

A Data Set of Brands and Their Characteristics

Mitchell Lovett

Simon Business School, University of Rochester, Rochester, New York 14627, mitch.lovett@simon.rochester.edu

Renana Peres

School of Business Administration, Hebrew University of Jerusalem, Jerusalem, Israel 91905, peresren@huji.ac.il

Ron Shachar

Arison School of Business, The Interdisciplinary Center Herzliya-IDC, Herzliya, Israel 46150, ronshachar@idc.ac.il

Brands stand at the core of marketing. They are central to positioning, marketing communications, word of mouth, customer relationships, and firm profits. Brands have been studied from multiple perspectives using a variety of measures and scales. We offer a data set that contains 136 different measures of the brand characteristics for almost 700 of the top U.S. national brands across 16 categories measured by 2010. These measures cover a broad range of characteristics including brand personality, satisfaction, age, attributes related to Rogers' innovation scheme such as complexity, and the four brand equity pillars of Young and Rubicam's BrandAsset Valuator. The data were collected from a combination of sources including an original survey on 4,769 subjects. In addition, we provide quarterly data on the variables available from the BrandAsset Valuator for two and a half years between 2008 and 2010. These data can be used as a building block in research that aims to explore the antecedents of brand perceptions or connect brand characteristics with market and financial outcomes. This paper describes the data and some relevant research questions.

Data, as supplemental material, are available at <http://dx.doi.org/10.1287/mksc.2014.0861>.

Keywords: brands; brand characteristics; brand personality; complexity; differentiation; esteem

History: Received: October 19, 2013; accepted: February 15, 2014; Preyas Desai served as the editor-in-chief and Peter Fader served as associate editor for this article.

Overview

Brands are viewed as one of the most valuable marketing assets and are a core concept in marketing. Research on brands over the last 70 years covers many topics including brand management, brand extensions, umbrella brands, positioning, and the role of brands in consumers' lives. Today, research on brands continues to be a vibrant area of study with recent contributions linking brands to stock market reactions (Cao and Sorescu 2013) and global venture success (Steenkamp and Geyskens 2014).

Because of their centrality, brands have been studied from multiple perspectives (e.g., strategic and behavioral). Each has shed light on specific elements of brands and developed its own measures and scales to characterize brands. These measures and scales include (1) the four pillars of the BrandAsset Valuator (BAV) tool of Young and Rubicam (Y&R) measuring brand equity via its perceived strength and emotional capital (Mizik and Jacobson 2008, Stahl et al. 2012); (2) the Interbrand ranking that assesses brand equity based on the price premium of branded products (Ailawadi et al. 2003); (3) brand personality traits (Aaker 1997); (4) brand identity and its role in consumers' expression of self-worth (Fournier 1998, Shachar et al. 2011); (5) satisfaction, loyalty, and other customer relationship management (CRM)-related measures (Oliver 1999);

and (6) the characteristics introduced by Rogers (1995) for understanding the diffusion of innovation (such as complexity and perceived risk) that were used both in the context of product categories as well as brands.

The focus of each perspective on its own measures has certainly advanced our understanding of brands. However, it seems likely that a comprehensive and integrative approach that will rely on all of these measures and scales can yield some new insights. Indeed, such an approach was proven effective in understanding brands' role in simulating word of mouth (Lovett et al. 2013).

To encourage research taking the comprehensive and integrative approach to brands and to facilitate cross-perspective research, we provide a unique database on brand characteristics. The data set contains 136 measures of brand characteristics for a cross section of almost 700 top national U.S. brands across 16 product and service categories measured by 2010 (available as supplemental material at <http://dx.doi.org/10.1287/mksc.2014.0861>). These characteristics include the Y&R BAV pillars, brand personality components, Rogers' attributes, satisfaction, and many other characteristics such as age and type of good. The data come from market research companies, as well as from our own data collection. Parts of this data set were used in Lovett et al. (2013). Here we describe the

data set and offer a list of potential research questions it can help to address.

Data Set Description

The data set contains multiple characteristics for 697 major U.S. national brands (both corporate and product) from 16 broad product categories. Since the data set was originally used in the context of word of mouth, the brands are selected to have large volumes of word-of-mouth mentions off-line and over the Internet. However, this list of brands is consistent with other lists used by brand research agencies such as Young and Rubicam or Interbrand. For example, 92 brands of Interbrand's top global 100 list for 2009 are part of our data set. Note that oil and tobacco categories are not included in the data set. Table 1 displays the distribution of brands across categories.

