ONLINE APPENDIX 1

MODEL SPECIFICATION AND ESTIMATION

The basic simultaneous model of Purchase and Interest involves an individual-specific baseline, a commercial-specific baseline, repressors specific to each dependent variable and Normally-distributed errors, correlated across viewing interest and purchase intent. We allow for correlation between the estimation of both regressions by specifying a full covariance matrix of the error term but restrict the variance of the Probit regression to 1 for identification reasons. Using Y^P and Y^I as the latent utilities associated with the intention to purchase and to view the ad fully, respectively, the model can be written hierarchically as:

$$\begin{pmatrix} Y_{ia}^{P} \\ Y_{ia}^{I} \end{pmatrix} = \begin{pmatrix} \mu_i^{(P)} + \alpha_a^{(P)} + B^{(P)} X^{*} \\ \mu_i^{(I)} + \alpha_c^{(I)} + B^{(I)} X \end{pmatrix} + \boldsymbol{\varepsilon}_{ict}, \qquad \begin{pmatrix} \mu_i^{(P)} \\ \mu_i^{(I)} \end{pmatrix} \sim \begin{matrix} N(W_i^{(3)}\Lambda^{(P)}, V_\lambda^{(P)}) \\ N(W_i^{(3)}\Lambda^{(I)}, V_\lambda^{(I)}) \end{matrix}$$

$$\begin{pmatrix} \alpha_c^{(P)} \\ \alpha_c^{(I)} \end{pmatrix} \sim \begin{matrix} N(Z_c^{(4)}\mathrm{K}^{(P)}, V_\kappa^{(P)}) \\ N(Z_c^{(4)}\mathrm{K}^{(I)}, V_\kappa^{(I)}) \end{matrix}$$

$$\boldsymbol{\varepsilon}_{ict} \sim N_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_I^2 = 1 & \sigma_{IP} \\ \sigma_{PI} & \sigma_P^2 \end{bmatrix} \right)$$

In the structure above, we assume additive separability of strictly individual and commercial baseline interest as well as purchase. In line with parsimony, we shrink the random effect parameters to aggregate covariate specific ones via Hierarchical Bayes. Lastly, we endogenize viewing interest in the purchase intent model, which is appended to X^* but not to X, which contains prior brand entertainment.

In order to estimate the unique observation equation via Gibbs sampling, let the full parameter set be $\Phi = \left\{ B^{(P)}, B^{(I)}, \mu_{i=1}, ..., \mu_{i=I}, \Lambda, V_\lambda, \alpha_{a=1}, ..., \alpha_{a=A}, \mathrm{K}, V_\kappa, \Sigma \right\}$ and $\Omega_t = \left\{ X, X^*, W_i, Z_a \right\}$ be the complete data. The following algorithm describes the estimation steps along with full conditionals for each 'sweep' (iteration) of the Gibbs sampler. All model parameters are estimated simultaneously, by recursively sampling from their full conditional posterior distributions, with diffuse conjugate priors:

$$\mathrm{B}^{(P)} \sim N_4\left(n_0 = 0, S = 10^5 I\right)$$

$$\mathrm{B}^{(I)} \sim N_4\left(n_0 = 0, S = 10^5 I\right)$$

$$\beta_1, \ldots, \beta_6 \sim N\left(n_0 = 0, S = 10^5\right)$$

$$\Lambda^{(P)}, \Lambda^{(I)} \sim N_4\left(\Lambda_0 = 0, \Sigma_\Lambda = 10^7 I\right)$$

$$\mathrm{K}^{(P)}, \mathrm{K}^{(I)} \sim N_2\left(\mathrm{K}_0 = 0, \Sigma_\mathrm{K} = 10^7 I\right)$$

$$V_\lambda^{-1(P)}, V_\lambda^{-1(I)} \sim G\left(\rho_\lambda = 4 + 50\ , R_\lambda = 10^{-8}\right)$$

$$V_\kappa^{-1(P)}, V_\kappa^{-1(I)} \sim G\left(\rho_\kappa = 2 + 50\ , R_\kappa = 10^{-8}\right)$$

$$\Sigma_\sigma^{-1} \sim W_2\left(\rho_\sigma = 2 + 10^0, R_\sigma = 10^{-4} I_2\right)$$

ONLINE APPENDIX 2

AD CONTENT CODING

Twenty professionals from an ad agency were recruited to participate in the coding of creativity, information content and degree of brand association in the 82 ads. Two were excluded for failure to complete the task. Coders had between 0 and 20 years of experience (mean=3.4, s.d.= 3.4) and were paid $10 per hour of coding. To minimize fatigue, either 20 or 21 ads were given to each coder in order to have at least four coders per ad. Coders were asked to view each ad once before answering questions regarding the ads. After that, they were told that they would be allowed to watch each ad, pause or rewind as many times as necessary to answer each question with confidence.

