Quantitative Marketing and Economics, 1, 223–244, 2003.

The Effectiveness of Demographic and Psychographic Variables for Explaining Brand and Product Category Use

GERALDINE FENNELL
Consultant, Bridgeport, CT
E-mail: fenl@onetel.com

GREG M. ALLENBY*
Ohio State University, 540A Fisher Hall, 2100 Neil Ave., Columbus, OH 43210
E-mail: allenby.1@osu.edu

SHA YANG
New York University, 44 West 4th Street, Suite 9–77, New York, NY 10012
E-mail: shayang@stern.nyu.edu

YANCY EDWARDS
University of Delaware, Business Administration, 222 MBNA America Hall, Newark, DE 19716
E-mail: edwardsy@lerner.udel.edu

Abstract. The predictive relationship of a large and comprehensive set of personal descriptors to aspects of product and brand use is examined. The descriptors comprise demographic and general psychographic variables frequently used in segmentation studies and studies of consumer purchase behavior. The evidence is overwhelming that the covariates are related to brand use in an identical way for all brands, indicating that they are not useful for predicting relative brand preference. The covariates are shown to be predictive of product use. Discussion of the explanatory content of the variables is offered.

Key words. brand preference, basis variables, segmentation

JEL Classification: C25, C53, D12, M31

1. Introduction

The search for useful explanatory variables for brand and product use constitutes a major research stream in marketing. Studies since the late 1960s have examined the effectiveness of variables to explain consumption rates and product use, examining demographics, psychographics, cultural values, personality and situational variables. Other studies demonstrated the association of these variables to being deal-prone, convenience oriented, and price and advertising sensitive. The bulk of this early

*Corresponding author.

Table 1. Studies relating demographic and psychographic variables to aspects of product category and brand use.

Studies using product category data:	
Consumption rates	Frank et al. (1967)
	Bass et al. (1968)
	Wildt and McCann (1980)
	Schaninger (1981)
Product category use	Peters (1970)
	Sparks and Tucker (1971)
	Frank et al. (1972)
	Belk (1974)
	Wells (1975)
	Kinnear and Taylor (1976)
	Henry (1976)
	Horton (1979)
Deal prone	Blattberg et al. (1978)
Convenience oriented	Anderson (1971)
Price and advertising sensitive	McCann (1974)
Studies using brand choice data:	
Price sensitivity and category expenditure	Allenby and Rossi (1991)
	Mittal (1994)
	Hoch et al. (1995)
	Bawa et al. (1997)
	Kalyanam and Putler (1997)
PurchasetTiming	Chintagunta and Haldar (1998)
Latent classes of consumers	Gupta and Chintagunta (1994)
	Bayus and Mehta (1995)

research, summarized in Table 1, focused on aspects of product category use, not on the use and preference for specific brands.

The availability of scanner panel data has made possible the study of brand choice behavior among product customers. However, researchers have not been successful in associating demographic and psychographic variables to brand-specific model parameters. While these variables have been related to aspects of price sensitivity and category expenditure, purchase timing, and latent classes of demand in finite mixture models, the evidence of a predictive relationship to the demand for specific brands is much weaker (see Table 1). For example, Rossi et al. (1996) show that the information content contained in demographics for explaining brand preference is very low, and Bucklin et al. (1995) report weak to no evidence of a predictive relationship between demographics and brand preference.

The existence of a predictive relationship for various aspects of product category use but not for brand preference use calls for examination. Product categories are groupings, or sets, that are composed of brands. Since brands are exchanged in commercial transactions, and not products categories, any measured predictive relationship between the variables and product category use must originate between the variables and the brand. Therefore, some predictive relationship between the variables and brand use must be present for the existence of a relationship at the abstracted category level. The absence of a predictive relationship conditional on category use does not necessarily imply absence of an unconditional relationship.

The purpose of this paper is to document the explanatory power of major classes of independent variables commonly used in studies of consumer preference, and to elucidate the difference in findings for product categories and brands. We present results of an analysis of more than 50 product categories from the Study of Media and Markets (SMM) database, examining the influence of five different sets of covariates (e.g., demographics, psychographics) commonly used in studies of consumer behavior. SMM is a survey of the American population, measuring media habits, product and brand use, and psychological descriptors of 20,000 individuals, which we describe in the next section. The results show a consistent pattern across product categories and sets of covariates, thereby providing generalizable results about the predictive usefulness of the variables explaining product and brand use.

We replicate previous findings concerning product use and volume, but also find that the variables examined predict brand use in an equal way for all brands. The variables therefore do not predict conditional (i.e., relative) brand preference. Demographics and psychographics are therefore not useful for predicting brand preference given product category use. We offer a conceptual discussion of why this is so and of the kind of variable that may be more effective.

The remainder of the paper is organized as follows. In Section 2, we briefly describe the SMM survey and, in Section 3, our selection of the product categories and brands for analysis. In Section 4, we describe the model used to assess the covariates' predictive relationship with product use and brand choice. Results are presented in Section 5, and discussion in Section 6.

2. The SMM data

The Simmons SMM is a comprehensive survey of American consumers, conducted annually. The study is based on a probability sample of adults aged 18+ living in the contiguous United States, stratified to ensure geographic and demographic representation. Data are made available on a semi-annual basis, based on rolling year samples of approximately 20,000 completed interviews. The data used in this analysis were obtained from all completed mail questionnaires administered between June 1996 and June 1997.

The survey comprises questions in three distinct categories: (1) media usage, including magazines, newspapers, television (cable and broadcast), outdoor advertising, and use of on-line services; (2) product and brand use for almost 2000 products and services, and (3) demographic/psychographic information about the respondents and their households. Since media use is outside the scope of the present paper, we limit our description of SMM information to the second and third categories.

Dependent variables. For each topic covered, the SMM asks a broadly similar set of questions, including three that are the focus of our interest here, i.e., product use, frequency of product use, and brand use. Using the product category as section heading, e.g., toothpaste, the first SMM question in each section establishes whether or not the respondent uses the product: "Do you yourself use it," with response alternatives: Yes, No, DK/NA (don't know, not applicable). Only respondents who answer, Yes, continue answering questions in that section. Such questions ask about decision-maker, types and forms of the product and of the container used. Then, SMM asks the respondent to indicate frequency of product use (in a form appropriate to the product category, e.g., for toothpaste, "times used each day," with response options: 4 or more, 3, 2, 1, less than 1). SMM then lists brands in the product category and asks the respondent, "for each brand you use," to check "most often," or "also use." SMM distinguishes four ways to consider the resulting data, i.e., Sole (only one brand checked), Primary (multiple brands checked in "most often" column), Secondary (brands checked in "also use" column), and All (sum of sole, primary, and secondary).

Independent variables. Tables 2 through 6 provide details of five groups of independent variables in the SMM survey, i.e., demographics, including socio-economic categories (Table 2), self-concept (Table 3), buying style (Table 4), attitudes, interests, and opinions, the latter divided into statements of personal views

Table 2. Demographic variables.*

1. Respondent age (years)
2. Attended grade school but not high school
3. Attended high school but did not graduate
4. Graduated from high school but attended less than 1 year of college
5. Attended less than three years of college
6. Graduated from college or more
7. Respondent is employed
8. Respondent is retired
9. Respondent is married
10. Respondent is male
11. Respondent is white
12. Respondent is black
13. Number of people in household
14. Respondent's employment income ($)
15. Number of children under 6 years of age (0, 1, 2, 3+)
16. Number of children age 6 to 17 (0, 1, 2, 3+)
17. Value of residence ($)
18. Residence ownership

*All variables coded as dummy (indicator) variables unless otherwise indicated in parentheses.