The data have three main sources. Figure 1 displays the sources and the variables extracted from each:

1. *Young and Rubicam's BrandAsset Valuator*. The BAV consulting group of Young and Rubicam conducts a quarterly survey among a representative panel of the U.S. population (17,000 individuals, where each respondent answers on 250 brands each quarter). The survey measures a broad array of perceptions and attitudes for a large number of brands. The data set provided here includes information on 629 of the 697 brands and describes the most recent data point that was available at the end of the second quarter of 2010. The data set also contains a separate file with the quarterly information for quarter 1 of 2008 to quarter 2 of 2010. While some brands were measured in each one of the quarters, others were measured in only a subset—e.g., Circuit City was measured in only 5 out of 10 quarters. The opening worksheet in the quarterly data file shows the data availability for each brand for each of the 10 quarters. Included are not only the values of the specific questions in the Y&R BAV

survey, but also of the four pillars of brand equity that Y&R BAV constructs from them. More details on the measures can be found at <http://bavconsulting.com/>.

2. *Survey*. We developed a survey to measure additional characteristics. The survey was administered online to a representative sample of the U.S. population via the platform of Decipher, Inc., during September–October 2010. We collected data from 4,769 respondents, and each brand was evaluated by at least 35 respondents. Respondents were screened to ensure a high level of brand familiarity. Hence, although the overall sample is representative of the U.S. population, the brand ratings are representative of those familiar with the brand. The variables measured in this survey include product involvement and brand familiarity, excitement, complexity, visibility, and perceived risk. We used existing measurement scales whenever possible. An annotated version of this questionnaire, plus a description of the quotas and response rates, is described in the supplementary material.

3. *Secondary data, Interbrand, and the American Customer Satisfaction Index*. For the remaining measures in the database, we use various secondary data sources. From Interbrand we use the list of brands that were ranked in the top 100 places in 2009. From the American Customer Satisfaction Index (ACSI) we use the measure of brand-level satisfaction. We determine other variables such as the age of the brand, whether it is a premium or a value brand, or whether it is a product or a service from the business press and industry reports or based on independent judges.

We next provide a detailed description of each brand characteristic, the scale, the measures, and the source used to collect it. The first worksheet of the data file contains a dictionary with all of the variable names and a brief description of each of them.

Data from Young and Rubicam's BrandAsset Valuator

1. *Usage*. Usage measures the percentage of respondents who stated that they use the brand occasionally or often.

2. *Consideration*. This is the percentage of respondents who indicate that this is the brand, or one of the several brands, they would consider to buy or use.

3. *Energized_Differentiation*. This Y&R BAV pillar captures the extent to which the brand is perceived as differentiated from other brands. It is measured through items stating whether the brand is different, distinctive, unique, dynamic, and innovative.

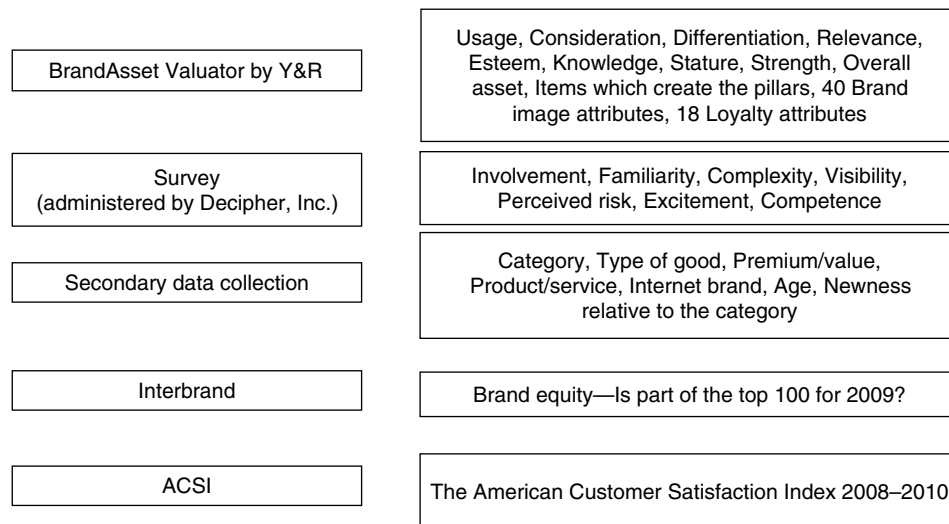
4. *Relevance*. This Y&R BAV pillar captures the average on the question, How appropriate is the brand for you personally?

5. *Esteem*. This Y&R BAV pillar captures the extent to which people hold a brand in high esteem. It is measured through items asking about the regard, leadership, reliability, and quality of the brand.

Table 1 Distribution of the 697 Brands in the Data Set Across Categories

Category	Number of brands	% of brands
Food and dining	105	15.1
Media and entertainment	103	14.8
Beverages	66	9.5
Technology products and stores	56	8.0
Beauty products	52	7.5
Clothing products	51	7.3
Cars	47	6.7
Financial services	39	5.6
Travel services	34	4.9
Health products and services	27	3.9
Telecommunications	25	3.6
Household products (cleaning ingredients, etc.)	24	3.4
Sports and hobbies	21	3.0
Children's products	19	2.7
Department stores	15	2.2
Home design and decoration	13	1.9

Figure 1 Data Sources and Variables



6. *Knowledge*. This Y&R BAV pillar indicates the level of intimate understanding of the brand.