To code for informativeness, they were required to assess the presence or absence of twelve cues pertaining to Resnik and Stern's (1977) scale. A commercial is informative the more it contains the following cues about the product, service, or institution: price or value, quality, performance, components or contents, availability, special offers, taste, pack-aging or shape, guarantees or warrantees, safety, nutrition or independent research. Coders also assessed the overall ad's infomativeness using a 7-point scale. As for coding creativity, Smith et al. 2007 developed a scale to measure multiple dimensions of ad creativity, compared to the average TV ad. We use their global measures of overall creativity consisting of 3 items: creativeness, award worthiness and inventiveness using 7-point Likert scales. Lastly, coders assessed the strength of the association between entertainment and the brand claim(s) using a four-item, 7-point semantic differential measure of relevance (correspondence, relationship, fit and consistency) based on research by Bryant, Hezel, & Zillman, 1979. The scales ranged from 1 (very low) to 7 (very high). The items coded, wording of questions and types of measure are provide in the following table.

AD CODING QUESTIONS

Measure	Construct	Question	Type
Numbers of cues	Information	Does the ad provide any of the following cues about the product, service or brand?	Presence of 12 cues of product or brand information
Informative	Information	How informative do you judge this ad to be?	7-point scale, anchored by 'Not at all' to 'Very'
Creativeness	Creativity	The ad is very creative.	7-point Likert scale
Award worthiness	Creativity	The ad should win an award for creativity.	7-point Likert scale
Inventiveness	Creativity	The ad is very inventive and displays much creativity in its design.	7-point Likert scale
Correspondence	Brand association	The strength of correspondence between entertainment and the brand claim(s) is:	7-point scale, anchored by 'Very low' to 'Very high'
Relationship	Brand association	The strength of relationship between entertainment and the brand claim(s) is:	7-point scale, anchored by 'Very low' to 'Very high'

Fit	Brand association	The strength of fit between entertainment and the brand claim(s) is:	7-point scale, anchored by 'Very low' to 'Very high'
Consistency	Brand association	The strength of consistency between entertainment and the brand claim(s) is:	7-point scale, anchored by 'Very low' to 'Very high'

The items were averaged by construct and coders to form a single measure for information, another measure for creativity and a third for Brand association for each ad. The following table provides summary statistics of these measures, which were used as controls in the regressions.

AD CODING SUMMARY STATISTICS

Variable	Variation unit	N*	Mean	Std. Dev.	Minimum	Maximum
Information	ad	82	3.0	1.1	1.0	5.1
Creativity	ad	82	3.8	1.1	1.5	6.6
Brand association	ad	82	4.1	1.0	1.5	5.9

Note: Correlation between Entertainment and Information is 0.16 and between Entertainment and Creativity is 0.32.

The cross-item reliability of the items proved to be high with moderate to high intercorrelations within construct items. Due to the small number of professional coders, four to five per ad, and the well-known challenge of measuring subjective constructs such as creativity, intercoder reliability as measured by the conservative Krippendorff's alpha is low. Ordinarily, more coders would need to be used to increase reliability. But the challenge of finding more of these professionals precluded this approach. And given that these variables are used as additional controls versus variables of interest, lower alphas measures are tolerable.

AD CODING ITEM AND CODER RELIABILITY AND INTERCORRELATIONS

Construct	**Items**	**Cronbach alpha**	**Range of intercorrelation**	**Krippendorff's alpha**
Information	Numbers of cues	0.67	(0.54, 0.54)	0.254
	Informative			0.445
Creativity	Creative	0.90	(0.74, 0.82)	0.233
	Award winning			0.238
	Inventive			0.193
Brand association	Correspondence	0.91	(0.59, 0.89)	0.193
	Relationship			0.180
	Fit			0.064
	Consistency			0.154

Note – Inter-item reliability is asses via Cronbach's alpha and intercoder reliability via Krippendorff's alpha.