Table 3. Self-concept variables.*

1. Affectionate, passionate, loving, romantic
2. Amicable, amiable, affable, benevolent
3. Awkward, absent-minded, forgetful, careless
4. Brave, courageous, daring, adventuresome
5. Broad-minded, open-minded, liberal, tolerant
6. Creative, inventive, imaginative, artistic
7. Dominating, authoritarian, demanding, aggressive
8. Efficient, organized, diligent, thorough
9. Egocentric, vain, self-centered, narcissistic
10. Frank, straightforward, outspoken, candid
11. Funny, humorous, amusing, witty
12. Intelligent, smart, bright, well-informed
13. Kind, good-hearted, warm-hearted, sincere
14. Refined, gracious, sophisticated, dignified
15. Reserved, conservative, quiet, conventional
16. Self-assured, confident, self-sufficient, secure
17. Sociable, friendly, cheerful, likeable
18. Stubborn, hard-headed, headstrong, obstinate
19. Tense, nervous, high-strung, excitable
20. Trustworthy, competent, reliable, responsible

*All variables originally collected on a 1–5 scale (1 = agree a lot, 5 = disagree a lot) and converted to dummy variables (1 = agree, 0 = not sure or disagree).

(Table 5) and statements people have made (Table 6). Such variables provide comprehensive, general descriptions of the respondent, and span the range of questions typically considered in segmentation research.

Whenever possible, we employ dummy variable coding of the covariates to avoid making unnecessary assumptions about the scaling of the variables. All variables

Table 4. Buying style variables.*

1. *Environmentalists:* I buy products that use recycled paper in their packaging
2. *Economy Minded:* I shop around a lot to take advantage of specials or bargains
3. *Cautious:* I do not buy unknown brands merely to save money
4. *Impulsive:* When in the store, I often buy an item on the spur of the moment
5. *Experimenters:* I like to change brands often for the sake of variety and novelty
6. *Brand Loyal:* I always look for the brand name on the package
7. *Product Label Readers:* I usually read the information on product labels
8. *Conformists:* I prefer to buy things that my friends or neighbors would approve of
9. *Ad Believers:* In general, advertising presents a true picture of the products of well-known companies
10. *Style Conscious:* I try to keep abreast of changes in styles and fashions
11. *Planners:* I generally plan far ahead to buy expensive items such as automobiles
12. *Ecologists:* All products that pollute the environment should be banned
13. *Environmentalists:* I buy paper products (napkins, towels, toilet paper, etc.) that are recycled

*All variables originally collected on a 1–5 scale (1 = agree a lot, 5 = disagree a lot) and converted to dummy variables (1 = agree, 0 = not sure or disagree).

Table 5. Attitudes, interests, and opinions—statements of personal views.*

1. I am perfectly happy with my standard of living
2. I am a workaholic
3. I try to keep up with new developments in modern technology
4. I like spending most of my time at home with my family
5. I would be willing to volunteer time to an environmental organization
6. In a job, security is more important to me than money
7. It's important to me to attend religious services
8. I like to be well-organized and follow a routine
9. I'd rather have a boring job than no job at all
10. I am no good at saving money
11. I tend to spend money without thinking
12. I don't like the idea of being in debt
13. I am very good at managing money
14. I like to try out new food products
15. I really enjoy cooking
16. I would like to buy a home computer
17. I only give flowers when I can't think of any other present
18. I consider myself to be a conservative, evangelical Christian
19. I choose a car mainly on a basis of its looks
20. It's nearly always worth paying extra for quality goods
21. There is too much sponsorship of "the arts" and sporting events these days
22. It is important to me to look well-dressed
23. I am not too concerned about my appearance
24. I know I should exercise more than I do
25. I like to keep up with the latest fashions
26. Most of the time I'm trying to lose weight
27. I really enjoy shopping for clothes
28. The kitchen is the most important room in my home
29. I'm always looking for new ideas to improve my home
30. I always think of the calories in what I eat
31. I try to eat healthier food these days
32. I make sure I take regular exercise
33. I am concerned about health but tend not to do much about it
34. I decide what I want before I go shopping
35. I look for the lowest possible prices when I go shopping
36. I often enter contests and sweepstakes
37. I constantly watch the amount of calories I take in
38. I am dieting (not a physician ordered diet)
39. I recycle paper, glass, cans, etc. because it is required by law in my community
40. I enjoy watching religious television programs.

*All variables originally collect on a 1–5 scale (1 = agree a lot, 5 = disagree a lot) and converted to dummy variables (1 = agree, 0 = not sure or disagree).

were originally collected on a five point scale, with "1" indicating agree a lot, and "5" indicating disagree a lot, and were converted to dummy variables with "1" indicating agreement and "0" indicating not sure or disagree. We report the sensitivity of this recoding in the results section below. The number of covariates within each group ranges from 13 for buying style, to 40 for attitudes, interests, and opinions—statements of personal views.

Table 6. Attitudes, interests, and opinions—statements people have made.*

1.	I would rather travel in the United States than outside the country
2.	I often buy take-out meals to eat at home
3.	I expect advertising to be entertaining
4.	I'm very happy with my life as it is
5.	I like to behave as I please without worrying about other people's opinions
6.	Money is the best measure of success
7.	I don't like to stand out from the crowd
8.	I prefer to spend a quiet evening at home than to go out
9.	I like other people to think I'm a financial success
10.	How I spend my time is more important than how much money I make
11.	It is more important to do your duty than to live for your own enjoyment
12.	I'm willing to sacrifice time with my family in order to get ahead
13.	I'm careful with my money
14.	I consider myself an intellectual
15.	Just as the Bible says, the world literally was created in six days
16.	I like to be outrageous
17.	I feel very alone in the world
18.	I worry a lot about myself
19.	Women are more suited to running their homes than holding public office
20.	I like to do things that are unconventional
21.	There's little I can do to change my life
22.	I rely on newspapers to keep in touch
23.	I entertain more often on the spur of the moment
24.	If I win the lottery, I would never work again
25.	I hate doing any form of housework
26.	Men's fashions are more exciting now-a-days
27.	I am good at fixing mechanical things
28.	I do some sport/exercise at least once a week
29.	I enjoy eating foreign foods
30.	I always use money-off coupons

*Variables originally collected in binary (agree/disagree) format. Dummy variable coding used for all variables.

3. SMM data selected for analysis

Our analysis concerns 52 product categories of non-durable packaged goods, about which SMM questions both males and females. We selected frequently purchased categories, favoring products that have been previously studied in the academic literature, or that we suspected might have relatively strong predictive relationship to the covariates. We included all brands in the analysis that have a use share of greater than 1% of the total sample. That is, brands were included in the analysis if more than 200 of the 20,000 respondents reported using it. This criterion resulted in a relatively minor reduction in the number of brands analyzed. Furthermore, we employed a case-wise deletion of respondents from the analysis whenever we encountered missing observations. While this resulted in excluding many respondents, there still remained sufficient information for our analysis. For example, many respondents did not provide information on household income, resulting in fewer respondents available for analysis for the demographic compared with other, analyzes of covariate relationships. Moreover, as a DK/NA response to the dependent variable is uninformative with regard to the predictive relationship, the casewise deletion procedure does not result in any distortion in its measurement (see Green, 1993). Table 7 reports the product categories, number of brands (in database and analysis), and the number of respondents for each set of covariates.