7. *Stature, Strength, and Overall Asset*. The scores of the four pillars are combined into three scores termed stature, strength, and overall asset. See the data files for details.

8. *The items that created the pillars*. These are the individual items that make up *Energized_Differentiation* and *Esteem*. Note, that due to quota weighting issues, the score of the pillars cannot be derived directly from the items in the data set (i.e., averaging the items for *Energized_Differentiation* for a brand is not necessarily equal to the *Energized_Differentiation* score).

9. *Brand image*. Respondents were asked to check whether they can associate the brand with 40 attributes, such as *arrogant*, *authentic*, *friendly*, etc. For each item, the data set contains the percentage of respondents who checked this attribute with respect to the brand. Then, these 40 attributes are factor analyzed into eight brand personality factors, which are also part of the data set. These traits are *cutting-edge*, *classic*, *superior*, *chic*, *customer-centric*, *outgoing*, *no-nonsense*, and *distant*. The exact factor analysis coefficients are proprietary to Young and Rubicam.

10. *Loyalty*. This is a set of 18 attributes that capture brand attitudinal loyalty. For each item, the data set contains the percentage of respondents who checked this attribute with respect to the brand.

Data from Our Survey

11. *Involvement*—survey questions Q2–Q4. We use the involvement scale of Ratchford (1987). This is a three-item, five-point scale, which asks about the importance of the purchase decision, the amount of thought invested in the decision, and the risk of making the wrong decision. For simplicity, and since there is not much variability in involvement levels of different

brands within a category, involvement was measured at the category level. Therefore, the involvement score is the same for all of the brands in a category.

12. *Familiarity*—Q5. This single item familiarity question asks to what extent the respondent is familiar with the brand.

13. *Complexity*—Q6_1–Q6_5. These five items form a measure based on Moore and Benbasat (1991) and Speier and Venkatesh (2002).

14. *Visibility*—Q6_6–Q6_10. This variable comes from the observability construct of Rogers (1995) using a five-item, five-point scale based on Moore and Benbasat (1991).

15. *Perceived risk*—Q6_11–Q6_13. These items are based on Rogers' (1995) definition of perceived risk as the functional, financial, and emotional uncertainty associated with the product. We use the full three-item scale of Ostlund (1974).

16. *Excitement*—Q7_1–Q7_11. These items come from Aaker's (1997) excitement scale.

17. *Competence*—Q7_12–Q7_20. These items come from Aaker's (1997) competence scale.

Data from Secondary Sources

18. *Category*. Brands are classified into 16 categories as depicted in Table 1. This classification is based in principle on identifying the industry the brand belongs to. In case of multiple possible categories, the largest share of business for the brand was chosen. The Keller-Fay Group, a market research company involved in linking us to the data set, performed the classification task.

19. *Type of good*. We used the classification of Nelson (1974) and Laband (1986) to divide the brands into search, experience, and credence goods. Using the definitions from the literature, two independent judges separately classified the subcategories. The intercoder

agreement was 72%, and the judges resolved all disagreements by consensus.

20. *Premium/value*. Each brand was classified as one of the following: premium, value, or middle. Two independent judges classified the brands relative to the product type (e.g., Clinique was evaluated relative to beauty products, and Hilton with respect to other hotels). The intercoder agreement was 70%, and the judges resolved all disagreements by consensus. In formulating these classifications, the judges used secondary data on various aspects such as the brand's price relative to the price of other brands in the category.

21. *Product/service*. Two independent judges classified each brand on the list into one of the following: product, service, or mix, if the brand offers a mix of a product and a service. The intercoder agreement was 82%, and the judges resolved all disagreements by consensus.

22. *Internet brand*. Seventeen of the brands on our list such as eBay, Amazon, Expedia, and Google are Internet-based services. We provide a binary variable for whether the brand is an Internet brand or not. Intercoder agreement was 100%.

23. *Age*. We define age as the time (in years) elapsed from the commercial launch of the brand to the reference date, August 1, 2010. We obtained the data from brand publications and from historical business and press data.

24. *Newness relative to the product type*. We define newness as the time interval (in years) between the U.S. national commercial launch of the first brand of this product type to the brand's commercial launch. For example, the toothpaste brand Crest was introduced in 1955, but the first U.S. national commercial toothpaste brand dates back to 1800 with a brand called Crème Dentifrice. Therefore its relative newness is 155 years. Classification was done relative to the product type.

25. *Interbrand top list*. Based on Interbrand's list of top 100 brands for 2009, we code a binary variable indicating whether the brand is on the list or not.

26. *Satisfaction (ACSI)*—The American Customer Satisfaction Index is a standard measure of satisfaction for American corporate brands (Fornell et al. 1996), collected each quarter using 250 customer telephone interviews per brand on a rolling set of brands with each receiving at least one measure each year. Of our list of brands, 209 have an ACSI score (with Heinz having the highest score and Charter Communications having the lowest). The data set contains the average ACSI index for 2008–2010.

