Journal of Business Research 66 (2013) 224–232

Contents lists available at SciVerse ScienceDirect

Journal of Business Research

Investigating the accuracy of self-reports of brand usage behavior

Magda Nenycz-Thiel *, Virginia Beal [1], Gosia Ludwichowska [1], Jenni Romaniuk [1]

Ehrenberg-Bass Institute for Marketing Science, University of South Australia, GPO Box 2471 Adelaide SA 5001, Australia

ARTICLE INFO

Article history:
Received 1 May 2011
Received in revised form 1 December 2011
Accepted 1 December 2011
Available online 29 September 2012

Keywords:
Self-report surveys
Panel data
Purchase frequencies
Brand usage
Claimed data

ABSTRACT

This paper increases understanding of the accuracy of consumers' self-reports about using brands and categories. The researchers select television viewing as the category of usage firstly, due to the availability of robust panel data for validation of the claimed data (i.e. self-reported) and secondly, because watching television and purchasing fast moving consumer goods have similar underlying structures in consumer behavior (Ehrenberg, 1969; Goodhardt, Ehrenberg, & Collins, 1975). The results show that light users (viewers) are the main source of error at both brand (program) and category (total television viewing) levels. At brand level, the data shows underestimation of once-only events, which suggests that those who engage in behavior infrequently either forget that the event has occurred, or do not form a representation of the event in memory. At category level, light users tend to generalize their responses to reflect the regularity of the behavior, which manifests in fewer non-users in claimed data. Regardless of the measurement level, the main questioning challenge is getting less frequent users to accurately report an event occurring. The paper provides recommendations for brand researchers on how to minimize the errors caused by responses from light users, which will increase the accuracy of the usage metrics overall.

1. Introduction

Consumer surveys are a common data source in research by academia and industry. Steenkamp, de Jong, and Baumgartner (2010) report that 30% of all empirical articles in the *Journal of Marketing* and *Journal of Marketing Research* from 1996 to 2005 employ surveys. Surveys are also the most common method to gather data in commercial market research, with global survey research valued at $18.9 billion in 2010 (Casro: The Voice & Values of Research, 2011).

Despite its common use, survey data have limitations. The key limitation is the reliance on respondents to remember, and report on, their own behavior retrospectively. This reliance on respondent recall introduces the potential for respondents to engage in activities such as telescoping, projecting, and omission when giving responses (East & Uncles, 2008; Tourangeau, 2000). These factors lead to errors in consumers' recalled responses; these errors bring into question the degree to which survey responses reflect real world behavior.

These errors are especially problematic when measuring routine consumer behavior such as purchasing of categories or brands. Purchasing is an extremely common, and important, area of questioning in survey research. Wind and Lerner (1979 p. 46) state that "marketing research if it is to be of practical value, must ensure the reliability and validity of the most basic measures — the measure of past usage behavior." Research on past usage behavior collects consumer buying metrics. Two such metrics are how many consumers buy a category or a brand, referred to as penetration, and how many times they have bought, referred to as frequency.

Consumer usage metrics have several purposes. The first is to generate market estimates for brand buying and market share calculations. This need for estimates is particularly relevant in markets where scanner data or household panels that record actual purchases are not available, such as impulse categories or in emerging markets. The second use of the data is to segment survey responses into categories based on usage levels (light to heavy), which can be used to screen suitable respondents for questions about future loyalty or intent to recommend (such as in Reichheld, 2003), which are asked of a brand's customers only.

The third use by researchers is as a dependent variable in marketing studies. This need for a reliable dependent variable is of particular relevance to brand equity researchers given the considerable evidence that past consumer buying and usage behavior impact on consumer's awareness, perceptions and attitudes to the brand (Barnard & Ehrenberg, 1990; Barwise & Ehrenberg, 1987). To control for this bias, researchers need accurate measurement of consumer's usage of brands. Therefore, accurately collecting brand usage behavior is important for making sure academic and commercial studies use robust variables in analyses. A brand manager's performance assessments and researcher's modeling quality can rely on these survey responses being good representations of real consumer behavior.

* Corresponding author. Tel.: +61 8 8302 9172.
E-mail addresses: Magda.Nenycz-Thiel@MarketingScience.info (M. Nenycz-Thiel), Virginia.Beal@MarketingScience.info (V. Beal), Gosia.Ludwichowska@MarketingScience.info (G. Ludwichowska), Jenni.Romaniuk@MarketingScience.info (J. Romaniuk).
[1] Tel.: +61 8 8302 9172.

0148-2963/$ – see front matter
http://dx.doi.org/10.1016/j.jbusres.2012.07.016

Ideally, a researcher would capture consumer buying from longitudinal records of purchases collected in panel or scanner data. However, prohibitive costs (Lee, Hu, & Toh, 2000) lack of technology infrastructure, and various category-buying anomalies mean that this option is often not feasible. Researchers also often require other consumer based brand equity (CBBE) variables, such as awareness, perceptions and attitudes (Christodoulides & de Chernatony, 2010; Keller, 2003), alongside usage behaviors; however panel companies are reluctant to tax panelists further by surveying them to collect CBBE metrics. Therefore, many researchers and marketers are, and will continue to be, reliant on consumer's self-report responses to measure buying behavior.

