

Web Appendix

Appendix 1: Brand List

The list of brands was constructed using the Keller and Fay TalkTrack free recall data. We dropped non-commercial brands (e.g., political candidates, cities) and brands that are periodic (e.g., Olympic games, WorldCup) or discontinued (e.g., Linens and Things), and focused on the top brands in terms of offline WOM. We then confirmed that this list of brands contained all the brands listed in commonly referenced online top brands lists (e.g., from NMIcrite BuzzMetrics). The respondents provide the brand mentions in free entry format so that the list is constructed after the fact based on analysis of the text. Since respondents had no limits on the brands they chose to recall, the list is not constrained by company definitions. For instance, we have in some cases the corporate and product specific brands related to the same company such as iPhone and Apple or Xbox 360 and Microsoft.

Beauty Products

Always
Arm And Hammer
Aussie Conditioner
Aveda
AVEENO
AVON
AXE
Bath & Body Works
Caress
Chanel
Charmin
Clairol
Clean & Clear
Clinique
Colgate
CoverGirl
Crest
Degree
Dial Soap
Dove (Personal Care)
Estee Lauder
Garnier Fructis
Gillette
Head & Shoulders
Herbal Essence
Irish Spring
Ivory
Jergens
Kleenex
Kotex
Lancome
Listerine
Loreal
Mary Kay
Maybelline
Neutrogena
Nivea
Noxzema
Oil of Olay
Old Spice
Pantene
Playtex
ProActiv
Revlon
Scott Tissue
Secret
Sephora
St. Ives
Suave
Tampax
Tresemme
Zest

Beverages

7UP
A and W Root Beer
Absolut
Anheuser Busch
Aquafina
Arizona Beverage

Bacardi
Bud Light
Budweiser
Canada Dry
Captain Morgan
Cherry Coke
Coca-Cola
Coca-Cola Diet
Coca-Cola Zero
Coors
Coors Light
Corona
Crystal Light
Dasani Water
Diet Mountain Dew
Diet Pepsi
Dr Pepper
Dr Pepper Diet
Fanta
Folgers Coffee
Fresca
Gatorade
Grey Goose
Guinness
Heineken
Jack Daniels
Jose Cuervo
Juicy Juice
KoolAid
Lipton
Maxwell House
Michelob
Mikes Hard Lemonade
Miller Brewing
Miller Lite
Minute Maid
Monster Energy Drink
Mountain Dew
Mug Root Beer
Nestea
Ocean Spray
Patron Tequila
Pepsi
Poland Spring
Powerade
Propel Fitness Water
Red Bull
Sam Adams
Sierra Mist
Smirnoff
Snapple
Sobe
Sprite
Sunkist
Sunny Delight
Tropicana
V-8 Juice
Vault Energy Drink
Vitamin Water
Welch

Cars

Acura
Audi
Autozone
BMW
Buick
Cadillac
Chevrolet
Chrysler
Corvette
Dodge
ExxonMobil
Ferrari
Firestone
Ford
GM
GMC
Good Year Tires
Harley Davidson
Honda
Hyundai
Infiniti
Jaguar
Jeep
Jiffy Lube
Kia
Lamborghini
Land Rover/Range Rover
Lexus
Lincoln
Mazda
Mercedes Benz
Mercury
Mini Cooper
Mitsubishi
Napa
Nissan
Pep Boys
Pontiac
Porsche
Saab
Subaru
Suzuki
Toyota
Toyota Scion
Volkswagen
Volvo
Yamaha

Children's Products

Babies R Us
Carter's
Children's Place
Enfamil
Fisher Price
GAP Baby
Gerber
Graco
Gymboree
Huggies
LeapFrog
Lego
Little Tikes
Luvs
Mattel
OshKosh
Pampers
Playskool
Toys R Us

Department stores

Amazondotcom
Barnes & Noble
Big Lots
BJs
Borders
Costco
Kmart
Meijer
Office Depot
Overstock.com
Sams Club
SEARS
Staples
Target
Walmart

Home design and decoration

Ashley Furniture
Bed Bath & Beyond
Behr Paint
GE
Home Depot
Ikea
Kenmore
La-Z-Boy
Maytag
Menards
Pier 1 Imports
Pottery Barn
Whirlpool

Clothing Products

Abercrombie & Fitch
Adidas
Aeropostale
American Eagle
Ann Taylor
Armani
Baby Phat
Banana Republic
Belk
Bloomingdales
Chicos
Claire's
Coach
Converse
Dick's Sports
Dillard's
Dolce & Gabbana
Eddie Bauer
Foot Locker
Forever 21
GAP
Gucci
H&M
Hollister
Hot Topic
JC Penney
JCrew
JoAnn Fabrics
Kohls
Lane Bryant
Levis
Louis Vuitton
Lowe's
Macy's
Marshalls
New Balance
Nike
Nordstrom
Old Navy
Pac Sun (Pacific Sunwear)
PayLess
Polo
Prada
Ralph Lauren
Reebok
Ross
TJ Maxx
Tommy Hilfiger
Under Armour
Victorias Secret
Wilson

Financial services

AIG
Allstate
American Express
Bank of America
BB&T Bank
Capital One
Charles Schwab
Citibank
Citizens Bank
Discover Card
Dow Jones
Edward Jones
Etrade
Fidelity Investments
Fifth Third Bank
GEICO
H&R Block
HSBC
ING Direct
JP Morgan Chase
Mastercard
Merrill Lynch
MetLife
Morgan Stanley/Dean Witter
National City Bank
Prudential
Regions Bank
Scottrade
Smith Barney
State Farm
SunTrust
TD Ameritrade
TRowe Price
US Bank
USAA
Vanguard
VISA
Wachovia
Wells Fargo

Food and Dining

Albertsons
Applebees
Arbys
Banquet
Bertolli
Betty Crocker
Bob Evans
Boston Market
Breyers Ice Cream
Burger King
Butterball
Campbell
Cheerios
Cheesecake Factory
Cheetos
Chick-Fil-A
Chilis
Chipotle
Cracker Barrel
Dairy Queen
Dannon
Del Monte
Dennys
Digiorno
Dole
Dominos Pizza
Doritos
Dunkin Donuts
Fred Meyer
Frito Lay
General Mills
Giant Eagle
Giant Food
Great Value
Healthy Choice
HEB Grocery
Heinz
Hershey
Hostess
Hot/Lean Pockets
Hunts
Ihop
Jack in the Box
Jello
Kelloggs
KFC
Kraft
Kroger
Lays Chips

Lean Cuisine
Long John Silvers M&M
Marie Callender
McDonalds
Nabisco
Nestle
Olive Garden
Oreos
Oscar Mayer
Outback Steakhouse
Panda Express
Panera
Papa Johns
Pathmark
Perdue Chicken
PF Chang
Pillsbury
Pizza Hut
Popeyes
Post Cereal
Prego
Publix
Quaker Oats
Quiznos
Ragu
Ralphs Grocery
Red Lobster
Red Robin
Romanos Macaroni Grill
Ruby Tuesdays
Safeway
Sara Lee
Shaw's Supermarket
Shop Rite
Slim Fast
Snickers
Sonic
Starbucks
Stouffers
Subway
Swansons
Taco Bell
Texas Roadhouse
TGI Fridays
Tostitos
Trader Joes
Tyson
Velveeta
Vons
Wegmans
Wendys
White Castle
Whole Foods
Winn Dixie
Yoplait

Health Products

Advil
Aetna
Aleve
Band Aid
Bayer
Benadryl
Blue Cross
Blue Cross/Blue Shield
Cigna
Claritin
CVS
Excedrin
GNC
Humana Healthcare
Ibuprofen
Johnson & Johnson
Kaiser Permanente
Liptor
Medicare/Medicaid
Merck
Motrin
Pfizer
Prilosec
Rite Aid
Tylenol
United Health
Walgreens
Weight watchers

Household Products

Ajax
Black & Decker
Cascade
Cheer
Clorox
Dawn
Downy
Febreze
Gain
Hoover
Kitchen Aid
Lysol
Mr Clean
Oxy Clean
P&G
Palmolive
Pet Smart
Pine Sol
Pledge
Purex
Purina
Swiffer
Tide
Windex

Media and entertainment

24TVShow
300 (the movie)
ABC
Amazing Race
AMC Theater
American Idol
America's Got Talent
America's Next Top Model
BBC
BET
Big Brother
Blockbuster
Cars (the movie)
Cartoon Network
CBS
Closer
CNBC
CNN
Comedy Central
Criminal Minds
CSI
Dancing with the Stars
Deal or No Deal
Desperate Housewives
DirectTV
Discovery Channel
Disney
E!
Ebay
ER
ESPN
Everybody Loves Raymond
Facebook
Family Guy
Food Network
Fox
Fox News
Friends
Fringe (TV Show)
General Hospital
Google
Gossip Girl
Greys Anatomy
Hallmark
Hancock (the movie)
Harry Potter
HBO
Hells Kitchen
Heroes (TV show)
House (TV Show)
I am Legend (the movie)
Incredible Hulk (the movie)
Indiana Jones (the movie)
Iron Man (the movie)
Jeopardy
Juno (the movie)
Law And Order
Lifetime Television
Lost
Money Magazine

MSN
MSNBC
MTV
MySpace.com
NBC
NCIS
NetFlix
Nickelodeon
NY Times
One Tree Hill
Oprah
PBS
People Magazine
Pirates of the caribbean (the movie)
Prison Break
Scrubs
Sex and the City (the movie)
Shrek (the movie)
Simpsons
Sirius
Smallville
So You Think You Can
Dance (the movie)
Sopranos
South Park
Spiderman
Sponge Bob Square Pants
Star Trek (the movie)
Survivor
The Dark Knight (the movie)
The Office
Time Warner
TNT
Transformers: Revenge of the Fallen (the movie)
Twilight (the movie)
Two and a Half Men
Ugly Betty
Universal Studios
Vh1
Wall Street Journal
Wall-E (the movie)
Wheel of Fortune
Yahoo
You Tube

