

Can we fix the errors in self-reported buying frequencies?

Malgorzata Ludwiewska-Alluigi

MLing, PostGradCert(BusCom)

A thesis in fulfilment of the requirements for the degree of
Master of Business (Research)

Supervisors:

Associate Professor Jenni Romaniuk and Dr. Magda Nenycz-Thiel

University of South Australia
School of Marketing, Division of Business



Ehrenberg-Bass
Institute for Marketing Science

gosia.ludwiewska@marketingscience.info
University of South Australia, North Terrace, Adelaide, SA, 5000

August 2013

Abstract

Retrospective survey reports are prone to memory biases. Academia and industry raised questions about the validity and reliability of the measures obtained from this source. Errors of under- or over-reporting of behavioural events have been found. Nevertheless, academia and industry rely on surveys as their primary data source for cost and time efficiency reasons, or lack of other data sources (panel infrastructure). Consequently, the survey is the most viable alternative for longitudinal panel data. For marketers and academics, it is one of the most common methods of gathering marketing metrics to track consumer behaviour. Therefore, knowledge about how to improve the accuracy of marketing metrics in surveys is important.

The thesis focuses on past brand usage that is the major component of consumer related studies. Brand usage relates in the purchase behaviour context to two key measures (a) how many people buy from the category and brand and (b) how often they buy in the given time frame. These measures are expressed in terms of key brand performance metrics (BPMs), such as average purchase frequency (usage) and penetration (users). They are important independent, dependent and confounding variables in marketing studies. The accuracy of these metrics relies on the respondents' memory to accurately recall buying episodes.

The most commonly used questioning approach to gather usage data is by means of a frequency question such as, "In the last X months, how many times have you bought X?". However, the recall strategies that respondents use to formulate a frequency response can result in specific errors and give inaccurate estimates.

This thesis aims to improve the accuracy of usage frequency questions by (1) identifying the systematic errors in consumer responses and (2) testing eight questioning approaches to minimise the errors. The domain of the research is fast moving consumer goods (FMCG) in the UK. The period of recall tested was six months and the four categories were breakfast cereals, shampoo, tea and toothpaste. Analysis was conducted at category and brand level. A total of 36 category and 180 brand frequencies were analysed.

A matched sample quota approach was used to validate survey reports with panel records provided by TNS-Kantar Worldpanel ($N \approx 23,525$ households). A split sample approach was used to test the questioning approaches. The study has the following key findings:

1. Respondents over-report infrequent purchases and under-report frequent purchases.
2. Purchase reports of infrequently purchased categories have a higher error than purchase reports of frequently purchased categories.
3. Zero and light buyers are the main contributors to the overall over-reporting error in the data.
4. Medium and heavy buyers result as the most overstated because of the dominating error of the infrequent buyers.
5. The error is higher for smaller share brands than larger share brands.
6. Errors from light buyers are not due to forgetting to report past episodes, but rather telescoping them or inferring them from the overall purchase pattern.
7. Errors from heavy buyers are generated due to averaging or rounding the estimates.
8. Of the eight questioning approaches tested, the bounded recall approach, asking the respondent first about purchases in a longer time frame of 12 months and then about a shorter target period of 6 months, results as the most effective in questioning respondents about brand buying in the target time frame. This approach reduces the error across all brands on average by 40% in brands' average purchase frequencies and by 60% in brands' penetrations.

Therefore, to improve the accuracy of brand frequency data and the category penetration, the results of this thesis suggest a bounded recall questioning approach:

1. 'In the last 12 months, which of the ready to eat breakfast cereal brands have you bought? Please tick the appropriate answer.'
- List of brands
2. 'In the last 12 months, how many times have you bought the selected brands? Please write a number next to the brand.'
3. 'In the last 6 months, which of the ready to eat breakfast cereals have you bought? Please tick the appropriate answer.'
- List of brands
4. 'In the last 6 months, how many times have you bought the selected brands? Please write a number next to the brand.'

An incidence like question could be used instead of a frequency question to decrease the amount of effort on the respondent to retrieve purchase occasions in a longer time frame, 'In the last 12 months, have you ever bought breakfast cereals?'