Web Appendix

A. Simulation Study I

For the first simulation study, we generated synthetic data according to Model 1 which assumes individual level structural heterogeneity. Here, we let I = 100 consumers, J = 5 brands, T=2 dimensions, and R=7 situations of which we use the first 6 situations for calibration purpose using the last situation for hold-out sample (as in our drug application). Next, given arbitrarily assigned brand positions, we randomly generated vectors v_{it}, ideal points w_{it}, and additive parameters b_i and d_i from a normal distribution. Note that we set the scale parameter $c_i = -1$ for ease of interpretation of the derived joint space. We then generated individual level structural heterogeneity indicators for 100 consumers in which 50 consumers display vector utility, while the remaining 50 consumers display ideal point utility. We then calculate the true preference data with known structural heterogeneity indicators and model parameters, and add random error generated from the standard normal distribution.

Next, we estimated the four different models discussed in the manuscript (vector only, ideal point only, Model 1, and Model 2) for t = 1,…,3 dimensions. Here, we execute the MCMC procedure for 20,000 iterations, where the last 10,000 iterations are used for inference. Convergence was checked by starting the chain from multiple starting points and by inspecting iteration series plots. As mentioned in the manuscript, Model 2 (which assumes occasion level structural heterogeneity) is a generalized version of Model 1, as Model 2 becomes Model 1 if all situations of an individual are classified to either vector or ideal point model. As shown in Table A1, both Model 1 and Model 2 which both assume structural heterogeneity clearly outperform models without structural heterogeneity across all dimensions. Second, models without structural heterogeneity tend to show improved fit with the increase of dimensionality, while models with structural heterogeneity identify the true number of dimensions.

[Insert Table A1 here]

As shown in Table A1, all model selection criteria suggest that two dimensions are most parsimonious for Models 1 and 2 for this simulated data as expected. A cursory look at Table A1 shows all the model selection criteria point to Model 1 as the best fitting model as expected. Note that Model 2 is a generalized version of Model 1. As such, Model 2 is expected to perform as well as Model 1. A close inspection of each model selection criterion shows that while Newton and Raftery's fourth estimate $\hat{p}_4(D)$ positively supports Model 1, the Harmonic Mean $\hat{p}_1(D)$ shows little difference between Model 1 and Model 2. The BIC is clearly better for Model 1 over Model 2. This is due to the large number of situation level structural heterogeneity indicator parameters (i.e., number of situations × number of subjects) in Model 2. Note that BIC is calculated as $BIC = -2\log\left(Pr\left(Data\middle|\hat{\Theta}_k, H_k\right)\right) + \log(N)\,NP$, where $\log\left(Pr\left(Data\middle|\hat{\Theta}_k, H_k\right)\right)$ is the maximum log likelihood given estimated parameters ($\hat{\Theta}_k$) for Model H_k for a specified number of iterations, N is the number of observations, and NP is the number of parameters. On the other hand, method independent validation measures such as RMSE (Root Mean Squared Error) and VAF (Variance Accounted For) show almost identical results between Model 1 and Model 2. A close inspection of log likelihoods for both Model 1 and Model 2 in Figure 1 shows that Model 1 has slightly higher Log Likelihoods than Model 2 during the MCMC draws, even though the two models have the same mode as shown in Figure A1. As a result, the RJMCMC measure favors Model 1 (Prob. = 0.6706) to Model 2 (Prob. = 0.3280), although both models correctly identified all individuals' structural heterogeneity indicators.

[Insert Figure A1 Here]

Now we turn our attention to the recovery of true parameters for Model 1. First, we calculate RMSE for all parameters in interest to test recovery of true parameters. Table A2 shows the average RMSE and information regarding standard deviations for brand coordinates x_{jt}, vectors v_{it}, ideal points w_{it}, and additive parameters b_i and d_i across the MCMC draws . As shown, the proposed model uncovers the true parameters with high accuracy and small standard deviations. This can be confirmed by the comparison between true joint space and estimated joint space in Figure A2 (after appropriate configuration matching).

[Insert Table A2 and Figure A2 Here]

Next, we inspect the posterior estimates of the hyperpriors. As shown in Table A3, the means and standard deviations of the posterior estimates of the hyperpriors indicate that different levels of uncertainty are reflected in these estimates. For instance, the hyperprior for brand coordinates τ_x^2, additive constants for the ideal point model τ_d^2, and vectors τ_v^2 show relatively high means and standard deviations, indicating higher levels of uncertainty.