Technology products and stores

Acer
AMD
Apple
Best Buy
Bose
Brother
Call of Duty (the video game)
Canon
Circuit City
Compaq
Creative Labs
Dell
emachines
Epson
Fuji
Garmin
Lexmark
LG
Microsoft
Nikon
Nintendo
Nintendo DS/DSLite
Norton/Symantec
Palm/Treo
Panasonic
Phillips Magnavox
Pioneer
PlayStation 2
PlayStation 3
PSP
Radio Shack
RCA
Samsung
SanDisk
Sanyo
Sharp
Sony
Sony PlayStation
Super Mario Brothers (the video game)
Tivo

Sports and hobbies

24 Hour Fitness
Atlanta Braves
Bally Fitness
Boston Celtics
Boston Red Sox
Chicago Cubs
Curves
LA Lakers
MLB (Major Baseball League)
NASCAR
NBA
NCAA
New England Patriots
NFL (National Football League)
NHL (National Hockey League)
NY Giants
NY Jets
NY Yankees
Pittsburgh Steelers
WWE
YMCA

Telecommunications

AOL
AT&T
Blackberry
Boost Mobile
BrightHouse
Cellular One
Charter Communications
Comcast
Cox
Cricket
Dish Network
iPhone
Motorola
Motorola Razr
Nokia
Qwest
Road Runner
Sidekick
Sprint
TMobile
Tracfone
US Cellular
Verizon
Virgin Mobile
Vonage

Travel services

AirTran
Alamo
Alaska Air
American Airlines
Amtrak
Avis Rental Cars
Best Western
British Airways
Budget Car Rental
Carnival Cruise Lines
Comfort Inn
Continental Airlines
Days Inn
Delta Airlines
Enterprise Car Rental
Expedia
Frontier Airlines
Hampton Inn
Hertz
Hilton
Holiday Inn
Hyatt
Jet Blue
Marriott
Orbitz
Priceline.com
Princess Cruises
Royal Caribbean Cruises
Sheraton Hotels
Southwest Airlines
Spirit Airlines
Travelocity
United Airlines
US Air

Appendix 2: Annotated Questionnaire

The questionnaire is a dynamic Internet survey done on a representative sample of the US population. Respondents were recruited through a sample provider and then referred to a password protected website to do the survey. The survey process was complex for two main reasons: First, the system had to dynamically allocate brands to respondents so respondents would answer on brands they are familiar with on one hand, and so enough responses were obtained for each brand on the other hand. Second, due to the large number and diversity of brands on the list, some questions had to be modified for particular categories. For example the item: “iPhone is not very visible in my environment” was adjusted to “Boston Celtics fans are not very visible in my environment”, for sports teams. We bring here a generic wording. The survey started with screening questions on gender, age, education, ethnicity, zip code. If we still needed people in their profile, they continued to the next stage.

(TEXT) We are conducting a survey in order to understand the thoughts and perceptions of customers on the brands and products they use. You will be asked questions on several brands, please answer all the questions for each brand.

Category familiarity

Q1. To what extent are you familiar with each of the following types of products and services? (A five- point scale with “1” for “unfamiliar” and “5” for “familiar”. Order of categories is randomized).

Category list: Beverages, Beauty products, Cars, Children’s Products, Clothing products, Department stores, Financial services, Food and dining, Health products and services, Household Products (cleaning ingredients etc.), Media and entertainment, Sports and hobbies, Technology products and stores, Telecommunications, Home design and decoration, Travel services

(SELECT UP TO THREE CATEGORIES PUNCHED “4” OR “5” AT Q1 BASED ON LEAST-FILLED QUOTA TO ASK FOR Q2-Q4.)

Category involvement

Q2. For each of the following product types, how important to you is the decision of which specific product to choose within that type? (A five-point scale with “1” for “very unimportant” and “5” for “very important”. Order of categories is randomized. Use same category list at Q1).

Q3. How much thought is required to choose a specific product of the following types? (A 5-point scale with a “1” for “decision requires little thought” and a “5” for “decision requires a lot of thought”. Order of categories is randomized. Use same category list at Q1).

Q4. How much risk do you take when choosing a specific product of the following types? (Insert a 5-point scale with a “1” for “little to lose if i choose the wrong product” and a “5” for “a lot to lose if i choose the wrong product”. Order of categories is randomized. Use same category list at Q1).

(SELECT ONE OF THE SELECTED CATEGORIES TO BE USED FOR THE BRAND SECTIONS BASED ON LEAST-FILLED QUOTA.)

Brand familiarity

Q5. To what extent are you familiar with the following brands? (Insert a 5-point scale with a “1” for “unfamiliar” and a “5” for “familiar”. Select 20 brands from the selected category based on least-filled quota. Randomize brands).

Brand perceptions

(SELECT UP TO 10 BRANDS FOR WHICH THE RESPONDENT SELECTED EITHER “4” OR “5” AT Q5 BASED ON LEAST-FILLED QUOTA. NEED TO ACHIEVE A MINIMUM OF 30 RATINGS PER BRAND.)

(ASK Q6-Q7 AS A SERIES FOR EACH BRAND BEFORE PROCEEDING TO THE NEXT BRAND. STORE THE ORDER IN WHICH THE BRANDS ARE SHOWN TO THE RESPONDENT IN THE SERIES.)

Q7. To what extent do you agree with the following statements for (BRAND). (Insert a 5-point scale with a “1” for “totally disagree” and a “5” for “totally agree”. Randomize statements).

Complexity

[Q6_1] Getting used to (BRAND) requires a major learning effort .

[Q6_2] Getting used to (BRAND) takes a long time before one can fully understand the advantages. [Q6_3] The product concept of (BRAND) is difficult to evaluate and understand.

[Q6_4] Overall, I believe that using (BRAND) is easy.

[Q6_5] Using (BRAND) requires a lot of mental effort.

Visibility

[Q6_6] I have seen how another person is using (BRAND).

[Q6_7] In my environment, one sees (BRAND) quite often.

[Q6_8] (BRAND) is not very visible in my environment.

[Q6_9] I have had plenty of opportunities to see someone else using (BRAND).

[Q6_10] It is rare that I see someone else using (BRAND).

Perceived risk

[Q6_11] When using (BRAND) I am sure I will get what I expected.

[Q6_12] Using (BRAND) may cause me additional unplanned expenses [FOR THE MEDIA CATEGORY, USE THE FOLLOWING SENTENCE: I have frequently felt that watching (BRAND) was a waste of my time

[Q6_13] Using (BRAND) might cause me social embarrassment.

Q7. Now please think of (BRAND) as if it were a person. This may sound unusual, but think of the personality traits or human characteristics that come to mind when you think of (BRAND). (A 5-point scale with a “1” for “not at all descriptive” and a “5” for “extremely descriptive”. Randomize statements.)

Excitement

Daring
Trendy
Exciting
Spirited
Cool
Young
Imaginative
Unique
Up to date
Independent
Contemporary

(Competence. Measured but eventually not used for the analysis due to multicollinearity with other variables)

Reliable
Hard-working
Secure
Intelligent
Technical
Corporate
Successful
Leader
Confident

Demographics

To end the survey, the subscriber is asked a set of demographic questions on employment, occupation, income, marital status and children living at home.

Appendix 3: Markov Chain Monte Carlo Estimation

We estimate this model separately for online and offline WOM data using Markov Chain Monte Carlo posterior simulation. We break the sampling into four blocks with three blocks using the Metropolis-Hastings algorithm and one block using a conditionally conjugate draw. The procedure for the posterior simulation is a modification of the `rnegbinRw` routine in the R package, `bayesm`. Formally, the four sampling blocks are

1) Draw $\left(\beta^m, \left\{ \gamma_k^m \right\}_{k=1}^K \mid \alpha^m, \delta^m, \sigma_m^2, X_{ijk}^I \right)$ using a Metropolis-Hastings random walk algorithm.

We tune the variance of this random walk using the Hessian from the MLE that assumes the γ_k^m are fixed effects. This variance-covariance is scaled down to improve the efficiency of the sampler. In the final exploration of the posterior, the scaling factor was .025 in both channels.

- 2) Draw $\left(\alpha^m \mid \beta^m, \left\{ \gamma_k^m \right\}_{k=1}^K, X_{ijk}^I \right)$ using a Metropolis-Hastings random walk algorithm. We use the posterior sampling procedure as written in `rnegbinRw` contained in the `bayesm` R package.
- 3) Draw $\left(\delta^m, \sigma_m^{-2} \mid \left\{ \gamma_k^m \right\}_{k=1}^K \right)$ using conditionally conjugate draws from the multivariate normal and chi-square distributions, respectively. To accomplish these draws, we use `rmultireg` contained in the `bayesm` R package.
- 4) Draw $\left(X_{ijk}^I \mid \beta^m, \left\{ \gamma_k^m \right\}_{k=1}^K, \alpha^m, X_{ijk}^C \right)$ independently using a Metropolis-Hastings random walk algorithm. We use the variance of the X_{ijk}^C scaled to achieve efficiency. The final scaling was 3 for the online channel draws and .5 for the offline channel draws.