Throughout, all respondents were retained in the analysis, whether product category users or not. The reason for this approach is that the deletion of individuals who are not product users would result in a censored sample based on the dependent variable. For example, since consumers can be product users only if they use one or more brands, there exists a logical relationship between product and brand use. Restricting an analysis of brand usage to only those consumers who are product users results in an implicit censoring based on the dependent variable. This form of censoring would lead to inconsistent parameter estimates and distorted statistical evidence about the covariate relationships. The incidence of product use for the categories under study is provided in Table 8.

4. Analytic approach

The goal of our analysis is to determine the strength of predictive relationship, if any, between the five sets of covariates described in Tables 2 through 6, and various aspects of product and brand use. As noted in Section 2, all respondents are asked to answer the question about product use, so for each of the 52 categories, we used the binomial responses in yes/no form as the dependent variable. Only respondents who answer "yes" to product use are asked to respond to the within-product category questions, including frequency of product use and brand use. Respondents are presented with a listing of brands and given the opportunity, for each, to check "most often" or "also use." For the present analysis, we combine responses to these two questions to obtain a single binomial measure of "any use of brand *i*," i.e., sole, primary, and/or secondary. As noted above, non-category users were included in the brand and frequency analysis by imputing values of the dependent variables. In the case of "any use of brand *i*," for non-category users the code for "No" is entered in the dataset for each brand included in the analysis. In the case of frequency of product use, we recoded the data to reflect zero frequency for respondents who were not category users.

Our approach for establishing the existence and strength of the predictive relationships of interest is to employ a latent variable model. We assume that a particular respondent will report use of a brand if the value they receive from it (net of price) is sufficiently high. We further assume that the distribution of values for a

Table 7. Description of product categories.

Product category	Brands in database	Brands analyzed	Respondents available for analysis: Demographics	Self-concept	Buying style	Attitudes	Opinions
Charcoal, for barbecues	5	5	8447	12,780	14,394	10,900	11,877
Tuna, canned	8	8	8691	13,049	14,699	11,085	12,111
Potatoes, packaged instant	8	8	8338	12,608	14,156	10,705	11,695
Frosting, store bought	9	9	8513	12,884	14,504	10,955	11,978
Jams, jellies and preserves	9	9	8742	13,089	14,748	11,102	12,136
Mustard	9	9	8675	12,988	14,621	11,021	12,052
Spaghetti, dry packaged	10	9	8570	12,889	14,521	10,960	12,003
Eye drops and eye wash, nonprescription	10	7	8462	12,622	14,202	10,762	11,732
Kitchen wrap, plastic-type	10	10	8598	12,897	14,537	10,940	11,965
Cola drinks, diet or sugar free, carb.	11	9	8807	13,025	12,113	11,069	13,025
Drain cleaner	12	9	8505	12,803	14,450	10,914	11,938
Floor wax or polish	12	10	8347	12,657	14,243	10,769	11,767
Pickles	12	11	8692	12,981	14,639	11,039	12,068
Snack cakes, ready to eat	12	12	8530	12,855	14,497	10,940	11,954
Bacon	13	12	8601	12,953	14,615	11,025	12,047
Bug spray	13	12	8310	12,560	14,120	10,680	11,661
Cough syrup, nonprescription	13	11	8758	13,002	14,652	11,059	12,072
Batteries, household	14	11	8895	13,152	14,785	11,155	12,170
Dishwashing liquid, not for automatics	14	14	8697	13,026	14,694	11,041	12,101
Orange juice, frozen	14	14	8563	12,908	14,523	10,959	11,985
Scouring pads and scouring sponges	14	13	8619	12,931	14,553	10,980	12,013
Brownie mixes, dry	15	14	8509	12,820	14,450	10,915	11,936
Pancake table Syrup	15	14	8675	12,992	14,643	11,022	12,055
Furniture polish	17	15	8557	12,831	14,452	10,890	11,906
Garbage bags, trash can liners plastic	17	16	8712	13,037	14,667	11,055	12,081
Bleach	18	14	8686	12,992	14,647	11,020	12,068
Tea, regular in bags, packages	18	18	8676	12,981	14,638	11,042	12,061
Cake mix, dry	19	18	8488	12,785	14,378	10,838	11,858
Vegetables, frozen	21	20	8708	13,046	14,709	11,093	12,131
Candy, hard roll	21	21	8727	12,968	14,596	11,012	12,044
Scotch whisky	21	9	8505	12,782	14,374	10,916	11,898
Facial tissue	22	17	8703	13,019	14,664	11,027	12,071
Toilet bowl cleaners, in tank	22	18	8516	12,848	14,472	10,921	11,923
Mouthwash/dental rinse	23	22	8853	13,086	14,717	11,107	12,135
TV dinners, frozen complete	23	23	8388	12,636	14,233	10,749	11,770
Toilet paper	24	21	8735	13,045	14,727	11,062	12,089
Laxatives, nonprescription	25	7	8521	12,780	14,376	10,888	11,874
Potato chips	25	23	8661	12,993	14,638	11,027	12,044
Sparkling water/seltzer/natural sodas	25	19	8471	12,667	14,196	10,775	11,754
Margarine/margarine spread	26	26	8503	12,801	14,413	10,879	11,900
Paper towels	26	23	8753	13,063	14,720	11,088	12,111
Salad dressing, prepared	31	31	8747	13,092	14,750	11,072	12,125
Toothpaste	32	32	9043	13,252	14,937	11,230	12,284
Frozen casseroles or rntrees	36	34	8327	12,557	14,108	10,681	11,664
Yogurt, not bought frozen	37	24	8698	12,931	14,538	11,002	11,998
Coffee, ground/whole bean	38	28	8525	12,819	14,462	10,927	11,943
Beer, regular domestic	38	18	8591	12,847	14,464	10,981	11,946
Toilet soap, bar	39	36	8771	13,093	14,763	11,110	12,149
After-shave lotion and cologne	46	24	8786	12,989	14,616	11,063	12,048
Cookies, ready to eat	47	46	8705	13,046	14,713	11,089	12,112
Headache and pain relievers, nonprescription	47	27	9027	13,250	14,935	11,230	12,277
Candy bars and packs, full size	74	74	8901	13,119	14,766	11,134	12,145

Table 8. Incidence of product use: 52 product categories.