Next, we describe the data and correlations among the variables.

descriptions of some of the variables. Although we cannot feasibly present all of the variables, we select some of the main variables that we think are likely to be of interest to users of this data set. Table 2 presents the descriptive statistics of these variables, and Table 3 contains the category averages. We point to a few interesting patterns in these data. From Table 3, we see high usage and consideration for food, department stores, and household products. The last two categories also enjoy high brand strength and stature. Also interesting is the low satisfaction for service categories such as health, telecom, and travel relative to that of goods categories. Financial services score high on complexity and low on esteem, whereas technology, which is also perceived as complicated, is highly esteemed by respondents.

Table 4 contains correlations for these same variables. As expected, usage is highly correlated with consideration and familiarity, and also with the Y&R BAV pillars. It is interesting to see its negative correlation with risk and complexity, a correlation that might stem either because less used brands are perceived as risky and complex, or the other way around—their risk and complexity leads to lower usage. Other interesting correlations include the positive correlations of risk and complexity, and the negative correlations between risk and satisfaction, as well as the Y&R BAV pillars.

As mentioned above, for the Y&R BAV, data are available for 10 quarters. To illustrate the nature of variation in the data, we focus on the brands that are available across all 10 quarters. Figure 2 describes the car, food and dining, and household product categories. It depicts the category average over time for the four Y&R BAV pillars. As the figure indicates, time variation is not large, and only the knowledge pillar shows any systematic pattern (decline). Table 5 presents the descriptive statistics for the four Y&R BAV pillars and the eight Y&R BAV personality factors. We present standard summary statistics as well as the mean squared errors (MSEs) from a variable-by-variable analysis of variance (ANOVA) containing brand and time factors. As the mean squared error measures (significance in parentheses) indicate, both factors are significant, suggesting both sources of variation are present. More importantly, brand plays a much larger role than time in explaining the variation. In fact, for most variables, brand dwarfs time, suggesting cross-sectional variation is much more important than time variations. We also conducted an unreported ANOVA including category and category * time interactions. The category by time interactions were not significant, suggesting that the time trends are relatively similar across categories.

Descriptive Statistics

This section provides a brief view into the rich detail and variation available in the data set by presenting

Potential Research Questions

This data set can be used on its own or with other data to shed light on managing and building brands as

Table 2 Descriptive Statistics

	Mean	Std. dev.	Min.	Max.	Max. value brand	Min. value brand	Obs.
Age	54.19	39.54	1.10	204.66	Colgate	<i>Transformers: Revenge of the Fallen</i>	697
Newness relative to the category	64.48	61.67	0.00	370.07	*	Multiple (e.g., Amazon, American Express)	697
Satisfaction	79.40	6.56	55.00	89.33	Heinz	Charter Communications	209
Involvement ^a	3.72	0.36	3.09	4.38	Financial Services	Beverages	697
Familiarity	3.30	0.62	0	4.62	Band-Aid	HEB Grocery	697
Complexity	1.84	0.38	1.01	3.31	Medicare	Pledge	693
Visibility	2.99	0.38	1.54	3.99	Microsoft	Lamborghini	693
Perceived risk	1.80	0.31	1.02	2.72	Medicare	Dr. Pepper	693
Excitement	3.32	0.36	2.16	4.44	iPhone	Medicare	695
Competence	3.50	0.28	2.61	4.45	iPhone	Diet Mountain Dew	695
Usage	33.35	22.35	0.28	89.31	Band-Aid	Porsche	618
Consideration	40.48	16.63	9.18	82.72	Hershey	Regions Bank	618
Energized differentiation	0.50	0.16	0.17	1.12	Food Network	Days Inn	629
Relevance	2.74	0.72	1.39	4.75	Kraft	Saab	629
Esteem	0.61	0.30	0.09	1.67	Tylenol	<i>Ugly Betty</i>	629
Knowledge	3.54	0.90	0	5.16	Walmart	Shaw's Supermarket	630
Brand stature	2.33	1.54	0.12	8.10	Tylenol	Pacific Sunwear	629
Brand strength	1.36	0.56	0.31	4.24	Discovery Channel	Alamo	629
Overall asset	3.66	3.63	0.10	23.98	Google	Pacific Sunwear	629
Type of good				Search, 20.5%; experience, 73.2%; credence, 6.3%			697
Premium/value				Premium, 24.8%; middle, 50.5%; value, 24.7%			697
Product/service				Product, 52.5%; service, 44.3%; mix, 3.2%			697
Internet				Internet, 2%; non-Internet, 98%			697
Interbrand top list				Part of Interbrand top 100, 12%; not a part of Interbrand top 100, 88%			697

^aInvolvement is measured at the category level.