For the full model, this MCMC sampler was run for 100,000 iterations for three different initial seeds. Convergence was clearly achieved within a few thousand iterations as determined by visual inspection of the three chains (which all converged to the same stationary distribution). We discarded the first 10,000 iterations of each chain as burn-in and thinned at 10 for inference. This essential procedure for diagnosing convergence and for inference was used in estimating each of the submodels as well. All marginal likelihoods are calculated using the harmonic mean estimator.

Appendix 4: Moderation Analysis

We evaluate whether the type of good (Search, Experience, or Credence) influences the nature of the relationship between the brand characteristics and online and offline WOM. We do this by running the full model analysis with the addition of interactions between the type of good and our theoretical variables of interest.

Our dataset contains relatively few observations of credence goods (40 in the final sample). While this is a reasonable sample to estimate a main effect of *Credence*, it is too small to make statistical conclusions on moderation effects given that we have 14 parameters of interest (leaving less than 3 observations per variable for these interaction effects). For search goods, we have 127 observations, which gives us a large number for a main effect and about 9 observations per parameter for the interaction effects, which is a more reasonable number. Hence, we only present results relating to interactions with *Search* as compared to goods that are either credence or experience. We chose to group these two together since they seem to be more similar. The results are displayed in Table 1.

We focus first on the three novel characteristics introduced by our theoretical framework.

Energized-Differentiation. For the online data, both the main effect of *Energized-Differentiation* and the interaction with search are positive and significant (though slightly weaker than the five percent level). Offline, the effect is positive and significant for search goods but not for experience and credence goods.

Excitement. In both the online and offline data, the main effects of *Excitement* is significant and positive (as in our main results), and the interactions are not significant. This means that the association between *Excitement* and WOM does not depend on the type of good.

Complexity. The results for offline WOM are quite similar to those reported in Table 5 – *Complexity* is positive and significant. Furthermore, given that the interaction is not significant, we can conclude that the association between *Complexity* and WOM does not depend on the type of good. In the online data the association is also positive but insignificant for both the main effect and the interaction. We have witnessed the difference between the online and offline estimates also in our primary results reported in Table 5, though then the association for online was marginally significant and negative.

Of the remaining significant variables from Table 5 of the main paper, all but *Middle (VP)* for online and *Relevance* for offline are the same direction and maintain a similar level of significance in the main effects. While *Middle (VP)* is simply insignificant for online here, offline *Relevance* is still positive and significant for search goods but not for credence and experience ones. Hence, even in a model that accounts for type of good interactions, experience and credence goods largely follow the patterns discussed in the main text.

Further, for most variables, there are no significant interactions. Some interactions, however, are significant and may point to potential interesting avenues for research. For the online interactions, we find that *Relevance* has a significant positive coefficient, *Esteem* a significant negative one, and *Knowledge Factor* a marginally significant negative one. For the offline interactions, *Relevance*, as pointed above, has a significant positive interaction, *Esteem* and *Premium (VP)* significant negative ones, and *Age* significant positive one (indicating that while newer experience and credence goods received more WOM, the opposite holds for search goods).

While these interaction effects are complex and interesting, our main results are largely consistent with most of the data. Furthermore, these differences suggest that there is more nuanced heterogeneity in the effect of brand characteristics on WOM. Hence, we leave the exploration of such interactions for future research.

Appendix 5: Comparison of Submodel Estimates to Full Model Estimates

In this appendix we compare the results of the estimates from the submodels with those of the full model. We first note that the signs and significance of the coefficients are largely consistent with the full model, despite an obvious concern with omitted variable bias. That said, in a few cases these effects differ, likely as a result of the omitted variable *Satisfaction*. The estimates are presented in Tables 2 and 3.

Considering first the online results, we find that *Esteem*, *Visibility*, *Excitement*, *Age*, *Credence*, and *Knowledge Factor* all have the same sign as the full model results, and those variables that are significant remain significant while those that are not remain insignificant. *Differentiation*, *Premium (VP)*, and *Perceived Risk* in the same direction as the main results and in all but one case are significant. *Search* and *Middle (VP)* are both not significant, while they were significant in the full model. The most notable change is *Complexity*, which switches from negative and insignificant to positive and significant. We note, however, that this effect appears to result from

an omitted variable bias (i.e., as noted in Web Appendix 6.2, *Satisfaction* and *Complexity* are highly related and dropping *Satisfaction* leads to apparent omitted variable bias in the estimated effect of *Complexity*).

Turning to the offline results, we find that *Middle (VP)*, *Visibility*, *Search*, and *Knowledge Factor* all have the same sign as the full model results, and those variables that are significant remain significant, while those that are not remain insignificant. *Differentiation*, *Esteem*, *Excitement*, *Credence*, and *Complexity* all have the same sign and in all but one submodel are significant, while *Age* loses significance for two models and *Relevance* for all three. *Perceived Risk* is more significant in some cases. We note, however, that this again appears to result from omitted variables bias, which we discuss in the Appendix 6.2.

Despite decreased significance in some variables, overall, these results are quite consistent with those presented in the full model. In fact, it appears that including *Satisfaction* largely strengthens our ability to tease apart the role of some of the variables, including *Differentiation*, *Esteem*, *Relevance*, *Excitement*, and *Complexity* for offline and *Differentiation and Premium (VP)* for online. However, *Satisfaction* also appears to explain the role of *Complexity* in online WOM and *Perceived Risk* in offline WOM, which suggests we need *Satisfaction* in the model to get the correct significance levels.

Appendix 6: Robustness Checks

In this appendix, we conduct a series of robustness tests to examine whether our results hold up to four potential issues with our data: selection bias, multicollinearity, outliers, and missing data.

6.1 Selection Bias

Our sample is not a random sample – it includes the brands with the highest amount of WOM. This selection may lead to bias. In order to get a sense of the issue we can decrease the number of brands, estimate the model on a partial set of the brands, and examine the sensitivity of the results to it. In the rest of this subsection we describe this analysis and show that overall it appears that the main results indeed stand up. In particular, we consider the top 200, 450, 500, and 550 WOM brands, where the list of brands is specific to the channel (i.e., top 450 online WOM brands or top 450 offline WOM brands). The results of these analyses are presented in Table 4 for the online data and Table 5 for the offline data.

We turn first to the relative importance of the three drivers (i.e., social, functional and emotional). Note that the results for this task do not rely on the missing data model. The results indicate that the most important driver is consistently social for online mentions and emotional for offline mentions. The only exception is for the online top 200 dataset, where the differences in the log marginal likelihoods are quite small compared to the other models. In other words, the order of importance of these drivers in the two channels does not depend on the size of the sample. Thus, there appears to be no meaningful selection biases involved in the relative importance results.

We next turn to the three novel variables introduced by our theoretical framework – *Excitement*, *Differentiation*, and *Complexity*.

- *Excitement* is still significantly related to online and offline WOM, which we take as strong evidence supporting a high degree of robustness to selection bias.
- *Differentiation* retains significance and the expected direction in all cases online (though only marginal significance in the smallest dataset). In the offline analysis, *Differentiation* is in the expected direction and is significant in all but the smallest dataset. In the smallest dataset, it becomes insignificant. This suggests a high degree of robustness and no obvious selection bias.
- *Complexity*, which was marginally significant in the full online dataset, maintains the negative sign throughout, but with varying significance levels: it is significant in the 550 dataset and insignificant for the smaller datasets. For the offline dataset, *Complexity* is positive and significant for the 550 and 500 datasets, but insignificant and positive for the 450 and negative and insignificant for 200. Overall, these results suggest that *Complexity* may be less robust to the sample size, and that it may play a smaller role in the very top brands.

Finally, we present the influence on the remaining parameters.

Online. A glance at the results of the full model suggests no changes in the direction of the effect and only few changes in significance in the relevant variables. There are no meaningful changes in any of the following: *Relevance*, *Esteem*, *Middle (VP)*, *Premium (VP)*, *Satisfaction*, *Knowledge Factor*, and *Involvement*. *Perceived Risk* has one model with less significance, but the signs are consistent. However, *Age*, becomes more significant and *Visibility* less so for the Top 200 dataset. Thus, it seems that the selection is not a major issue for our online data and results.

Offline. For the offline setting, all of the datasets have very similar estimates in terms of the direction of effect, while the significance is somewhat lower in some cases. Of the significant variables from the full model, *Differentiation*, *Esteem*, *Middle (VP)*, *Premium (VP)*, *Visibility*, *Excitement*, *Satisfaction*, *Search*, *Credence*, *Knowledge Factor*, and *Perceived Risk* all have the same signs and significance in the Top 450, Top 500, and Top 550, but in some cases the significance changes in the Top 200. In addition, *Relevance* and *Age* are not significant in at least one of the datasets. For the Top 200, we find that several variables seem to strengthen or weaken. *Visibility* weakens to insignificance, while *Middle (VP)*, *Premium (VP)*, and *Age* strengthen and become significant and potentially stronger. Nonetheless, on most variables until reaching the smallest datasets both online and offline, we find very similar directional results, but with a few cases of reduced significance.

Overall, we take these results for the relative importance and individual variable effects as supportive of the selection biases not being too severe. There are a couple of cases of potential issues (e.g. *Complexity*), but these cases suggest that while some caveats need to be made with regard to selection biases, the magnitude of these potential problems appears fairly limited.

Another takeaway from this exercise is that the relationship between brand characteristics and WOM is a bit different for the top 200 brands – for example, *Age* plays a much more important role for this subset of brands both online and offline, while the role of *Visibility* weakens.

6.2 Multicollinearity

Since our study accounts for many factors, there is potential for multicollinearity. Such a problem results in inflated standard errors (or in our case inflated credible interval sizes) arising from highly correlated factors in the model.