Product	Proportion who use the product
Toothpaste	0.9605
Toilet paper	0.9493
Paper towels	0.9275
Toilet soap, bar	0.9169
Headache and pain relievers, nonprescription	0.8911
Mustard	0.8792
Batteries, household	0.8754
Garbage bags, trash can liners plastic	0.8747
Dishwashing liquid, not for automatics	0.8644
Jams, jellies and preserves	0.8522
Salad dressing, prepared	0.8504
Tuna, canned	0.8396
Facial tissue	0.8322
Scouring pads and scouring sponges	0.8191
Bleach	0.8140
Kitchen wrap, plastic-type	0.8095
Pickles	0.8083
Pancake table syrup	0.8054
Potato chips	0.7954
Cookies, ready to eat	0.7816
Vegetables, frozen	0.7580
Furniture polish	0.7551
Tea, regular in bags, packages	0.7124
Bacon	0.7050
Margarine/margarine spread	0.7049
Candy bars and packs, full size	0.6762
Mouthwash/dental rinse	0.6386
Cake mix, dry	0.6346
Candy, hard roll	0.6049
Coffee, ground/whole bean	0.5740
Spaghetti, dry packaged	0.5394
Frosting, store bought	0.5377
Cough syrup, nonprescription	0.5287
Bug spray	0.5257
Orange juice, frozen	0.5012
Brownie mixes, dry	0.4893
Drain cleaner	0.4863
Snack cakes, ready to eat	0.4792
Yogurt, not bought frozen	0.4723
TV dinners, frozen complete	0.4546
Cola drinks, diet or sugar free, carbonated	0.4533
Toilet bowl cleaners, in tank	0.4526
After-shave lotion and cologne	0.4318
Frozen casseroles or entrees	0.3808
Potatoes, packaged instant	0.3725
Eye drops and eye wash, nonprescription	0.3417
Charcoal, for barbecues	0.3364
Beer, regular domestic	0.3261
Floor wax or polish	0.2914
Sparkling water/seltzer/natural Sodas	0.2704
Laxatives, nonprescription	0.1632
Scotch whisky	0.1270

brand is related to the covariates through a linear model:

$$\begin{aligned}&\Pr(\text{any use of brand } i) = \Pr(\mathbf{x}'\boldsymbol{\beta}_i + \varepsilon_i > \text{threshold}_i),\\ &\varepsilon_i \sim \mathrm{EV}(0,1),\end{aligned} \tag{1}$$

where $\mathbf{x}'$ denotes the vector of covariates, $\boldsymbol{\beta}_i$ is a vector of coefficients for brand i, and ε_i is an error term that captures the distribution of residual values. Equation (1) is equivalent to a standard discrete choice model

$$\Pr(\text{any use of brand} i) = \Pr(\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i + \varepsilon_i > 0), \tag{2}$$

where β_{0i}, the model intercept for brand i, is a combination of the threshold value in (1) plus the intercept from the regression function $\mathbf{x}'\boldsymbol{\beta}_i$. For ε_i distributed extreme value, the brand choice probabilities take on a logit form:

$$\Pr(\text{any use of brand } i) = \frac{\exp(\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i)}{1 + \exp(\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i)}. \tag{3}$$

Since there are multiple brands in any product category, the data from a respondent can be regarded as a multivariate binomial vector of zeros and ones. More than one element of the vector can take on a value of one (indicating that more than one brand is used). Since a product is "used" whenever at least one of the brands are used, we have:

$$\begin{aligned}\Pr(\text{use of product}) &= \Pr(\text{at least one brand used})\\ &= \Pr(\max\{\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i + \varepsilon_i\} > 0)\\ &= \Pr(\mathbf{x}'\boldsymbol{\beta} + \max\{\beta_{0i} + \varepsilon_i\} > 0),\end{aligned} \tag{4}$$

where the equality in the third line occurs only if the coefficients β are equal for all brands. For ε_i distributed extreme value (i.e., $\varepsilon_i \sim$ EV(0,1)), then it can be shown that the distribution function of the maximum of $\{\beta_{0i} + \varepsilon_i\}$, equal to the product of the individual distribution functions (see Mood et al., 1974, p. 182), is also extreme value but with a new location parameter:

$$\max\{\beta_{0i} + \varepsilon_i\} \sim \mathrm{EV}\left(\ln \sum_i \exp(\beta_{0i}), 1\right). \tag{5}$$

This leads to a model for product use that is also of logit form but with an intercept,

β_0^*, that is related to the brand intercepts in equation (3) in a non-linear manner:

$$\Pr(\text{use of product}) = \frac{\exp(\beta_0^* + \mathbf{x}'\boldsymbol{\beta})}{1 + \exp(\beta_0^* + \mathbf{x}'\boldsymbol{\beta})}. \tag{6}$$

In other words, if we find that the slope coefficients $\boldsymbol{\beta}_i$ are the same across the i brands, then it makes sense to estimate a product use model of similar form. If the slope coefficients are instead different, then any failure to detect an association with equation (6) can be explained in part by the need for a more complicated model.

An advantage of using independent extreme value errors is that it leads to probability expressions that are easy to evaluate. This ease of evaluation allows us to include many brands per category in our analysis. As reported in Table 6, the median number of brands per category is 18 and the maximum number of brands is 74. The use of a correlated error structure, such as a multivariate normal, is difficult to implement because of the need to evaluate the model likelihood for testing purposes. Moreover, the assumption of a dependent error structure will only affect the standard errors of the estimated coefficients in equation (2), while the assumption of a specific error distribution will only serve to rescale the coefficients by a specific constant.

To understand the influence of our modeling assumptions, consider the effect of assuming a multivariate normal distribution for the errors in equation (2). Estimates from a model that incorrectly assumes independent normal errors are still statistically consistent (i.e., in large samples the estimates converge almost surely to the true values), but are no longer efficient (i.e., do not have the smallest standard errors). This property is also present in multivariate regression models—i.e., ignoring dependencies in the error structure still yields consistent estimates of regression coefficients (see Johnson and Wichern, 1998, Chapter 7). Furthermore, as discussed in Green (1993), the assumption of extreme value versus normal errors results in a simple rescaling of the estimated coefficient vector by $\pi/\sqrt{3}$ because of difference in the variances of the two distributions. We report on a sensitivity analysis of the results to the assumed error distribution below.

One could consider using the multinomial "most preferred" data instead of the multivariate binomial "any use of brand i" data as a dependent variable. The use of either variable provides information about the latent values of the brands. The "most preferred" data reveal the one brand with highest value (i.e., $\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i$ is max for all i), while "any use of brand i" reveals all brands with value above the threshold (i.e., $\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i > 0$). For analyses involving many brands, the "any use of brand i" data therefore contain more information about the latent values for the entire set of brands.

Finally, we investigate the predictive relationship between the covariates and frequency of use through a censored regression (tobit) model:

$$\begin{aligned} &y = 0 \quad \text{if } y^* \leq 0, \\ &y = y^* \quad \text{if } y^* > 0, \quad \text{where } y^* = \mathbf{x}'\gamma + \varepsilon \text{ and } \varepsilon \sim \text{Normal}(0, \sigma^2) \end{aligned} \tag{7}$$

where γ denotes the vector of coefficients and y is the frequency data with imputed zeros for non- product use. We assume that if a respondent is a non-user, then utility for the product is below some minimum threshold. More formally, the rate of consumption of a product is dependent on its utility ($\beta_0^* + \mathbf{x}'\boldsymbol{\beta}$) and the utility of all other goods (see Hanneman, 1984; Arora et al., 1998). Since we do not observe expenditures for the various products, we cannot specify a formal model of consumption frequency that is related to brand use (equation (3)) or product use (equation (7)). Instead, we investigate the relationship between the covariates and frequency through a simple predictive model that accounts for censoring.

The predictive relationship between the covariates and brand preference is investigated by estimating a series of constrained and unconstrained models. We estimate three models with the multivariate binomial brand use data (see equation (3)): (1) a model without covariates ($\boldsymbol{\beta}_i = 0$ for all i), (2) an unconstrained model ($\boldsymbol{\beta}_i \neq 0$ for all i), and (3) a model where the slope coefficients for all brands are constrained to be equal ($\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all i). We fit the models using the method of maximum likelihood, and test the restrictions by calculating the Bayesian Information Criterion (BIC) proposed by Schwarz (1978). This criterion provides a large-sample estimate of the marginal (or predictive) density of the data used in Bayesian hypothesis testing. Bayesian methods of testing are often employed when analyzing large datasets because they help avoid the problem of almost sure rejection of sharp hypotheses (see later) in large samples (see Allenby, 1990). In other words, this approach helps avoid committing a Type I error where a rejection of a null hypothesis is incorrectly made. In addition, Bayesian approaches to testing allow for evidence against the null hypothesis because the test statistic does not condition on the null being true. For the product use (equation (6)) and frequency (equation (7)) data, we fitted models with and without the covariates.