*There are multiple versions as to the exact name of the first commercial soap maker, all providing convergent evidence that it dates back to 1620.

well as the role of brands in marketing and economics. Here are some initial ideas:

1. *The antecedents of brand perceptions.* Understanding what influences brand perceptions is an important line of research that these data can support. For example, one can study the dependence of these perceptions

on market factors, past investments, date of launch, competition, or the presence of similar brands in the category.

2. *The connection between brand characteristics and features of social networks.* Brand characteristics were already shown to be associated with word of mouth

Table 3 Variables Averages—Breakdown by Category

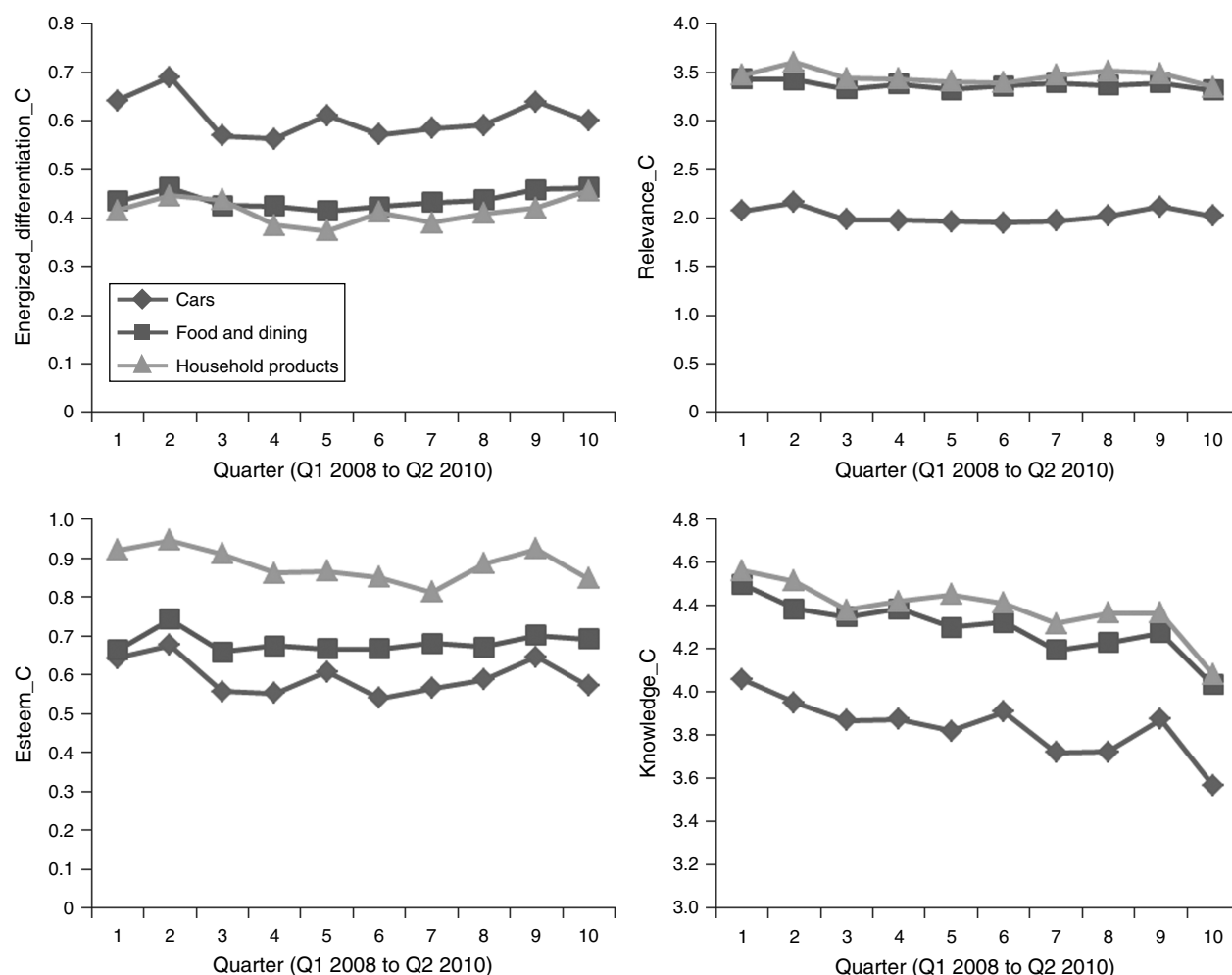
	Beauty	Bevr.	Cars	Child.	Cloth.	Dept. stores	Financ.	Food	Health	Home	House	Media	Sports	Tech.	Telecom	Travel
Age	70.55	61.16	60.63	57.25	59.37	41.02	80.71	63.57	68.58	65.42	70	20.32	73.73	43.53	21.68	48.77
Newness relative to the category	108.86	90.33	43.89	67.72	113.7	61.94	101.72	38.78	67.7	80.62	55.69	56.92	20.47	31.74	24.78	67.83
Satisfaction	85.14	84.21	82.61	85	77.08	78.64	75.89	81.17	72.5	80.83	85.17	74.33		74.11	69.92	71.02
Involvement	3.69	3.09	4.32	3.82	3.52	3.18	4.38	3.55	4.17	3.95	3.48	3.62	3.4	4.03	3.91	3.97
Familiarity	3.62	3.19	3.15	4.01	2.95	3.43	2.99	3.31	3.62	3.8	3.92	3.15	3.33	3.26	3.01	3.51
Complexity	1.55	1.62	1.89	1.85	1.95	1.95	2.38	1.55	1.97	1.69	1.36	1.91	2.1	2.04	2.14	2
Visibility	2.94	3.06	3.09	3.26	2.91	3.07	2.68	3.07	3.02	3.08	3.24	2.93	2.63	3.01	2.96	2.98
Perceived risk	1.58	1.6	1.94	1.83	2.03	1.95	2.11	1.62	1.92	1.74	1.38	1.84	2.09	1.8	2.09	1.91
Excitement	3.41	3.22	3.36	3.82	3.44	3.26	3	3.17	2.71	3.55	3.23	3.57	3.4	3.48	3.27	3.24
Competence	3.63	3.21	3.55	3.84	3.4	3.57	3.48	3.41	3.34	3.72	3.65	3.5	3.57	3.71	3.46	3.54
Usage	39.04	38.46	6.4	27.87	29.09	50.12	11.32	51.68	34.84	37.41	55.71	39.79	31.72	22.48	11.08	18.27
Consideration	41.75	38.64	29.18	43.18	35.62	50.46	22.69	52.12	40.89	51.55	57.79	39.5	31.38	39.62	25.35	42.84
Energized differentiation	0.46	0.45	0.56	0.46	0.54	0.51	0.36	0.44	0.37	0.61	0.46	0.65	0.47	0.62	0.45	0.35
Relevance	2.95	2.74	2.17	2.52	2.54	3.42	2.19	3.27	2.87	3.13	3.56	2.62	2.08	2.83	2.28	2.23
Esteem	0.68	0.51	0.59	0.77	0.53	0.86	0.45	0.68	0.73	0.89	0.95	0.49	0.36	0.71	0.44	0.53
Knowledge	3.69	3.69	3.72	3.35	3.16	3.88	2.77	3.74	3.55	3.77	4.19	3.66	3.62	3.28	3.02	3.34
Brand stature	2.65	2.01	2.3	2.73	1.89	3.55	1.39	2.78	2.82	3.51	4.04	1.91	1.36	2.52	1.47	1.84
Brand strength	1.34	1.23	1.18	1.17	1.37	1.73	0.79	1.45	1.04	1.82	1.65	1.72	0.99	1.77	1.03	0.78
Overall asset	3.71	2.88	2.96	3.56	3.04	6.66	1.31	4.45	3.43	6.55	6.86	4.06	1.5	5.12	1.84	1.47