This subsection is split into two parts. In the first, we use three methods (variance inflation factors, principal factor analysis, and partial correlation analysis) in order to identify the variables that might be involved in multicollinearity. In the second, we examine whether the estimates of these variables are indeed affected by this issue.

Step 1: Identify potentially problematic variables.

Variance inflation factors. To evaluate this potential problem we first examine the Variance Inflation Factors (VIFs) for each of the brand characteristics (see Table 6). To do so, we include all variables (except for *Satisfaction* and *Involvement*) and the fixed effect for the category. *Satisfaction* was excluded due to its large amount of missing data, and *Involvement* since it is perfectly collinear with the category fixed effects. Based on these VIFs, only *Relevance* has a score higher than the traditional rule of 10. We can narrow the potential problem even further to only the online channel, since the effect of *Relevance* in the offline estimation is significant (i.e., the credible interval is not too large). As pointed above, we will soon check whether the estimate of the potentially problematic variable (i.e. *Relevance*) is indeed problematic.

Principal component analysis. We present a principal component analysis and examine which of our variables have large portions of their variances explained by the same factors (Table 7). We have bolded the cases in which one factor explains a high portion of variance for more than one variable. This analysis indicates that there are several variables that have some meaningful overlap in factors such as: *Complexity* and *Perceived Risk* on Factor 33 (0.3 and 0.36), *Relevance* and *Esteem* on Factor 31 (0.63 and 0.35), and *Visibility* and *Complexity* on Factor 32 (0.46 and 0.57). Also, not surprisingly given that they are dummy variables, *Middle (VP)* and *Premium (VP)* load on largely the same factors. We summarize our conclusions from these findings after the *partial correlation analysis*.

A partial correlation analysis. This analysis (Table 8) reveals the correlations between our conceptual variables after removing the joint correlation with (a) our control variables (*Interbrand Top 100*, *Usage*, *Mixed (PS)*, *Service (PS)*, and *Internet Brand*) and (b) the category effects. We remove this variation in order to avoid allocating it to the theoretical variables. This analysis reveals very similar results to those in the principal component analysis. We find very few pairs of variables have troubling correlations. Using a conservative rule of 0.5, we identify the following pairs: *Middle/Premium (VP)*, *Complexity/Perceived Risk*, and *Esteem/Relevance*.

Based on the latter two analyses, in all but two cases, the most highly related variables come from the same drivers (e.g., both *Esteem* and *Relevance* are in the Social driver). The exceptions are the *Complexity/Visibility* pair, which includes one variable from each of the Social and Functional drivers, and the *Complexity/Perceived Risk* pair, which includes one variable from the Hybrid drivers and one from the Functional drivers. Since our relative importance results are

consistent both with and without the Hybrid drivers, this latter case cannot explain our ordering. As a result, multicollinearity is unlikely to influence the relative importance results.

Step 2: Are the potentially problematic variables really problematic?

We are now ready to focus on some of the key pairs identified above, to evaluate whether multicollinearity is affecting any of the individual coefficients from our main results. These pairs are *Complexity/Perceived Risk*, *Complexity/Visibility*, and *Esteem/Relevance*. Also, since we did not evaluate *Satisfaction* in the previous correlational analyses due to the missing data, we will consider it here. Based on the correlation table in the main text (Table 4), we will consider the pairs *Satisfaction/Complexity* and *Satisfaction/Perceived Risk*.

Since multicollinearity lowers the significance of the estimates due to inflated standard errors, only non-significant coefficients are a concern to us. Of the above-mentioned variables, only three are not significant--*Perceived Risk* offline, *Complexity* online (it is marginally significant), and *Relevance* online. For each of these variables, we drop its paired variable (e.g. dropping *Esteem* for *Relevance*) and evaluate whether the credible interval (CI) decreases in size. If such a change is large enough, this is a likely signal of multicollinearity.

Tables 9 and 10 present the multicollinearity tests for *Complexity* online, *Relevance* online and *Perceived Risk* offline with the first column in each table showing the full model results.

- *Perceived Risk* offline (dropping *Complexity* or *Satisfaction*). The 95% CI range of the full model is 0.58, compared with 0.51 (<20% decrease) without *Complexity* and 0.68 (an increase due to model differences) without *Satisfaction*. These changes in credible interval size are relatively small, but because of a large increase in the mean, both effects are now significant. The posterior mean for the coefficient of *Perceived Risk* when *Complexity* or *Satisfaction* are dropped has the same direction as in the full model, but increases by a factor of 10 in magnitude. This suggests an omitted variable issue when we do not account for *Complexity* and *Satisfaction* and that we need to control for them in order to get the correct magnitude of effect for *Perceived Risk*. At the same time, it also means that multicollinearity is not a problem here.
- *Relevance* online (dropping *Esteem*). The 95% CI range in the full model is 0.68 compared with 0.54 (<20% decrease) when dropping *Esteem*. This is a relatively small decrease. The estimate itself switches from negative to positive, but in both cases, the estimate is insignificant. In other words, it seems that including or not including *Esteem* seems to have relatively little impact on the significance of *Relevance*.
- *Complexity* online (dropping *Perceived Risk*, *Visibility*, or *Satisfaction*). The 95% CI range for the full model is 1.03, compared with 0.94 (<20% decrease) without *Visibility*, 0.83 (about a 20% increase) without *Perceived Risk*, and 0.88 (<20% decrease) without *Satisfaction*. These changes in credible interval size are relatively moderate. However, in each case the mean shifts fairly dramatically. Dropping *Visibility* decreases the mean making the parameter more significant; dropping *Perceived Risk* increases the mean making it closer to zero; and dropping *Satisfaction* increases the mean so dramatically that *Complexity* reverses signs and becomes significant in the opposite direction. Thus, without accounting for these variables, the magnitude of the effect can change dramatically,

suggesting that while multicollinearity appears to be a minor issue, without controlling for these correlated variables we are introducing a meaningful omitted variable bias.

Of course in these exercises additional variables (other than our focal ones) are also affected. After careful inspection, these changes do not appear to be due to multicollinearity (i.e., the credible intervals do not change much in size). However, similar to our three variables, without controlling for the full set of variables, the effects can be different, suggesting again that our full set of variables helps to guard against omitted variable biases.

Overall, we conclude that there is little evidence of troubling multicollinearity. Perhaps more telling than any issue from multicollinearity is that we find (in almost every case we examined) the potential for omitted variable bias (if the full set of variables were not included). Thus, including our full set of variables is a valuable aspect of our analysis rather than a problematic source of insignificant results.

6. 3 Missing Data

Our full model accounts for the missing data in *Satisfaction*, by using a Bayesian approach that is likelihood based and naturally accommodates multiple imputation. This approach treats the missing data as (unobserved) parameters in a Bayesian sense, so the missing data are drawn in the MCMC sampling routine. The MCMC chain generates multiple imputations (samples) for the missing data that are a function of all of the observed data (i.e., it does not depend only on *Satisfaction* but also on the other variables in the model and their correlation with the observed values of *Satisfaction*) as well as our prior (which follows the observed distribution of *Satisfaction*). In other words, though "random" in a distributional sense, the distribution of the missing data accounts for the relationships between the missing data and the observed data in a statistically and logically desirable manner. This full model approach is consistent with the suggested approach of Shafer and Graham (2002), i.e., a likelihood-based missing data analysis.

Before we proceed it is important to note that our analysis of the relative importance of the three drivers does not use the missing data model, and thus could not be driven by the missing data model specification. That said, one might be concerned that the parameter estimates of the effects of brand characteristics might be driven by it. To examine the robustness of our full model results, we consider two different sets of analyses, both displayed in Table 11.

1. *Case Deletion*: We conduct an analysis using only the observations that contain data on all variables (i.e., not imputing the missing data). That is, we use a case deletion approach to solving the missing data problem. This results in a much smaller data set with sample size of only 201 brands.

The results for the online data (Table 11) indicate the same directional effects but with generally reduced significance. We still have significant and directionally consistent results for *Middle (VP)*, *Visibility*, *Excitement*, *Satisfaction (marginally significant)*, *Knowledge Factor*, *Perceived Risk*, *Interbrand Top 100*, and *Service (PS)*; also multiple variables that were not significant remain so. Some of the online parameters are the same direction, but no longer significant-- *Differentiation*, *Esteem*, *Premium (VP)*, and *Complexity*. That said, taken as a whole, these findings suggest that even with the much smaller dataset, we achieve very similar results for the online data.

For offline, we find (Table 11) the directions of effects are the same for all but a few, insignificant variables (from Table 5 in the main text), but usually (not surprisingly) with lower significance. We still have significant and directionally consistent results for *Esteem* (marginally significant), *Visibility*, *Excitement*, *Satisfaction*, *Credence* (marginally significant), *Complexity* (marginally significant), *Knowledge Factor*, *Interbrand Top 100*, and *Service (PS)*, while various other variables (e.g. *Perceived Risk*) remain insignificant. We find that while having the same direction of effect, the variables become insignificant in some cases--*Differentiation*, *Relevance*, and *Age*. Thus, the main impact of case deletion as compared to our full model is reduced significance, which is not too surprising for such a dramatic decrease in sample size.

2. *Single imputation of conditional means*: We next use a first stage linear regression of the observed *Satisfaction* on all of the other variables in the model. We take the observed data when *Satisfaction* is missing to predict the missing cases for use in the second stage main analysis.