An interesting test of the covariates involves whether the coefficients are equal for all brands within a product category. If we find support of the model where $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all brands, then this implies that the covariates have an equal association to all of the brands in a product category. That is, the effects of the covariates would cancel in a model of relative brand preference (i.e., $\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i$ is max for any i), although the covariates may be useful to predict "any use of brand i" (i.e., $\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i > 0$).

It should be noted that our analysis of the brand use data is different from that commonly employed in most practitioner studies. Practitioner studies typically examine the relationship between covariates and the brand "most used," not the "any use" brand variable. The reason is that the former variable is often directly interpreted as a measure of preference, and under this assumption, the relationship to covariates can be assessed through a cross-tabulation of the data instead of a latent variable model. However, this approach to analyzing the data quickly becomes awkward when studying many covariates because the dimension of the resulting table is of order 2^k, where k is the number of dichotomous covariates. The covariate relationship is more efficiently assessed through a latent linear model.

Finally, practitioner search for correlates of brand use often involves a multi-step procedure to group respondents into clusters, and then show brand use behavior for each cluster. There are several drawbacks to this approach. First, any potential relationship between the variables used to create the clusters and brand preference must be contained in the univariate distance measure used to create the clusters. Cluster analysis employs a one-dimension measure that summarizes differences among respondents on a large battery of variables. Our modeling approach allows for a unique relationship between each variable and each brand, avoiding use of a mediating distance variable. Second, the multi-step approach used by practitioners leads to over-confident inference unless uncertainty associated with cluster formation is taken into account. It is often the case that many different cluster solutions (e.g., 4, 5, 6 clusters) fit the data with similar levels of explained variance, and statistical differences in brand use profiles should account for the possibility of multiple cluster solutions. Moreover, there exists much evidence in marketing that the distribution of heterogeneity of brand preference is not well represented by a finite set of point masses (Allenby and Rossi, 1999), which is assumed to exist in a cluster analysis. Our approach to analyzing the relationship between the covariates and aspects of brand use is therefore more flexible and precise than other approaches.

5. Results

Results are reported in summary form because of the large number of models that required estimation. Instead of reporting results for all of the models fitted, we report results across products and make available the details of our results upon request. Approximately 50% of the coefficients were estimated to be statistically significant at the 0.05 level, and the algebraic signs of the coefficients are plausible. For example, regarding the dependent variable "any use of brand i", age is negatively related in the frosting, potato chips and candybars categories, and gender in after-shave. Individuals who report that they constantly watch the amount of calories they take in tend not to use beer, bacon, and brownie mixes.

Table 9 reports the number of products for which each set of covariates is predictively related to product use, any use of brand i, product use frequency, and relative brand choice. The inclusion of demographic variables, for example, resulted in an improvement in the BIC statistics for the product use model (equation (6)) in 44 out of 52 categories. Similarly, the demographic variables resulted in an improved BIC statistic for the any use of brand i models (equation (3)) in 48 out of 52 categories. Recall that the BIC is a predictive fit statistic that penalizes the in-sample model fit by the number of model parameters. Since the number of parameters in the

Table 9. Evidence of predictive relationships.[1]

	Demographics	Self-concept	Buying style	Attitudes	Opinions
Product use	44/52	19/52	22/52	28/52	11/52
Any use of brand[2]	48/52	29/52	36/52	33/52	17/52
Frequency of use	42/52	3/52	5/52	5/52	2/52
Relative brand preference	0/52	0/52	0/52	0/52	0/52

[1]Predictive evidence assessed with the BIC criterion (log likelihood – # parameters/2 × ln(sample size)). Reported is the number of product categories out of 52 for which the set of covariates is predictively useful.
[2]"Any use of brand" means that the respondent has checked the brand as the sole, primary or secondary brand used, as defined in the SMM (p. 230 of this article).

product use and brand use models differ, the evidence of a predictive relationship can also differ, even though a necessary condition for product use is the use of one of the brands.

The evidence of a predictive relationship between frequency of product use and the five sets of covariates is less strong. Demographic variables, for example, are predictively useful in 42 out of 52 categories. The demographic association is primarily related to variables such as family size and household income. None of the other sets of covariates is very useful in predicting frequency of use, with predictively useful results detected in just two to five of the 52 product categories.

An interesting result of the analysis is the relationship between the covariates and *relative* brand preference (i.e., $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all brands i). A predictively useful association is found in *none* of the 52 product categories for each set of covariates. This result is obtained by comparing the BIC of the constrained multivariate model with isolated binomial models. While the low predictive power of such general personal descriptors with brand-relevant behaviors has been previously reported (see Wedel and Kamakura, 2000, pp. 13–14), the weight of evidence, here, favoring independence is striking. It should be noted that this evidence is not based on a failure to detect a relationship between the covariates and brand use. Instead, the evidence is based on a sharp hypothesis that the covariate effects are equivalent for all brands in the product category ($\boldsymbol{\beta}_i = \boldsymbol{\beta}$ for all i). We find evidence of a predictive association between the covariates and any use of brand i, but also find that this association is the same for all brands. This implies that the five sets of covariates are not useful for predicting which brand consumers use, only that some brand is used, i.e., some use of the product class.

Toothpaste coefficient estimates for the buying style variables are reported in Table 10. There are 32 brands of toothpaste in our analysis and 13 buying style variables as reported in Table 4. Each row of Table 10 corresponds to an analysis with a different brand as the dependent variable. The columns correspond to one of the 13 buying style variables, plus an intercept, used as independent variables in the analysis. The cells of the table display the coefficient estimates, which are in bold if the estimate is more than two standard errors from zero. The bottom row of the table reports estimates of the constrained model. The common slope coefficients are

Table 10. Toothpaste coefficient estimates for buying style variables.[1]