Table 4
Correlations

	Age	New rel.	Satis.	Invol.	Famil.	Comp.	Visib.	Risk	Excite.	Compt.	Usage	Consid.	Ener. diff.	Relev.	Esteem	Know.	Stat.	Strngth.	Ovrl. asset
Age	1																		
New rel.	−0.161**	1																	
Satis.	0.134	0.180**	1																
Invol.	0.039	−0.067	−0.347**	1															
Familiarity	0.206**	−0.065	0.104	−0.018	1														
Complexity	−0.160**	0.025	−0.545**	0.382**	−0.548**	1													
Visibility	0.105**	−0.101**	0.126	−0.069	0.571**	−0.496**	1												
Risk	−0.182**	0.031	−0.598**	0.312**	−0.545**	0.813**	−0.504**	1											
Excitement	−0.319**	0.010	0.227**	−0.084*	−0.064	0.085*	0.118**	0.014	1										
Competence	0.113**	−0.112**	0.115	0.238**	0.288**	−0.013	0.301**	−0.215**	0.524**	1									
Usage	0.143**	−0.069	0.235**	−0.461**	0.683**	−0.654**	0.445**	−0.636**	0.195**	0.873**	1								
Consideration	0.227**	−0.085*	0.209**	−0.235**	0.787**	−0.644**	0.513**	−0.657**	0.191**	0.873**	0.873**	1							
Ener. diff.	−0.307**	−0.110**	0.194**	−0.235**	0.012	−0.259**	0.011	−0.259**	0.012	−0.259**	0.012	−0.259**	1						
Relevance	0.247**	−0.099*	0.250**	−0.259**	0.690**	−0.616**	0.535**	−0.668**	0.191**	0.873**	0.873**	0.873**	0.873**	1					
Esteem	0.348**	−0.145**	0.230**	0.266**	0.230**	0.230**	0.230**	0.230**	0.230**	0.230**	0.230**	0.230**	0.230**	0.230**	1				
Knowledge	0.238**	−0.091*	0.266**	−0.170**	0.780**	−0.527**	0.461**	−0.475**	0.022	0.741**	0.741**	0.741**	0.741**	0.741**	0.741**	1			
Stature	0.342**	−0.141**	0.234**	−0.053	0.749**	−0.479**	0.507**	−0.487**	0.109**	0.884**	0.884**	0.884**	0.884**	0.884**	0.884**	0.884**	1		
Strength	−0.090*	−0.143**	0.242**	−0.166**	0.394**	−0.263**	0.407**	−0.350**	0.340**	0.345**	0.503**	0.514**	0.514**	0.514**	0.514**	0.514**	0.514**	1	
Overall	0.186**	−0.153**	0.195**	−0.092*	0.613**	−0.367**	0.504**	−0.431**	0.044	0.355**	0.638**	0.710**	0.350**	0.776**	0.882**	0.636**	0.901**	0.799**	1