The results for online data are largely consistent with the full model. *Differentiation*, *Esteem*, *Premium (VP)*, *Visibility*, *Excitement*, *Complexity*, *Knowledge Factor*, *Perceived Risk*, and *Interbrand Top 100* all have the same sign and are statistically significant as before. While *Middle (VP)*, *Search*, *Satisfaction*, *Usage*, and *Service (PS)* become insignificant, *Relevance* and *Internet Brand* become significant. For offline, the results are also largely consistent with the full model. *Differentiation*, *Relevance*, *Esteem*, *Visibility*, *Excitement*, *Satisfaction*, *Age*, *Credence*, *Complexity*, *Knowledge Factor*, *Interbrand Top 100*, *Usage*, and *Service (PS)* are all directionally consistent and statistically significant. Of the significant effects reported in Table 5, only *Internet Brand* is no longer significant, while some others decrease in significance. Overall, while these single imputation results are a bit weaker than our full model results, they appear generally consistent.

Summarizing these robustness checks for the missing data model we use, we find that in general our main results are robust to a number of alternative missing data approaches. More importantly, the missing data model we employ is consistent with the recommended approach in the literature--accounting for the missing data through a likelihood-based analysis that imputes the missing data multiple times via a Bayesian MCMC approach (Shafer and Graham 2002). Furthermore, compared to the case deletion approach it is using a much larger sample for estimating the model and compared with the single imputation it is accounting for the errors in the missing data model.

6.4 Outliers

To examine outliers we ran log-linear regressions with fixed effects in the category (dropping *Involvement*, which is perfectly collinear with the category fixed effects and *Satisfaction*, which has a large number of missing observations). We identified cases having standardized residuals with absolute values greater than 3. The offline WOM regression had one case with the standardized residual of -3.58 (Harley Davidson). For online, we identified five cases with standardized residuals of -3.23 (Good Year Tires), -3.14 (Swansons), -3.16 (Whirlpool), -3.13 (Avis Rental Car), and -3.01 (Curves). Since all of these are moderately "outlying" and there are so few compared to our sample of over 600 observations, we suggest these outliers can have little effect on the results. To confirm this, we ran the models dropping these observations. As expected, there was no meaningful impact on our results.

7. Principal Component Analysis to form Knowledge Factor

Because both *Knowledge* and *Familiarity* represent the same underlying variable (Knowledge) and because these two variables are highly correlated (correlation of 0.8), we use principle component analysis to construct a single variables, *Knowledge Factor*. The loadings and proportion of total variance are displayed in Table 12. We see that 91% of the variation in the two variables is captured by the first component and that both variables enter positively. This ensures that the factor has the intended meaning. We use this component as the *Knowledge Factor* throughout the analysis in the paper and the appendix.

Table 1: Main effects and interactions (with *Search*)

Variable	Online		Offline	
	Main Effects	Search Interactions	Main Effects	Search Interactions
Differentiation	0.90* (-0.09, 1.83)	2.41* (-0.58, 4.82)	-0.09 (-0.56, 0.32)	3.64** (1.76, 6)
Esteem	1.70** (0.93, 2.36)	-2.36** (-4.02, -0.55)	0.52** (0.16, 0.88)	-1.72** (-3.43, -0.65)
Middle (VP)	-0.03 (-0.3, 0.23)	0.48 (-0.13, 1.23)	0.05 (-0.07, 0.16)	-0.20 (-0.5, 0.14)
Premium (VP)	0.29* (-0.02, 0.58)	0.14 (-0.55, 1.27)	-0.01 (-0.15, 0.14)	-0.44** (-0.9, -0.06)
Relevance	-0.63** (-0.95, -0.27)	1.44** (0.44, 2.65)	0.05 (-0.18, 0.24)	1.03** (0.3, 1.59)
Visibility	0.97** (0.56, 1.3)	0.11 (-1.3, 1.2)	0.80** (0.6, 1)	0.42 (-0.52, 1.19)
Excitement	0.58** (0.25, 0.93)	0.38 (-0.4, 1.38)	0.47** (0.27, 0.65)	-0.53 (-1.04, 0.08)
Satisfaction	4.25** (2.24, 6.55)	-4.10 (-17.88, 9.34)	5.02** (3.29, 6.22)	-3.65 (-12.8, 4.77)
Satisfaction^2	-2.54** (-3.83, -1.39)	0.34 (-4.04, 4.92)	-2.83** (-3.52, -1.88)	0.56 (-2.15, 3.45)
Age	-0.04 (-0.33, 0.26)	0.16 (-0.69, 1.13)	-0.22** (-0.35, -0.06)	0.62** (0.25, 0.97)
Search	-2.18 (-12.98, 9.73)		2.03 (-5.5, 9.98)	
Credence	-0.19 (-0.61, 0.21)		-0.43** (-0.66, -0.15)	
Complexity	0.41 (-0.26, 0.99)	0.66 (-0.71, 2.16)	0.25** (0.01, 0.56)	-0.55 (-1.75, 0.67)
Knowledge Factor	0.44** (0.26, 0.59)	-0.46* (-0.95, 0.08)	0.27** (0.17, 0.39)	0.16 (-0.17, 0.48)
Perceived Risk	0.56* (-0.15, 1.47)	0.00 (-2.54, 1.73)	0.10 (-0.21, 0.4)	0.24 (-1.16, 1.55)
Involvement	-0.10 (-1.49, 1.28)		0.08 (-0.95, 1.12)	
Category Average	5.56** (0.02, 11.11)		0.65 (-3.37, 4.67)	
Interbrand_top_100	0.78** (0.52, 1.06)	0.16 (-0.6, 0.85)	0.32** (0.18, 0.46)	0.25 (-0.37, 1.12)
Usage	-0.47 (-1.47, 0.63)	1.95 (-2.03, 5.5)	-0.41 (-1.03, 0.28)	-1.79 (-4.16, 0.41)
Mixed (PS)#	-0.04 (-0.56, 0.48)		-0.09 (-0.47, 0.27)	
Service (PS)	0.24 (-0.09, 0.55)	-0.36 (-1.17, 0.58)	0.46** (0.3, 0.61)	0.05 (-0.59, 0.75)
Internet Brand	0.80** (0.09, 1.46)	-1.39 (-3.5, 1.52)	-0.57** (-0.83, -0.33)	0.19 (-1.07, 1.46)
Dispersion	1.14** (1, 1.27)		4.15** (3.28, 5.05)	

** - indicates the 95% CI does not cover 0; * - indicates the 90% CI does not cover 0.

- insufficient observations of Search goods were available to identify the interactions with Mixed (PS)

Table 2: Estimation results with the hybrid variables

	Online						Offline					
	Social	Emotional	Functional	S & E	F & E	S & F	Social	Emotional	Functional	S & E	F & E	S & F
Differentiation	1.44**			0.61		1.54**	0.55**			0.41		0.42**
Esteem	1.48**			1.71**		1.14**	0.46**			0.49**		0.26
Middle (VP)	0.07			0.09		0.02	0.03			0.03		0.03
Premium (VP)	0.30**			0.27**		0.23*	-0.13			-0.14		-0.17**
Relevance	-0.52**			-0.48**		-0.55**	0.13			0.14		0.13
Visibility	1.42**			1.28**		1.31**	1.27**			1.25**		1.22**
Excitement		1.09**		0.60**	1.11**			0.37**		0.09	0.45**	
Satisfaction												
Satisfaction^2												
Age			-0.16		0.13	-0.04			-0.14		-0.02	-0.15**
Search			0.21		0.07	0.06			0.09		0.06	0.05
Credence			-0.17		-0.31	-0.1			-0.26		-0.30*	-0.25**
Complexity			0.71**		0.46**	0.53**			0.31**		0.23	0.52**
Knowledge Factor			0.69**		0.64**	0.46**			0.49**		0.50**	0.37**
Perceived Risk	1.29**	0.85**	0.39	1.42**	0.75**	0.93**	0.69**	0.18	-0.18	0.74**	0.01	0.31**
Involvement	0.07	0.13	0.87	-0.11	0.15	0.44	0.09	0.31	0.76	-0.07	0.12	0.4
Category Average	4.29	9.70**	2.74	5.64**	2.31	4.25	-0.32	4.68**	0.95	0.79	-0.74	2.59
Interbrand_top_100	1.00**	1.35**	1.11**	1.00**	1.10**	0.93**	0.41**	0.78**	0.61**	0.40**	0.56**	0.30**
Usage	0.79	1.99**	-0.95**	0.77	-0.39	-0.47	0.72**	2.74**	0.46	0.72**	0.50*	-0.47
Mixed (PS)	-0.12	0.03	0.24	-0.05	0.08	-0.11	-0.03	0.04	0.13	-0.04	0.02	-0.004
Service (PS)	0.18	0.33**	0.17	0.23	0.40**	0.2	0.49**	0.53**	0.61**	0.49**	0.61**	0.51**
Internet Brand	0.81**	1.06**	1.13**	0.86**	1.16**	0.80**	-0.21	-0.52**	-0.44**	-0.2	-0.38**	-0.22
Dispersion	0.93**	0.81**	0.81**	0.94**	0.87**	0.97**	2.22**	1.75**	1.89**	2.22**	1.95**	2.4**