	Int.	buy1	buy2	buy3	buy4	buy5	buy6	buy7	buy8	buy9	buy10	buy11	buy12	buy13
Brand1	**−1.13**	0.03	−0.02	0.01	0.04	−0.08	0.07	0.04	−0.04	−0.01	0.08	0.07	0.02	0.02
Brand2	**−1.76**	−0.02	0.08	−0.06	−0.04	**0.12**	−0.05	−0.01	−0.02	**0.16**	0.08	0.08	0.05	0.01
Brand3	**−2.18**	0.05	0.20	0.02	**0.15**	0.00	−0.05	0.01	**−0.25**	−0.02	**0.21**	−0.04	−0.06	0.05
Brand4	**−1.63**	−0.06	−0.02	−0.04	−0.08	0.03	0.03	**−0.14**	0.03	0.07	**−0.12**	−0.04	−0.08	−0.02
Brand5	**−2.36**	−0.05	**0.18**	−0.11	**0.13**	**0.25**	−0.09	0.02	**−0.16**	0.00	0.05	**−0.13**	0.10	0.05
Brand6	**−2.20**	**−0.17**	0.13	−0.09	−0.09	0.10	**0.14**	−0.01	0.15	0.07	−0.04	−0.06	−0.01	0.01
Brand7	**−2.22**	−0.19	0.08	−0.03	−0.10	**0.20**	−0.07	−0.04	0.04	**0.17**	0.02	−0.04	0.09	0.11
Brand8	**−2.62**	−0.08	0.08	−0.07	−0.11	**0.27**	−0.03	−0.02	0.06	**0.20**	0.05	−0.05	−0.03	**0.15**
Brand9	**−2.79**	**0.18**	0.07	−0.06	0.10	**0.21**	−0.13	0.08	**−0.29**	−0.07	0.07	−0.03	0.05	−0.04
Brand10	**−2.85**	−0.13	**0.27**	**−0.23**	0.06	**0.19**	−0.06	−0.08	0.08	−0.03	−0.09	−0.09	**−0.17**	0.09
Brand11	**−3.25**	−0.17	0.15	−0.10	0.03	0.12	−0.12	**0.29**	0.01	0.04	−0.08	0.06	0.00	0.13
Brand12	**−3.58**	0.13	**0.22**	0.09	0.00	0.07	0.02	**0.32**	0.10	−0.16	0.11	−0.02	0.10	−0.08
Brand13	**−3.20**	−0.08	**0.25**	−0.05	−0.06	0.19	0.03	−0.10	−0.09	0.14	0.15	−0.18	−0.03	−0.03
Brand14	**−3.22**	−0.19	0.14	−0.11	0.11	**0.25**	**−0.23**	−0.02	−0.06	0.09	0.14	−0.15	0.12	0.16
Brand15	**−3.22**	−0.10	**0.31**	**−0.26**	−0.11	**0.38**	−0.13	0.06	−0.08	0.12	0.05	−0.12	0.10	−0.05
Brand16	**−3.34**	−0.12	0.19	**−0.20**	0.11	**0.34**	−0.06	0.06	−0.14	0.26	0.16	**−0.27**	0.10	−0.02
Brand17	**−3.50**	0.04	0.20	−0.05	0.01	0.17	−0.17	0.11	0.00	−0.03	0.11	−0.10	0.06	0.10
Brand18	**−3.40**	−0.07	0.18	−0.12	−0.11	**0.38**	−0.01	0.08	−0.14	−0.07	**0.20**	−0.06	0.05	−0.03
Brand19	**−3.62**	**−0.39**	**0.41**	−0.16	−0.12	**0.23**	−0.12	0.11	−0.24	**0.31**	0.09	0.01	0.24	0.14
Brand20	**−3.24**	−0.21	**0.23**	**−0.30**	**−0.21**	0.08	0.00	−0.13	0.03	−0.06	0.16	0.00	−0.12	0.12
Brand21	**−3.71**	−0.02	0.20	0.07	0.02	**0.32**	−0.17	**0.26**	0.01	−0.03	0.04	−0.15	−0.03	0.00
Brand22	**−3.91**	−0.02	−0.10	−0.11	0.10	**0.46**	0.01	−0.02	−0.01	−0.14	**0.34**	0.04	−0.07	0.12
Brand23	**−4.14**	−0.03	**0.54**	−0.16	0.04	0.42	−0.18	−0.02	**0.36**	0.04	0.16	−0.04	−0.15	0.12
Brand24	**−3.79**	0.00	**0.57**	**−0.42**	−0.02	0.13	**−0.45**	**0.29**	−0.15	−0.08	0.02	−0.08	−0.12	0.04
Brand25	**−4.09**	0.03	0.18	−0.07	−0.04	0.14	−0.16	0.07	0.04	0.21	0.10	0.07	0.10	0.06
Brand26	**−4.31**	**0.32**	−0.13	−0.17	−0.03	**0.32**	−0.26	0.60	−0.24	−0.14	−0.03	−0.11	0.25	0.04
Brand27	**−4.14**	−0.09	0.28	**−0.29**	−0.11	**0.38**	−0.10	−0.03	0.11	0.25	0.16	0.11	−0.10	−0.15
Brand28	**−4.37**	0.05	**0.35**	−0.02	0.09	**0.29**	−0.06	−0.07	−0.28	−0.07	0.19	−0.03	−0.03	−0.02
Brand29	**−4.47**	−0.21	0.21	−0.27	0.19	**0.51**	−0.19	−0.24	0.06	0.10	0.19	−0.01	0.18	0.30
Brand30	**−4.58**	−0.11	0.32	0.04	−0.03	0.32	−0.26	−0.24	0.04	**0.48**	0.27	−0.10	0.02	0.19
Brand31	**−4.66**	0.02	0.19	0.02	0.00	0.16	−0.04	0.23	0.00	0.25	0.05	**−0.16**	−0.36	0.09
Brand32	**−4.58**	0.13	**0.42**	−0.12	−0.01	0.25	−0.06	−0.10	0.17	0.31	0.14	−0.28	−0.07	−0.27
constrained	*	**−0.04**	**0.13**	**−0.08**	0.00	**0.16**	**−0.05**	0.02	−0.04	**0.06**	**0.07**	**−0.04**	0.01	0.04

[1]Entries in bold are greater than two standard errors from zero. The constrained model allows for separate Brand intercepts (not reported) and common slope coefficients.

reported for the constrained model, but the unique brand intercepts are not reported. The fifth buying variable, "I like to change brands often for the sake of variety and novelty" is positively associated with nearly all brands, with the coefficient statistically significant for more than half the brands. The coefficient for this variable is significantly greater than zero in the constrained model. Similarly, the third buying variable "I do not buy unknown brands merely to save money" is negatively related to nearly all brands, and statistically significant in six of the 32 brand relationships. The estimate of the constrained coefficient is negative and also significantly different from zero. The coefficients reported in Table 10 are typical of our empirical findings—the coefficient estimates are generally similar across brands.

Table 11 reports the percentage improvement in model fit for each set of covariates across the 52 product categories analyzed. The model fit for the product

Table 11. Relative improvement in model fit due to covariates.

	Demographics	Self-concept	Buying style	Attitudes	Opinions
Product use	1.024	1.020	1.006	1.016	1.006
Any use of brand	1.001	1.001	1.000	1.001	1.007
Frequency of use	1.028	1.013	1.004	1.009	1.005

Model fit for product and brand use measured in terms of average hit probabilities (i.e., the average probability of observing the binary data). Model fit for frequency of use measured in terms of mean absolute deviation of predicted versus observed purchase frequency. Reported is the ratio of the fit statistics for the model with covariates to the model without covariates, averaged over the 52 product categories.

use and any use of brand i analysis is measured in terms of the average fitted probabilities, defined as $\exp(\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i)/(1 + \exp(\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i))$ if the brand or product is used, and $1/(1 + \exp(\beta_{0i} + \mathbf{x}'\boldsymbol{\beta}_i))$ if the brand/product is not used. Model fit for the analysis of the frequency data is measured in terms of the mean absolute deviation between the predicted and actual frequency data. We report the relative improvement in the fit statistic associated with use of the covariates. In general, we find that the effect sizes of the covariates are an order of magnitude larger for the product use as opposed to the any use of brand i data, and that the results for the frequency data are similar to the product use data. The degradation in fit for the any use of brand i data is primarily due to the fact that many of the brands in the analysis have low market shares, making it easy to obtain a good average fitted probability by simply having a small intercept coefficient. The covariates therefore often do not substantially improve the fit statistic beyond a simple, intercept-only, model.