Prior literature (e.g., Hu, Toh, & Lee, 1996; Lee et al., 2000; Ram & Hyung-Shik, 1990; Wind & Lerner, 1979) reports that claimed survey data are accurate in estimating rank order statistics, but the data tend to under or overestimate actual purchase frequencies. Nevertheless, as noted by Ram and Hyung-Shik (1990), there is little in-depth empirical evidence in the area, and most of the research simply compares the averages, ignoring the heterogeneity in respondents' behavior (Rust, Lemon, & Zeithaml, 2004), and assuming that the same type of error dominates everyone's responses. This study addresses the issue by examining and comparing the underlying distributions of responses from claimed self-reported survey data and panel data to identify where, how and with whom errors are concentrated.

The researchers conduct a pilot study in a supermarket confectionery category to test the questioning approaches. The main study takes place in a category where there is a comprehensive, and official, recording of consumer metrics — television viewing. The analysis involves comparing consumer responses obtained from an online survey with those reported by OzTAM, the television industry body officially tasked with providing television audience measurement in Australia. The testing is at overall television viewing (category) and program (brand) viewing levels. The next section discusses the inaccuracies in claimed data caused by memory biases.

2. Background

No data are perfect. Past research acknowledges problems with accuracy of survey data as well as recorded panel data (see e.g., Ehrenberg, 1960; Parfitt, 1967; Ram & Hyung-Shik, 1990; Sudman & Bradburn, 1974; Wright, 2002; Woodside & Willson, 2002). The main sources of error in panel data are forgetting to record, recording the information incorrectly, not recording (unknown) purchases of other members of the household or products consumed out of home or impulse purchases and wear out of the panel. However, despite the various sources of bias in recorded panel data, the consensus is that the panel data is a reasonably accurate representation of purchases (East & Uncles, 2008; Wright, Sharp, & Sharp, 2002). For ease of reference, this paper uses the term panel data to refer to longitudinal panels that record behavior from the same consumers over time as purchases occur, via scanning devices or other records of behavior such as purchase dockets.

Researchers are less confident about the accuracy of claimed data, because it is prone to memory biases. There are three major sources of inaccuracies in claimed responses: encoding and retrieval failure, memory decay, and telescoping (Cohen & Martin, 2008; Tourangeau, 2000). These processes lead to overestimation or underestimation in responses (e.g., Hu & Bruning, 1988; Lee et al., 2000; Neter & Waksberg, 1964).

2.1. Under-reporting in claimed data

One of the key reasons for under-reporting is memory failure. Respondents may not correctly recall whether they bought or consumed something, because they did not form a representation of the event in their memory (Tourangeau, 2000). However, even if encoding of an event happened in memory, the link may be so weak that one may fail to retrieve it because it did not enter long-term memory (Shiffrin & Atkinson, 1969). This encoding failure is more likely to occur when the event is frequent, quick or routine (Tourangeau, 2000). Events that are processed more deeply, as they are dramatic or unusual or involve elaborative processing are more effective in entering, and being retrieved from, long-term memory (Craik & Tulving, 1975). Behaviors such as buying a brand or watching television are usually frequent and non-vivid, which lowers the chance of forming rich representations in long-term memory, and increases the chance of insufficient encoding to allow for future retrieval (Blair & Burton, 1987). The nature of these activities leads to a general under-reporting of infrequent, but routine activities.

Even for items memory encodes, retrieval failure can still occur at any point in time. The easiest way to explain why retrieval failure happens is by realizing the amount of interlinked information that long-term memories store over time. The more links leading away from one central node, the lower the probability that the activation of any specific link will be strong enough for retrieval (Heil, Rösler, & Hennighausen, 1994). Therefore, accuracy will be lower as the behavior gets more specific, and there are more items for retrieval.

Finally, when asking about activities over a longer period, *memory decay* is likely to be a factor hindering the accuracy of responses (Tourangeau, 2000). The longer the passage of time from an event, the higher the probability that the event will be inaccessible from memory, which also causes under-reporting.

2.2. Over-reporting in claimed data

The main source of over-reporting in claimed data is *forward telescoping* (Cohen & Martin, 2008). Forward telescoping refers to perceiving distant events as more recent than they are, and therefore counting them even though they did not occur in the timeframe specified. Consequently, this process leads to over-reporting (Sudman & Bradburn, 1973).

2.3. Estimation based on generic information

Past research shows that frequent behavior can lead to use of rate-base heuristics in that responses reflect overall pattern of behavior (Hu, Toh, & Lee, 2000; Tourangeau, Rips, & Rasinski, 2000). For instance when asked about grocery shopping, one may think: *I shop every week, so I must buy X product once a week*, while when asked about watching television one may think: *I'm at home every night so I must watch television seven nights in a week*. Such reasoning leads to generalizing responses, which manifests in over-reporting of certain frequencies that are typical for particular behaviors in a given period.

The research begins with a pilot study, which compares survey and panel category statistics in the chocolate market in Australia (Stage 1). The researchers use the findings from the pilot study to refine questions in the main study conducted in the television viewing context (Study 2).

3. Stage 1: pilot study

3.1. Pilot study method

A pilot study in the chocolate category provides data for the first stage. The researchers chose this category for convenience purposes, with the opportunity to add these questions to a survey that was collecting data for another purpose. The survey data came from an online survey. The data collection took place in May 2010. Respondents were online panelists in Australia and the demographic characteristics of the datasets were representative of the Australian chocolate buying population — 51% female and 49% male, and dispersed across the following age groups; 8% 12–15 years, 17% 16–24 years, 19% 25–34 years, 22% 35–44 years,