*Correlation is significant at the 0.05 level (two-tailed); **correlation is significant at the 0.01 level (two-tailed).

Figure 2 Average Values as a Function of Time for the Four Y&R BAV Pillars for Three Categories



(Lovett et al. 2013), but they might also be related to other aspects of social networks (such as the speed that information diffuses through social networks or the effectiveness of seeding).

3. *Brand networks.* It was recently shown that brands exist as part of a network in which purchasing one is related to another not just due to substitution effects (e.g., Oestreicher-Singer et al. 2013). One could examine whether the nature of such networks and the connections within it are related to brand characteristics.

4. *Marketing activities and market outcomes.* Research on the relationship between marketing activities and market outcomes has a long history. With this data set one could study whether the relationship depends on brand characteristics. For example, a study of the efficiency of a certain advertising campaign, or a brand promotion, on sales, might benefit from including brand characteristics (e.g., type of good, age, differentiation, and visibility) as either moderators or controls.

5. *The interdependence of brand characteristics.* As illustrated above, there are some interesting relationships among the different characteristics. The data

Table 5 Time Variability of Some of the Y&R BAV Variables

	Mean	Std. dev.	Min.	Max.	Median	MSE brand	MSE time
Usage	34.92	23.43	0.07	91.87	31.18	5,553.4*	278.2*
Consideration	41.93	17.77	6.63	88.00	40.55	2,996.32*	305.78*
Energized differentiation	0.48	0.16	0.12	1.24	0.45	0.21*	0.15*
Relevance	2.78	0.77	1.11	4.92	2.70	5.65*	0.28*
Esteem	0.63	0.32	0.05	1.97	0.57	0.86*	0.19*
Knowledge	3.79	0.96	−1.00	5.66	3.97	8.15*	3.82*
Cutting-edge	8.05	2.75	2.66	19.72	7.49	66.7*	41.56*
Classic	12.91	4.80	2.15	32.93	12.27	191.49*	75.72*
Superior	12.08	4.67	2.56	35.97	11.18	195.88*	60.29*
Chic	7.00	4.07	1.61	29.60	5.65	164.87*	22.52*
Customer-centric	13.02	4.81	2.79	31.61	12.72	192.67*	92.35*
Outgoing	10.82	4.83	1.79	34.84	9.68	217.6*	56.15*
No-nonsense	8.26	2.58	2.49	18.40	8.04	56.21*	32.45*
Distant	5.55	2.21	1.58	19.34	5.03	37.21*	16.74*

*F value is significant at the 0.001 level.

can assist in directing and testing theories about these relationships.

6. *Substitution based on brand characteristics.* In typical models in marketing, products and brands are mapped into categories based on functional characteristics of the product, and brand substitution is measured based on purchases. Our data enable a different means of exploring competition by using brand characteristics to define the similarity of brands within a category. For example, are brands with similar complexity scores perceived as closer substitutes than brands that differ in their complexity? If two brands are perceived as high on excitement, do they compete more intensely with each other than with less exciting brands?

7. *The role of satisfaction.* The satisfaction–loyalty connection has been explored in the CRM literature (e.g., Richins 1983). This connection might depend on brand characteristics (e.g., for exciting brands, high satisfaction might convert more or less easily into an actual purchase or retention).

8. *Brand characteristics and brand loyalty.* Brand loyalty, both in terms of retention and attitude, is considered to be a desired outcome in the CRM literature. Our data can be used to test to what extent it depends on the brand characteristics versus the firm's CRM policy (e.g., are brands with certain characteristics more robust to service failures?; does retention rate or repeat purchase depend on the brand's level of differentiation or esteem?).

9. *Brand characteristics and the financial value of the brand.* Assessing the financial value of brands has a long tradition with various methodologies. Both the cross-sectional and the longitudinal data can be leveraged to shed new light on this question.

10. *The evolution of brand perception.* Using our time varying data for the Y&R BAV variables, as well as additional data collection on the other variables, one can study the evolution of brand perceptions over time.