To obtain a better understanding of the meaning of the model fit statistic, Figure 1 displays the improvement in the fitted product use probabilities due to the

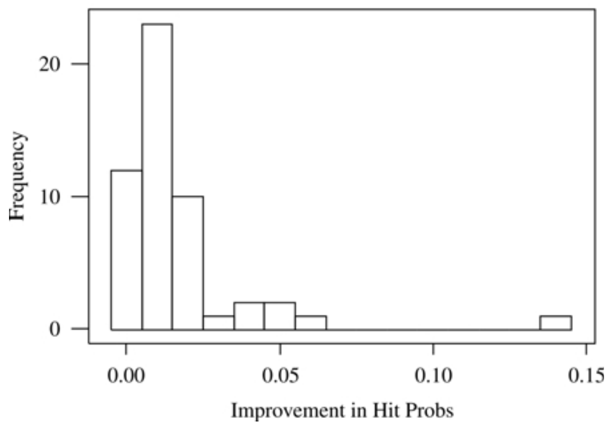

Figure 1. The effect of demograhic variables across products.

demographic variables. Plotted is the increase in model fit, measured in terms of the hit probabilities for each of the 52 product categories when the demographic variables are entered. Most of the product categories show relatively little improvement in fit, while in others the improvement in the fitted probabilities is greater than 0.02. For example, the improvement in model fit due to the covariates is of the order 0.05 for beer, frosting, brownie mix, and snack cakes, and is equal to 0.14 for aftershave with a respondent's sex (male) being highly predictive of product use. In other categories, such as toilet paper and batteries, the improvement is of the order 0.001. While the determination of whether these improvements are economically meaningful depends on factors such as size of potential market and the cost of advertising, we speculate that improvement in fitted probabilities of order $0.0x$ are often economically useful while improvements less than 0.01 are often not. We therefore conclude that the sets of covariates are sometimes useful for predicting product use and frequency of product use.

Sensitivity analysis. Various tests were conducted to assess the sensitivity of the results to the number of brands and the functional forms employed in the analysis. We reran our analysis using the five highest share brands in each product category and found no difference in the results. In addition, for these five brands in each category, we conducted an analysis that employed correlated normal errors instead of independent extreme value errors, and an analysis using the most preferred data for category users. The correlated normal error model, or multivariate binomial probit model, lead to minor changes in the predictive evidence reported in Table 8. For example, the number of categories where demographics are predictively related to any use of the brand changes from 48/52 to 34/52, and relative brand choice entry changes from 0/52 to 2/52. It should be noted that these difference are due to two sources—analyzing a smaller number of brands which leads to a loss of information and employing an error distribution that leads to an increase in the amount of information for testing. As mentioned above it was not feasible to obtain an accurate estimate of the likelihood function (needed here for testing purposes) for a high dimensional multivariate binomial probit model. We do find, however, generally consistent and robust results with regard to the assumed error distribution.

The potential benefit of non-linear functional forms was investigated by regressing model residuals (observed minus expected) upon all squared covariates, and all two and three way interactions of the covariates. This regression approach can be shown to be a method of scoring that yields evidence of model misspecification. The R-squared (unadjusted) of these regressions averaged 0.030 with an inter-quartile range of (0.013, 0.040), indicating that use of the linear model was predictively acceptable. Finally, we found that the use of the raw scale values for the pyschographic variables (versus our use of dummy variables) results in changes to the estimated coefficients of order $0.00x$. The sensitivity analysis indicates that the reported results are robust to our modeling assumptions.

6. Discussion

Since the early seventies, it has been known that the relationship between demographic and general psychographic variables and product use is present but not strong. The search for correlates for relative brand preference (e.g., using scanner data) has been less successful, with results suggesting that demographic and general psychographic variables are generally not predictive of brand use. Our study examines the relationship of these variables with product use, brand use, and relative brand preference across 52 product categories, providing evidence that these variables predict product use and unconditional brand use, but do not predict brand choice conditional on product category use. A rationale for our results discusses considerations relating to the nature of demographic and general psychographic variables as candidate explanatory variables on the one hand, and the differing status of product use and brand preference as dependent variables on the other.

The variable of product use can be considered a surrogate for the activities of human lives, i.e., for the tasks and interests that individuals pursue and for which they may consider using some marketplace offering. Using denture cleansers, for example, corresponds to the activity of cleansing an upper denture, which is likely found to occur more often in seniors compared with young adults. Hence, age is a shorthand for certain conditions such as genetic factors, patterns of oral hygiene, diet, sustaining accidental or criminal physical injury, presence of fluoride in the water supply, whose values lead in some cases to an individual's wearing, and hence, needing to care for, an upper denture. Information from the demographic variable of age tells us only that values of such unknown variables favoring the activity of caring for an upper denture are present increasingly with age. Demographic correlates of product use do not provide a direct, substantive understanding of conditions leading to product use. In a marketing context, these variables act as a link with other data sources that marketers use in implementing strategy.

When management has the complete profile of denture wearers, expressed in terms of demographic, trait, attitude, opinion, lifestyle, and interest variables, it will still lack information relevant to formulating, or even devising a promotional strategy for, a brand of denture cleanser. For example, does this denture wearer eat or drink substances that leave intractable stains? Which kinds of stain? Is the individual troubled by denture breath, by arthritic hands that make using a brush difficult, by difficulty dispensing water? Does the individual enjoy certain aromas, and oral sensations, which could be produced by a cleanser that leaves appropriate residue on the denture? No information of this kind can be ascertained from general descriptors, i.e., descriptors that lack information about the specific conditions that pertain within the context for product use (e.g., using denture cleanser) or engaging in the focal activity (e.g., caring for one's upper denture).

General descriptors provide a profile of individuals that is useful in linking with other information that has marketing relevance, e.g., media exposure, retail outlet patronage. While they may provide predictive information as to the type of person that engages in a particular activity for which some version of the product class is useful, they do not provide information on the conditions that make for a satisfactory outcome from using a specific brand. General descriptors such as demographics are too broad-scoped to point to the specific concerns and interests that lead to consumers preferring one brand over another—i.e., they may describe for us who is a prospect for the product category but not what prospects seek in the brands. Accordingly, there is no basis for expecting that demographics and general descriptors are appropriate for explaining relative brand preference (e.g., Fennell, 1978, 1995, pp. 114–115).

Variables that can be expected to explain brand preference must reflect the substantive conditions that lead people to action and potential brand use. Consider the likely nature of such contexts as the product class changes from personal care (e.g., bathing, hair care) to multi person transport to (e.g., autos) to treatment of minor ailments (e.g., remedies for headache, cough) to small appliances (e.g., shavers, pressing irons) to large appliances (refrigerators, snow blowers) to leisure goods of all kinds. The relevant personal and environmental elements vary as we move from one activity domain to another. Within any one domain, brands are valued if their attributes address the specific motivating conditions that lead people to action. These conditions, expressed as concerns and interests associated with a specific occasion of an activity, have been successfully related to brand and attribute preference (Yang et al., 2002) and can be used as basis variables in market segmentation analysis (Allenby et al., 2002; Fennell and Allenby, 2002). Since the motivating conditions arise from the intersection of personal and environmental systems, they reflect a finer (more granular, less coarse) explanation of behavior than demographic and psychographic variables provide, and are more appropriate variables for explaining relative brand preference.

The variables that describe such conditions differ from the independent variables, i.e., the demographics and general psychographic descriptors that we studied here, in being selected from within the substantive domain of the product category. Lacking such substantive content, demographic and general psychographics cannot be regarded as candidate explanatory variables for brand preference.