** - 95% CI does not cover 0; * - 90% CI does not cover 0

Table 3: Estimation results without the hybrid variables

	Online						Offline					
	Social	Emotional	Functional	S & E	F & E	S & F	Social	Emotional	Functional	S & E	F & E	S & F
Differentiation	1.60**			1.00**		1.58**	0.51**			0.45		0.41*
Esteem	1.50**			1.69**		1.20**	0.57**			0.55**		0.29
Middle (VP)	0.09			0.11		0.02	0.02			0.02		0.02
Premium (VP)	0.34**			0.29*		0.21	-0.13			-0.13		-0.18**
Relevance	-0.77**			-0.78**		-0.70**	-0.04			-0.03		0.08
Visibility	1.24**			1.10**		1.23**	1.18**			1.18**		1.21**
Excitement		1.07**		0.46**	1.07**			0.35**		0.02	0.45**	
Satisfaction												
Satisfaction^2												
Age			-0.18		0.06	-0.09			-0.13		-0.03	-0.15*
Search			0.19		0.05	0.08			0.1		0.06	0.05
Credence			-0.24		-0.34	-0.16			-0.23		-0.32**	-0.27**
Complexity			0.90**		0.85**	0.97**			0.22*		0.24	0.66**
Knowledge Factor			0.70**		0.66**	0.49**			0.48**		0.50**	0.39**
Perceived Risk												
Involvement												
Category Average	7.95**	10.59**	7.81**	6.66**	7.01**	6.79**	1.90**	5.67**	4.27**	1.02**	1.88**	4.11**
Interbrand_top_100	1.07**	1.37**	1.11**	1.04**	1.08**	0.91**	0.43**	0.78**	0.61**	0.44**	0.57**	0.30**
Usage	0.70	1.27**	-1.13**	0.76	-0.72*	-0.39	0.67**	2.57**	0.51	0.68**	0.5	-0.56
Mixed (PS)	0.10	0.17	0.29	0.19	0.20	0.11	0.04	0.08	0.09	0.05	0.02	0.04
Service (PS)	0.43**	0.48**	0.19	0.49**	0.46**	0.32**	0.61**	0.56**	0.60**	0.60**	0.62**	0.53**
Internet Brand	0.95**	1.20**	1.15**	1.04**	1.22**	0.88**	-0.21	-0.51**	-0.44**	-0.2	-0.36*	-0.25
Dispersion	0.90**	0.80**	0.81**	0.91**	0.87**	0.95**	2.15**	1.75**	1.88**	2.15**	1.95**	2.39**

** - 95% CI does not cover 0; * - 90% CI does not cover 0

Table 4: Online Selection Bias Test

Online Selection Bias Tests (Posterior Mean and 95% CI)					
Variable	Full Data	Top 550	Top 500	Top 450	Top 200
Differentiation	1.78** (0.90, 2.65)	2.56** (1.36, 3.43)	2.56** (1.53, 3.57)	2.19** (1.16, 3.31)	1.57* (-0.08, 3.26)
Esteem	1.22** (0.66, 1.79)	1.26** (0.68, 1.88)	1.14** (0.54, 1.79)	0.98** (0.21, 1.73)	1.20** (0.15, 2.36)
Middle (VP)	0.50** (0.31, 0.69)	0.64** (0.43, 0.85)	0.57** (0.35, 0.79)	0.56** (0.31, 0.77)	1.32** (0.96, 1.69)
Premium (VP)	0.47** (0.19, 0.75)	0.64** (0.38, 0.91)	0.61** (0.30, 0.90)	0.66** (0.35, 0.98)	0.79** (0.38, 1.20)
Relevance	-0.26 (-0.62, 0.06)	-0.19 (-0.54, 0.16)	-0.14 (-0.53, 0.24)	-0.22 (-0.62, 0.19)	0.19 (-0.44, 0.88)
Visibility	0.92** (0.65, 1.17)	0.77** (0.42, 1.10)	0.72** (0.35, 1.06)	0.46** (0.01, 0.90)	0.42 (-0.20, 1.05)
Excitement	0.71** (0.39, 0.99)	0.60** (0.28, 0.90)	0.57** (0.23, 0.94)	0.76** (0.39, 1.12)	0.67** (0.15, 1.21)
Satisfaction	4.60* (-0.54, 9.75)	6.45** (0.57, 11.87)	5.94* (-0.93, 12.37)	6.63* (-0.04, 13.96)	3.32 (-2.83, 9.03)
Satisfaction^2	-3.56** (-5.25, -1.94)	-4.22** (-5.98, -2.33)	-4.06** (-6.21, -1.88)	-4.35** (-6.76, -2.18)	-3.40** (-5.52, -1.22)
Age	0.13 (-0.07, 0.37)	-0.03 (-0.29, 0.19)	-0.17 (-0.42, 0.09)	-0.19 (-0.46, 0.10)	-0.89** (-1.31, -0.48)
Search	-0.30** (-0.56, -0.06)	-0.37** (-0.64, -0.11)	-0.13 (-0.45, 0.18)	-0.12 (-0.48, 0.21)	-0.78** (-1.23, -0.27)
Credence	-0.01 (-0.37, 0.37)	0.06 (-0.34, 0.47)	0.03 (-0.38, 0.46)	0.11 (-0.30, 0.53)	0.66** (0.12, 1.27)
Complexity	-0.49* (-0.98, 0.05)	-0.70** (-1.29, -0.14)	-0.42 (-0.95, 0.10)	-0.17 (-0.88, 0.53)	-0.91 (-2.2, 0.23)
Knowledge Factor	0.49** (0.33, 0.65)	0.53** (0.37, 0.69)	0.59** (0.39, 0.75)	0.67** (0.49, 0.86)	0.61** (0.34, 0.85)
Perceived Risk	0.91** (0.30, 1.44)	0.95** (0.34, 1.64)	0.73** (0.17, 1.32)	0.54 (-0.23, 1.29)	1.56** (0.56, 2.45)
Involvement	-0.58 (-2.01, 0.85)	-0.54 (-1.93, 0.89)	-0.38 (-1.80, 1.04)	-0.45 (-1.89, 0.99)	-0.48 (-2.01, 1.03)
Category Average	8.36** (1.67, 15.37)	7.50** (0.76, 14.87)	7.46** (0.49, 14.56)	7.99** (0.81, 14.98)	10.2** (3.16, 17.74)
Interbrand_top_100	0.95** (0.74, 1.17)	0.85** (0.62, 1.11)	0.87** (0.59, 1.13)	0.89** (0.59, 1.17)	0.55** (0.22, 0.94)
Usage	-1.07** (-2.22, -0.13)	-1.21** (-2.29, -0.17)	-1.14** (-2.32, -0.08)	-1.23* (-2.51, 0.02)	-1.63* (-3.17, 0.03)
Mixed (PS)	-0.36 (-0.99, 0.34)	-0.49 (-1.18, 0.28)	-0.51 (-1.19, 0.25)	-0.47 (-1.22, 0.31)	0.21 (-2.49, 2.74)
Service (PS)	0.54** (0.23, 0.79)	0.60** (0.31, 0.87)	0.49** (0.23, 0.80)	0.30* (-0.01, 0.71)	0.01 (-0.52, 0.43)
Internet Brand	0.31 (-0.08, 0.70)	0.18 (-0.26, 0.65)	0.35 (-0.11, 0.82)	0.02 (-0.44, 0.50)	0.26 (-0.57, 1.13)
Dispersion	3.12** (2.75, 3.51)	3.15** (2.74, 3.57)	2.99** (2.59, 3.41)	3.09** (2.64, 3.55)	5.53** (4.38, 6.73)
Log Marginal Likelihoods / Importance Rank					
Social	-8394 / 1	-7535 / 1	-6883 / 1	-6168 / 1	-2786 / 3
Emotional	-8456 / 3	-7596 / 3	-6921 / 3	-6199 / 3	-2785 / 2
Functional	-8439 / 2	-7564 / 2	-6896 / 2	-6180 / 2	-2779 / 1

