo markets. It may be however, that this relationship is not linear and that there is a point at which markets contain enough brands and buyers begin to limit the size of their repertoire.

## **Future Research**

We have suggested some potential reasons for the results gained in this paper, but do believe that replication and extension of the analysis conducted here is necessary to better describe the relationship we see between standardised repertoire size and household spend in supermarkets.

Another area for future research is in the analysis of the impact of other covariates on repertoire size. Most data sets with covariates have more than one covariate recorded for each buyer, therefore using the technique discussed in this paper, there is a wide scope for further analysis.

## **Conclusion**

This paper has used DMD regression to analyse how standardized repertoire size changes as a function of weekly household spend. The results in this paper show that (i) for store choice, detergents and shampoos as spend increases standardized repertoire size remains approximately steady and (ii) for chocolate bars, as spend increases, standardized repertoire size decreases slightly. Thus in this market, heavier spenders are slightly more loyal to the brands in their repertoire than we would expect given the number of purchases they make from the category.

## **References**

- Banelis M, Rungie CM & Colombo R (2003). Empirical Verification of Two Expected Portfolio Size Formulae. Australian and New Zealand Marketing Academy Conference, Adelaide, Australia.
- Colombo R & Jiang W (2002). Portfolio Size of Brand Purchases in Dirichlet Markets. Australia and New Zealand Marketing Academy Conference.

Ehrenberg ASC (1988). Repeat-buying: Facts, Theory and Applications. London, Oxford University Press.

Ehrenberg ASC, Uncles MD & Goodhardt GG(2004). Understanding brand performance measures: Using Dirichlet benchmarks. *Journal of Business Research* 57(12): 1307-1325.

Goodhardt GJ, Ehrenberg ASC & Chatfield C (1984). The Dirichlet: A comprehensive model of buying behaviour. *Journal of the Royal Statistical Society* 147(part 5): 621-655.

Rungie CM & Laurent G (2004). *Research Note: Statistical Summary of the Generalized Dirichlet Model*. School of Marketing, University of South Australia.

Stern P, & Hammond K (2004). The relationship between customer loyalty and purchase incidence. *Marketing Letters* 15(1): 5-19.

### **Acknowledgement**

The authors wish to thank MarketingScan and GfK, and specially Laurent Battais, Raimund Wildner, and Gérard Hermet, for providing the data and for their assistance in this study.

**Melissa Banelis is a postgraduate student and Cam Rungie and Erica Riebe are Senior Lecturers at the University of South Australia. Lars Meyer-Waarden is an Assistant Professor at the University of Toulouse III Paul Sabatier**