"Product use," and its corresponding domain of human activity, provides groupings of ways in which people allocate their resources. Essentially, product category names, and names of activity, point to kinds of task and interest to which people allocate resources. It is not uncommon to find that authors use the term, "consumer behavior" or "purchase behavior" leaving it ambiguous whether they intend "product use," a synthetic class of the analyst, or "brand choice," which refers to specific instances of actual behavior (Fennell, 2000). Such a practice can lead, at best, to ambiguity and more likely to conceptual confusion, as it has in the elusive search for what demographics and general psychographic variables can explain.

Acknowledgments

The authors thank Jim Ginter, Rich Johnson and Michel Wedel for helpful comments.

References

Allenby, Greg M. (1990). "Hypothesis Testing with Scanner Data: The Advantage of Bayesian Methods", Journal of Marketing Research 27, 379–389.
Allenby, Greg M. and Peter E. Rossi. (1991). "Quality Perceptions and Asymmetric Switching Between Brands", Marketing Science 10, 185–205.
Allenby, Greg M. and Peter E. Rossi. (1999). "Marketing Models of Consumer Heterogeneity", Journal of Econometrics 89, 57–78.
Allenby, Greg, Geraldine Fennell, Albert Bemmaor, Vijay Bhargava, Francois Christen, Jackie Dawley, Peter Dickson, Yancy Edwards, Mark Garratt, Jim Ginter, Alan Sawyer, Rick Staelin, and Sha Yang. (2002). "Market Segmentation Research: Beyond Within and Across Group Differences", Marketing Letters 13(3), 231–241.
Arora, Neeraj, Greg M. Allenby, and James L. Ginter. (1998). "A Hierarchical Bayes Model of Primary and Secondary Demand", Marketing Science 17, 29–44.
Anderson, W. Thomas Jr. (1971). "Identifying the Convenience-Oriented Consumer", Journal of Marketing Research 8, 179–183.
Bawa, Kapil, Srini S. Srinivasan, and Rajendra K. Srivastava. (1997). "Coupon Attractiveness and Coupon Proneness: A Framework for Modeling Coupon Redemption", Journal of Marketing Research 34, 517–525.
Bass, Frank M., Douglas J. Tigert, and Ronald T. Lonsdale. (1968). "Market Segmentation: Group Versus Individual Behavior", Journal of Marketing Research 5, 264–270.
Bayus, Barry L. and Raj Mehta. (1995). "A Segmentation Model for the Targeted Marketing of Consumer Durables", Journal of Marketing Research 32, 463–469.
Belk, Russell, W. (1974) "An Exploratory Assessment of Situational Effects in Buyer Behavior", Journal of Marketing Research 11, 156–163.
Blattberg, Robert, Thomas Buesing, Peter Peacock, and Subrata Sen. (1978). "Identifying the Deal Prone Segment", Journal of Marketing Research 15, 369–377.
Bucklin, Randolph E., Sunil Gupta, and Sangman Han. (1995). "A Brand's Eye View of Response Segmentation in Consumer Brand Choice Behavior", Journal of Marketing Research 32, 66–74.
Chintagunta, Pradeep K. and Sudeep Haldar. (1998) "Investigating Purchase Timing Behavior in Two Related Product Categories", Journal of Marketing Research 35, 43–53.
Fennell, Geraldine. (1978). "Consumers' perceptions of the product-use situation", Journal of Marketing 42 (April), 38–47.
Fennell, Geraldine. (1995). Globalization and market segmentation: Back to basics. In S.B. MacKenzie and D.M. Stayman (eds.), Proceedings of the Society for Consumer Psychology. Washington, DC: Society for Consumer Psychology, (Division 23), American Psychological Association.
Fennell, Geraldine. (2000). "Product and Brand: Interchangeable Terms?" In R. Keith Tudor, Ernest Capozzoli, and Daryl McKee (eds.), Advances in Marketing.
Fennell, Geraldine and Greg M. Allenby. (2002). "No Brand Segmentation? Let's Not Rush to Judgement", Marketing Research 14, 14–18.
Frank, Ronald E., William F. Massy, and Harper W. Boyd. (1967). "Correlates of Grocery Product Consumption Rates", Journal of Marketing Research 4, 184–190.
Frank, Ronald E., William F. Massy, and Yoram Wind. (1972). Market Segmentation. Englewood Cliffs, NJ: Prentice Hall.
Green, William H. (1993). Econometric Analysis. New York: Macmillan.
Gupta, Sachin and Pradeep K. Chintagunta. (1994). "On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models", Journal of Marketing Research 31, 128–136.
Hanemann, W. Michael. (1984). "Discrete/Continuous Models of Consumer Demand", Econometrica 52, 541–561.
Henry, Walter A. (1976). "Cultural Values Do Correlate With Consumer Behavior", Journal of Marketing Research 13, 121–127.

Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi. (1995). ‘‘Determinants of Store-Level Price Elasticity’’, Journal of Marketing Research 32, 17–29.

Horton, Raymond L. (1979). ‘‘Some Relationships Between Personality and Consumer Decision Making’’, Journal of Marketing Research 16, 233–246.

Johnson, Richard A. and Dean W. Wichern. (1998). Applied Multivariate Statistical Analysis, 4th edition, New Jersey: Prentice Hall.

Kalyanam, Kirthi and Daniel S. Putler. (1997). ‘‘Incorporating Demographic Variables in Brand Choice Models: An Indivisible Alternatives Framework’’, Marketing Science 16, 166–181.

Kinnear, Thomas C. and James R. Taylor. (1976). ‘‘Psychographics: Some Additional Findings’’, Journal of Marketing Research 13, 422–425.

McCann, John M. (1974). ‘‘Market Segment Response to the Marketing Decision Variables’’, Journal of Marketing Research 11, 399–412.

Mittal, Banwari. (1994). ‘‘An Integrated Framework for Relating Diverse Consumer Characteristics to Supermarket Coupon Redemption’’, Journal of Marketing Research 31, 533–544.

Mood, Alexander M., Franklin A. Graybill, and Duance C. Boes. (1974). Introduction to the Theory of Statistics. Third Edition, McGraw-Hill: New York, NY.

Peters, William H. (1970). ‘‘Using MCA to Segment New Car Markets’’, Journal of Marketing Research 7, 360–363.

Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby. (1996). ‘‘The Value of Purchase History Data in Target Marketing’’, Marketing Science 15, 321–340.

Schaninger, Charles M. (1981). ‘‘Social Class Versus Income Revisited: An Empirical Investigation’’, Journal of Marketing Research 18, 192–208.

Schwarz, Gideon. (1978). ‘‘Estimating the Dimension of a Model’’, The Annals of Statistics 6, 461–464.

Sparks, David L. and W.T. Tucker. (1971). ‘‘A Multivariate Analysis of Personality and Product Use’’, Journal of Marketing Research 8, 67–70.

Wedel, Michel and Wagner A. Kamakura. (2000). Marketing Segmentation: Conceptual and Methodological Foundations. Boston, MA: Kluwer.

Wells, William D. (1975). ‘‘Psychographics: A Critical Review’’, Journal of Marketing Research 12, 196–213.

Wildt, Albert R. and John M. McCann. (1980). ‘‘A Regression Model for Market Segmentation Studies’’, Journal of Marketing Research 17, 335–340.

Yang, Sha, Greg M. Allenby, and Geraldine Fennell. (2002). ‘‘Modeling Variation in Brand Preference: The Roles of Objective Environment and Motivating Conditions’’, Marketing Science 21(1), 14–31.