Limitations

To some degree the data set arrives with an expiration date. For all research questions that require additional data sources (e.g., point 2 above), the brand data set is useful only if these other data are available for a similar time period. Otherwise, the measures of brand perceptions may have changed and may be less relevant. Of course, (1) in many cases it is easy to collect data that describe things as they were in 2010 (e.g., for the purpose of point 1 above), (2) there are many research questions that do not require additional data sources (e.g., point 5 above), and (3) in some cases having data from 2010 is an advantage (e.g., point 9 above).

In addition, one other important limitation is the sample selection. The set of brands is made up of large,

widely known brands and lacks smaller, lesser-known brands. Although in some competitive settings (e.g., telecom and computers) the data include all of the major players, for other settings the data may be sparse. This could limit the usefulness of the data set (without further data collection) for some purposes.

Supplemental Material

Supplemental material to this paper is available at <http://dx.doi.org/10.1287/mksc.2014.0861>.

Acknowledgments

The authors thank all those who helped to collect the data set. They thank their industry collaborators Brad Fay from the Keller Fay Group and Ed Lebar and John Gerzema from the BAV consulting group at Y&R for sharing their data. They thank Kristin Luck and the Decipher, Inc. team for programming and managing the survey. They gratefully thank their research assistants Christina Andrews, Linda Wang, Chris Webber-Deonauth, Derric Bath, Grace Choi, Rachel Amalo, Yan Yan, Niels Mayrargue, Nathan Pamart, and Fangdan Chen at Wharton, and Yair Cohen, Dafna Presler, Oshri Weiss, Liron Zaretsky, Anna Proviz, Tal Tamir, and Haneen Matar at the Hebrew University. They thank Christophe Van den Bulte and Shawndra Hill from Wharton who came up with the idea of making this data set public. They also thank the review team for helpful suggestions. This data collection was supported by the Marketing Science Institute, the Wharton Customer Analytics Initiative, the Israeli Internet Association, Kmart International Center for Marketing and Retailing at the Hebrew University of Jerusalem, the Israel Science Foundation, and the Marketing Department at the Wharton School.

References

- Aaker JL (1997) Dimensions of brand personality. *J. Marketing Res.* 34(3):347–356.
- Ailawadi KL, Lehmann DR, Neslin SA (2003) Revenue premium as an outcome measure of brand equity. *J. Marketing* 67(4):1–17.
- Cao Z, Sorescu A (2013) Wedded bliss or tainted love? Stock market reactions to the introduction of co-branded products. *Marketing Sci.* 32(6):939–959.
- Fornell C, Johnson MD, Anderson EW, Cha J, Bryant BE (1996) The American customer satisfaction index: Nature, purpose, and findings. *J. Marketing* 60(4):7–18.
- Fournier S (1998) Consumers and their brands: Developing relationship theory in consumer research. *J. Consumer Res.* 24(4):343–353.
- Laband DN (1986) Advertising as information: An empirical note. *Rev. Econom. Statist.* 68(3):517–552.
- Lovett M, Peres R, Shachar R (2013) On brands and word of mouth. *J. Marketing Res.* 50(4):427–444.
- Mizik N, Jacobson R (2008) The financial value impact of perceptual brand attributes. *J. Marketing Res.* 45(1):15–32.
- Moore GC, Benbasat I (1991) Development of an instrument to measure the perceptions of adopting an information technology innovation. *Inform. Systems Res.* 2(3):192–222.
- Nelson P (1974) Advertising as information. *J. Political Econom.* 82(4):729–754.
- Oestreicher-Singer G, Libai B, Sivan L, Carmi E, Yassin O (2013) The network value of products. *J. Marketing* 77(1):1–14.

- Oliver RL (1999) Whence consumer loyalty? *J. Marketing* 63(Special issue):33–44.
- Ostlund LE (1974) Perceived innovation attributes as predictors of innovativeness. *J. Consumer Res.* 1(August):23–29.
- Ratchford BT (1987) New insights about the FCB grid. *J. Advertising Res.* 27(4):24–38.
- Richins ML (1983) Negative word-of-mouth by dissatisfied consumers: A pilot study. *J. Marketing* 47(1):68–78.
- Rogers EM (1995) *The Diffusion of Innovations* (Free Press, New York).
- Shachar R, Erdem T, Cutright KM, Fitzsimons GJ (2011) Brands: The opiate of the nonreligious masses? *Marketing Sci.* 30(1):92–110.
- Speier C, Venkatesh V (2002) The hidden minefields in the adoption of sales force automation technologies. *J. Marketing* 66(3):98–111.
- Stahl F, Heitmann M, Lehmann DR, Neslin SA (2012) The impact of brand equity on customer acquisition, retention, and profit margin. *J. Marketing* 76(4):44–63.
- Steenkamp J-BEM, Geyskens I (2014) Manufacturer and retailer strategies to impact store brand share: Global integration, local adaptation, and worldwide learning. *Marketing Sci.* 33(1):6–26.