** = 95% CI does not cover 0; * = 90% CI does not cover 0.

Table 5: Offline Selection Bias Test

Offline Selection Bias Tests (Posterior Mean and 95% CI)					
Variable	Full Data	Top 550	Top 500	Top 450	Top 200
Differentiation	0.62** (0.16, 1.12)	0.85** (0.29, 1.41)	0.89** (0.31, 1.45)	0.77** (0.15, 1.48)	-0.36 (-1.42, 0.65)
Esteem	0.52** (0.22, 0.81)	0.30* (-0.04, 0.61)	0.53** (0.17, 0.90)	0.54** (0.11, 0.94)	1.26** (0.37, 2.23)
Middle (VP)	0.01 (-0.09, 0.10)	0.04 (-0.07, 0.14)	0.02 (-0.11, 0.13)	-0.02 (-0.14, 0.12)	0.54** (0.33, 0.76)
Premium (VP)	-0.07 (-0.21, 0.06)	-0.06 (-0.21, 0.10)	-0.20* (-0.36, -0.02)	-0.17* (-0.36, 0.01)	0.32** (0.09, 0.57)
Relevance	0.29** (0.10, 0.47)	0.43** (0.25, 0.63)	0.31** (0.09, 0.51)	0.17 (-0.06, 0.41)	0.70** (0.22, 1.15)
Visibility	0.72** (0.53, 0.91)	0.60** (0.39, 0.79)	0.57** (0.34, 0.78)	0.55** (0.33, 0.78)	0.19 (-0.28, 0.54)
Excitement	0.44** (0.27, 0.60)	0.37** (0.20, 0.55)	0.36** (0.16, 0.55)	0.47** (0.27, 0.68)	0.70** (0.40, 0.99)
Satisfaction	5.59** (3.21, 8.17)	5.05** (2.72, 7.72)	7.12** (3.65, 11.03)	7.24** (3.93, 10.85)	3.50 (-3.64, 9.21)
Satisfaction^2	-3.10** (-3.93, -2.30)	-2.90** (-3.80, -2.12)	-3.58** (-4.88, -2.43)	-3.61** (-4.76, -2.53)	2.91** (-4.72, -0.78)
Age	-0.17** (-0.29, -0.05)	-0.12* (-0.26, 0.02)	-0.14* (-0.28, 0.01)	-0.11 (-0.28, 0.06)	-0.66** (-0.99, -0.31)
Search	0.04 (-0.11, 0.27)	0.10 (-0.08, 0.26)	0.06 (-0.12, 0.26)	0.07 (-0.13, 0.29)	-0.32* (-0.63, 0.08)
Credence	-0.60** (-0.81, -0.39)	-0.54** (-0.80, -0.28)	-0.56** (-0.80, -0.28)	-0.48** (-0.75, -0.20)	-0.92** (-1.29, -0.51)
Complexity	0.43** (0.09, 0.76)	0.34** (0.01, 0.69)	0.55** (0.11, 1.01)	0.33 (-0.05, 0.69)	-0.21 (-0.88, 0.46)
Knowledge Factor	0.46** (0.36, 0.56)	0.45** (0.36, 0.54)	0.45** (0.34, 0.54)	0.46** (0.34, 0.57)	0.10 (-0.09, 0.29)
Perceived Risk	0.03 (-0.26, 0.32)	0.09 (-0.26, 0.42)	-0.04 (-0.37, 0.30)	0.01 (-0.35, 0.39)	0.45 (-0.18, 1.06)
Involvement	0.13 (-1.01, 1.25)	0.13 (-1.04, 1.29)	0.07 (-1.05, 1.21)	0.13 (-1.01, 1.26)	-0.02 (-1.31, 1.30)
Category Average	-0.73 (-5.34, 4.01)	-0.10 (-4.85, 4.64)	-1.26 (-6.48, 3.80)	-1.25 (-6.32, 3.83)	2.89 (-3.83, 10.69)
Interbrand_top_100	0.26** (0.14, 0.39)	0.29** (0.16, 0.43)	0.27** (0.13, 0.41)	0.273** (0.12, 0.43)	0.23* (-0.02, 0.50)
Usage	-0.84** (-1.36, -0.28)	-0.93** (-1.48, -0.33)	-0.63** (-1.23, -0.02)	-0.45 (-1.12, 0.20)	-1.18* (-2.36, 0.10)
Mixed (PS)	0.27 (-0.15, 0.62)	0.34* (-0.04, 0.78)	0.32 (-0.09, 0.76)	0.30 (-0.20, 0.74)	0.15 (-2.49, 1.91)
Service (PS)	0.62** (0.49, 0.77)	0.60** (0.45, 0.75)	0.57** (0.41, 0.74)	0.52** (0.33, 0.72)	0.65** (0.41, 0.89)
Internet Brand	-0.30** (-0.49, -0.10)	-0.46** (-0.69, -0.22)	-0.60** (-0.84, -0.35)	-0.45** (-0.71, -0.19)	-1.06** (-1.51, -0.58)
Dispersion	8.36** (7.44, 9.34)	8.02** (7.10, 9.02)	8.23** (7.23, 9.30)	8.05** (7.03, 9.13)	12.33** (9.77, 15.11)
Log Marginal Likelihoods/Importance Rank					
Social	-5708 / 3	-5002 / 3	-4604 / 3	-4068 / 3	-2025 / 3
Emotional	-5316 / 1	-4737 / 1	-4302 / 1	-3932 / 1	-1873 / 1
Functional	-5409 / 2	-4885 / 2	-4404 / 2	-3993 / 2	-1975 / 2

** = 95% CI does not cover 0; * = 90% CI does not cover 0.

Table 6: Variance Inflation Factors (VIFs)

Variable	VIF
Differentiation	2.561
Relevance	10.355
Esteem	5.659
Middle (VP)	1.835
Premium (VP)	1.968
Visibility	2.163
Excitement	2.518
Age	1.67
Search	2.237
Credence	1.751
Complexity	5.224
KnowledgeFactor	3.909
PerceivedRisk	4.696
Interbrand Top 100	1.294
Usage	8.922
Mixed (PS)	1.151
Service (PS)	2.922
Internet Brand	1.216

Table 7: Loading Table of the Principal Component Analysis

Factor	Diff	Estm	Midl (VP)	Prm (VP)	Rel	Visi	Exct	Age	Srch	Cred	Cmpl	Know Fac	Per Risk	IB 100	Use	Mix	Sev	Inter-net
6	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.11
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.21	0.00	0.01
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.10	0.00	0.06	0.00	0.04
17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.19	0.01	0.34
18	0.00	0.00	0.04	0.12	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.01	0.00	0.03	0.00	0.13	0.00	0.01
19	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.16	0.00	0.21	0.00	0.18
20	0.00	0.00	0.02	0.07	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.07	0.00	0.41	0.00	0.03	0.00	0.05
21	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.67	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.03
22	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.05	0.48	0.01	0.00	0.04	0.00	0.00	0.00	0.02	0.00	0.00
23	0.00	0.00	0.43	0.32	0.00	0.00	0.00	0.05	0.12	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.17	0.02
24	0.01	0.00	0.05	0.05	0.00	0.00	0.00	0.60	0.05	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.03
25	0.00	0.00	0.30	0.18	0.00	0.00	0.00	0.02	0.09	0.03	0.00	0.03	0.00	0.01	0.01	0.01	0.15	0.00
26	0.00	0.00	0.07	0.06	0.00	0.00	0.00	0.00	0.13	0.01	0.00	0.18	0.00	0.00	0.03	0.01	0.49	0.00
27	0.01	0.33	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.01	0.30	0.01	0.00	0.00	0.01	0.05	0.01
28	0.44	0.05	0.00	0.05	0.00	0.00	0.00	0.02	0.01	0.01	0.00	0.06	0.00	0.05	0.14	0.01	0.01	0.00
29	0.25	0.18	0.01	0.01	0.00	0.00	0.00	0.13	0.03	0.01	0.04	0.10	0.03	0.01	0.23	0.00	0.04	0.01
30	0.06	0.01	0.00	0.00	0.02	0.15	0.04	0.01	0.00	0.00	0.08	0.02	0.04	0.00	0.31	0.00	0.00	0.02
31	0.14	0.35	0.01	0.01	0.63	0.03	0.21	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.23	0.00	0.01	0.01
32	0.04	0.02	0.00	0.00	0.09	0.24	0.25	0.00	0.00	0.03	0.30	0.03	0.36	0.00	0.03	0.01	0.01	0.01
33	0.04	0.02	0.00	0.00	0.01	0.46	0.28	0.00	0.00	0.01	0.57	0.00	0.22	0.01	0.00	0.00	0.00	0.01
34	0.01	0.04	0.00	0.00	0.24	0.12	0.21	0.03	0.00	0.00	0.00	0.02	0.34	0.01	0.00	0.00	0.01	0.00

Table 8: Partial Correlation Table

	Diff	Estm	Midl (VP)	Prem (VP)	Rel	Visib	Exct	Age	Srch	Cred	Cmpl	Know Fact	Per Risk
Differentiation	1.00	0.17	-0.16	0.27	0.02	0.09	0.47	-0.20	0.02	0.05	0.15	-0.03	0.05
Esteem	0.17	1.00	-0.04	0.15	0.64	0.17	-0.02	0.31	-0.16	0.08	0.01	0.35	-0.12
MiddleVP	-0.16	-0.04	1.00	-0.56	-0.04	0.08	-0.09	0.04	0.02	0.05	-0.10	0.00	-0.04
PremiumVP	0.27	0.15	-0.56	1.00	0.04	-0.01	0.17	0.02	0.05	-0.02	0.13	-0.02	-0.01
Relevance	0.02	0.64	-0.04	0.04	1.00	0.23	-0.06	0.23	-0.10	0.09	-0.13	0.22	-0.27
Visibility	0.09	0.17	0.08	-0.01	0.23	1.00	0.23	0.01	0.03	-0.04	-0.27	0.16	-0.30
Excitement	0.47	-0.02	-0.09	0.17	-0.06	0.23	1.00	-0.17	0.02	0.01	0.04	-0.04	-0.12
Age	-0.20	0.31	0.04	0.02	0.23	0.01	-0.17	1.00	-0.06	-0.01	-0.05	0.17	-0.10
Search	0.02	-0.16	0.02	0.05	-0.10	0.03	0.02	-0.06	1.00	-0.11	-0.01	-0.05	-0.04
Credence	0.05	0.08	0.05	-0.02	0.09	-0.04	0.01	-0.01	-0.11	1.00	0.16	-0.03	0.01
Complexity	0.15	0.01	-0.10	0.13	-0.13	-0.27	0.04	-0.05	-0.01	0.16	1.00	-0.13	0.60
KnowledgeFactor	-0.03	0.35	0.00	-0.02	0.22	0.16	-0.04	0.17	-0.05	-0.03	-0.13	1.00	-0.04
PerceivedRisk	0.05	-0.12	-0.04	-0.01	-0.27	-0.30	-0.12	-0.10	-0.04	0.01	0.60	-0.04	1.00

Table 9: The Model for Online Data with Dropped *Satisfaction*, *Esteem*, *Visibility*, or *Perceived Risk*

Online Dropped Variable Results (Posterior Mean and 95% CI)					
Variable	Drop None (Full Model)	Drop Satisfaction #	Drop Esteem	Drop Visibility	Drop Perceived Risk
Differentiation	1.78**(0.90, 2.65)	0.89* (-0.03, 1.73)	2.71** (1.93, 3.46)	1.97** (1.03, 2.80)	1.83** (0.82, 2.79)
Esteem	1.22** (0.66, 1.79)	1.30** (0.58, 1.99)		1.52** (0.97, 2.11)	1.23** (0.62, 1.80)
Middle (VP)	0.50** (0.31, 0.69)	0.08 (-0.13, 0.32)	0.57** (0.39, 0.74)	0.49** (0.30, 0.67)	0.51** (0.31, 0.69)
Premium (VP)	0.47** (0.19, 0.75)	0.24* (-0.03, 0.49)	0.55** (0.3157, 0.82)	0.37** (0.13, 0.59)	0.5** (0.25, 0.72)
Relevance	-0.26 (-0.62, 0.06)	-0.52** (-0.91,-0.17)	0.07 (-0.21, 0.33)	-0.28* (-0.59, 0.03)	-0.42** (-0.72, -0.06)
Visibility	0.92** (0.65, 1.17)	1.18** (0.85, 1.49)	1.01** (0.74, 1.29)		0.93** (0.67, 1.2)
Excitement	0.71** (0.39, 0.99)	0.52** (0.15, 0.87)	0.55** (0.28, 0.83)	0.94** (0.63, 1.24)	0.64** (0.33, 0.95)
Satisfaction	4.60* (-0.54, 9.75)		4.59* (-0.62, 9.36)	6.03** (0.35, 11.17)	4.66* (-0.79, 9.76)
Satisfaction^2	-3.56** (-5.25, -1.94)		-3.57** (-5.17, -1.92)		-3.65** (-5.34, -1.90)
Age	0.13 (-0.07, 0.37)	0.00 (-0.27, 0.26)	0.22* (-0.00, 0.43)	0.26** (0.04, 0.48)	0.11 (-0.12, 0.33)
Search	-0.30** (-0.56, -0.06)	0.04 (-0.28, 0.38)	-0.36** (-0.61, -0.10)	-0.30** (-0.56, -0.03)	-0.29** (-0.59, -0.02)
Credence	-0.01 (-0.37, 0.37)	-0.17 (-0.59, 0.30)	0.01 (-0.36, 0.40)	-0.10 (-0.46, 0.28)	-0.10 (-0.47, 0.24)
Complexity	-0.49* (-0.98, 0.05)	0.45** (0.01, 0.89)	-0.52* (-0.97, 0.00)	-0.92** (-1.41, -0.47)	0.01 (-0.39, 0.43)
Knowledge Factor	0.49** (0.33, 0.65)	0.43** (0.24, 0.59)	0.54** (0.37, 0.71)	0.54** (0.39, 0.70)	0.48** (0.33, 0.64)
Perceived Risk	0.91** (0.30, 1.44)	1.05** (0.47, 1.67)	0.98** (0.47, 1.53)	1.00** (0.47, 1.54)	
Involvement	-0.58 (-2.01, 0.85)	0.02 (-1.33, 1.38)	-0.59 (-1.98, 0.79)	-0.28 (-1.71, 1.14)	-0.46 (-1.92, 1.00)
Category Average	8.36** (1.67, 15.37)	3.77 (-1.67, 9.28)	8.03** (1.49, 14.80)	8.94** (1.88, 16.16)	9.31** (2.42, 16.36)
Interbrand_top_100	0.95** (0.74, 1.17)	0.92** (0.66, 1.19)	1.08** (0.86, 1.28)	1.06** (0.85, 1.29)	0.95** (0.73, 1.17)
Usage	-1.07** (-2.22, -0.13)	-0.35 (-1.59, 0.74)	-1.17** (-2.15, -0.14)	-0.82 (-1.87, 0.15)	-0.78 (-1.83, 0.18)
Mixed (PS)	-0.36 (-0.99, 0.34)	-0.05 (-0.49, 0.56)	-0.32 (-0.94, 0.44)	-0.36 (-0.98, 0.35)	-0.37 (-1.01, 0.33)
Service (PS)	0.54** (0.23, 0.79)	0.27* (-0.03, 0.57)	0.49** (0.18, 0.76)	0.59** (0.34, 0.86)	0.60** (0.3, 0.9)
Internet Brand	0.31 (-0.08, 0.70)	0.91** (0.37, 1.51)	0.18 (-0.24, 0.57)	0.17 (-0.24, 0.56)	0.22 (-0.15, 0.58)
Dispersion	3.12** (2.75, 3.51)	0.97** (0.88, 1.07)	3.04** (2.67, 3.42)	2.94** (2.59, 3.31)	3.04** (2.67, 3.42)

- Dropping Satisfaction also drops the missing data model

** - 95% Credible Interval does not cover zero; * - 90% Credible Interval does not cover zero

Table 10: The Model for Offline Data with Dropped *Satisfaction*, *Complexity*, or *Esteem*

Offline Dropped Variable Results (Posterior Mean and 95% CI)			
Variable	Drop None (Full Model)	Drop Satisfaction #	Drop Complexity
Differentiation	0.62** (0.16, 1.12)	0.30 (-0.22, 0.85)	0.85** (0.38, 1.33)
Esteem	0.52** (0.22, 0.81)	0.29 (-0.11, 0.65)	0.54** (0.24, 0.85)
Middle (VP)	0.01 (-0.09, 0.10)	0.03 (-0.10, 0.16)	0.02 (-0.07, 0.11)
Premium (VP)	-0.07 (-0.21, 0.06)	-0.17** (-0.32,-0.02)	-0.08 (-0.23, 0.06)
Relevance	0.29** (0.10, 0.47)	0.15 (-0.07, 0.37)	0.28** (0.07, 0.48)
Visibility	0.72** (0.53, 0.91)	1.19** (0.99, 1.39)	0.64** (0.47, 0.80)
Excitement	0.44** (0.27, 0.60)	0.13 (-0.04, 0.32)	0.42** (0.27, 0.57)
Satisfaction	5.59** (3.21, 8.17)		5.19** (2.98, 7.67)
Satisfaction^2	-3.10** (-3.93, -2.30)		-2.96** (-3.82, -2.21)
Age	-0.17** (-0.29, -0.05)	-0.13* (-0.29, 0.02)	-0.15** (-0.27, -0.02)
Search	0.04 (-0.11, 0.27)	0.04 (-0.13, 0.25)	0.03 (-0.14, 0.18)
Credence	-0.60** (-0.81, -0.39)	-0.25 (-0.53, 0.04)	-0.51** (-0.71, -0.31)
Complexity	0.43** (0.09, 0.76)	0.50** (0.20, 0.81)	
Knowledge Factor	0.46** (0.36, 0.56)	0.37** (0.28, 0.47)	0.45** (0.37, 0.55)
Perceived Risk	0.03 (-0.26, 0.32)	0.36** (0.01, 0.69)	0.28** (0.03, 0.54)
Involvement	0.13 (-1.01, 1.25)	-0.05 (-1.08, 1.03)	0.16 (-1.02, 1.30)
Category Average	-0.73 (-5.34, 4.01)	0.34 (-3.85, 4.42)	-0.01 (-4.71, 4.78)
Interbrand_top_100	0.26** (0.14, 0.39)	0.31** (0.14, 0.48)	0.28** (0.16, 0.41)
Usage	-0.84** (-1.36, -0.28)	-0.5 (-1.23, 0.13)	-0.90** (-1.48, -0.34)
Mixed (PS)	0.27 (-0.15, 0.62)	-0.01 (-0.34, 0.32)	0.27 (-0.17, 0.68)
Service (PS)	0.62** (0.49, 0.77)	0.50** (0.33, 0.68)	0.62** (0.49, 0.76)
Internet Brand	-0.30** (-0.49, -0.10)	-0.21 (-0.55, 0.14)	-0.27** (-0.47, -0.08)
Dispersion	8.36** (7.44, 9.34)	2.40** (2.15, 2.68)	8.30** (7.38, 9.25)

- Dropping Satisfaction also drops the missing data model

** - 95% Credible Interval does not cover zero; * - 90% Credible Interval does not cover zero

Table 11: Alternative methods for missing data

Variable	Online			Offline		
	Full Model#	Casewise Deletion	Single Imputation	Full Model#	Casewise Deletion	Single Imputation
Differentiation	***	1.77	1.22**	***	0.21	0.60**
Esteem	***	0.45	1.55**	***	0.53*	0.41*
Middle (VP)	***	0.61**	0.06	+	-0.04	0.05
Premium (VP)	***	0.22	0.34**	-	-0.22	-0.06
Visibility	***	1.10**	1.02**	***	0.76**	1.11**
Relevance	-	-0.2	-0.39**	***	0.13	0.25**
Excitement	***	0.69**	0.53**	***	0.46**	0.16*
Satisfaction	+	3.99	-0.32	***	3.14	0.92
Satisfaction^2	**	-2.00*	-0.64	**	-1.78**	-0.89**
Age	+	0.24	-0.01	**	-0.13	-0.14*
Search	**	-0.49	0.09	+	0.07	0.04
Credence	-	0.15	-0.23	**	-0.58*	-0.34**
Complexity	*	-0.71	0.38*	***	0.59*	0.48**
Knowledge Factor	***	0.60**	0.39**	***	0.52**	0.35**
Perceived Risk	***	1.91**	0.86**	+	0.09	0.16
Involvement	-	-0.92	0.14	+	0.1	-0.11
Category Average	***	4.69	5.80**	-	-0.17	1.53
Interbrand_top_100	***	0.94**	0.91**	***	0.30**	0.26**
Usage	**	-0.46	-0.95	**	-0.42	-0.98**
Mixed (PS)	-	0.26	-0.07	+	0.44	-0.05
Service (PS)	***	0.95**	0.08	***	0.74**	0.37**
Internet Brand	+	-0.18	0.95**	**	-0.31	-0.14

- Full Model results present only the direction of the effect statistical significance. For details see the main paper results.

** - 95% Credible Interval does not cover zero; * - 90% Credible Interval does not cover zero

Table 12: Principle Component Analysis Results

Element	Component 1	Component 2
Proportion of Variance	91%	9%
Loading on Familiarity	0.508	-0.861
Loading on Knowledge	0.861	0.508