[Insert Table A3 Here]

B. Simulation Study II

For the second simulation study where we assume situation level structural heterogeneity, we again use I = 100 consumers, J = 5 brands, R = 7 situations, and T = 2 dimensions. Similar to the first simulation study, we first generated the true parameters from Normal distributions. Next, we randomly assigned some individuals into vector (44%) and ideal points (39%) models without situation level structural heterogeneity, and randomly generated situation level structural heterogeneity indicators using a binomial distribution with prob. = 0.5 for the remaining individuals (17%). Note that we set the scale parameter $c_i = -1$ for this second simulation study as well. We then calculated the true preference data with known structural heterogeneity indicators and model parameters, and added random error generated from the standard normal distribution. Next, we estimated the same four different models discussed in Simulation A above for t = 1,…,3 dimensions. Here too, we executed the MCMC procedure for 20,000 iterations, and the last 10,000 iterations were used for inference.

In this second simulation where we assume the presence of situation level structural heterogeneity, all model selection criteria clearly point to the two dimensional Model 2 solution for this data. It is quite surprising that Model 2 has an almost 20,000 log marginal likelihood improvement over Model 1, which assumes only individual level structural heterogeneity. Further, the RJMCMC method clearly shows that Model 2 has Prob. = 1; all other model selection criteria show a similar consistent picture in Table A4. As in the first simulation study, models without structural heterogeneity result in higher dimensionality as both vector only and ideal point only models show their best fitting in three dimensions.

[Insert Table A4 Here]

Inspection of recovery of model parameters measured by RMSE in Table A5 shows that the RMSE for vectors v_{it} is relatively high compared to that of other parameters. This is perhaps due to the lack of number of situations as the RMSE for one particular occasion seems to be extremely high (i.e., 3.87) as shown in Table A6. This is not surprising since the parameters are randomly drawn from a normal distribution such that $P(v_{it}|\sim)\sim N(\bar{v}_i, V_{v_i})$ if there is any situation assigned to the vector utility, and $P(v_{it}|\sim)\sim N(0, \tau_v^2)$ otherwise. If there is only one situation that can contribute to the vector utility, one can expect that the random draws of the vector parameters would switch between these two distributions. Since the posterior deviates of the hyperprior τ_v^2 show extremely high uncertainty, the resulting posterior estimates of vector parameter would reflect this higher level of uncertainty. Otherwise, Model 2 uncovers the true parameters with high accuracy and small standard deviations. This can be confirmed by the comparison between true joint space and recovered joint space (after appropriate configuration matching) displayed in Figure A3.

[Insert Tables A5, A6 and Figure A3 Here]

Finally, we inspect the posterior estimates of hyperpriors. As shown in Table A7, the posterior estimates of the hyperprior for additive constants for the ideal point model τ_d^2 has high mean and standard deviation, reflecting higher levels of uncertainty.

[Insert Table A7 Here]

C. <u>Summary</u>

We have conducted two simulation studies where different levels of structural heterogeneity are assumed. The results indicate that the proposed model can recover the true underlying structure, and that the model selection criteria show consistent results. Together, these results suggest that a model that incorporates both structural and preference heterogeneity outperforms models without structural heterogeneity when individual and/or situation level structural heterogeneity is present in the observations. Moreover, the incorporation of structural heterogeneity can lead to more parsimonious preference structures as models without structural heterogeneity seem to require higher dimensionality.

Table A1: A Comparison of Model Selection Criteria for Simulation I

				Calibration Data				Hold out Data	
Dimension	Model	Newton and Raftery	Harmonic Mean	BIC	RJMCMC	RMSE	VAF	RMSE	VAF
	Vector only	-128138	-128110	257806	0.0000	9.142	0.670	9.197	0.668
1	Ideal Point only	-128158	-128158	257846	0.0000	9.146	0.668	9.200	0.666
	Model 1	-101164	-100468	199628	0.0000	7.258	0.729	7.284	0.729
	Model 2	-102490	-101791	204259	0.0000	8.111	0.728	8.147	0.728
	Vector only	-5907	-5853	13950	0.0000	1.437	0.992	1.437	0.992
2	Ideal Point only	-9651	-9566	21445	0.0000	2.131	0.981	2.122	0.982
	Model 1	**-4298**	**-4262**	11746	**0.6706**	**1.002**	**0.996**	**1.059**	**0.996**
	Model 2	-4302	-4263	15750	0.3280	1.002	0.996	1.058	0.996
	Vector only	-4708	-4606	12141	0.0000	1.110	0.996	1.177	0.996
3	Ideal Point only	-4694	-4530	12210	0.0000	1.088	0.996	1.176	0.995
	Model 1	-4349	-4322	12627	0.0008	1.022	0.996	1.098	0.996
	Model 2	-4357	-4321	16662	0.0006	1.021	0.996	1.122	0.995

Table A2: RMSE and Standard Deviation of MCMC

Parameters	RMSE*	Average Std.Dev.#	Maximum Std.Dev.#	Minimum Std.Dev.#
x_{jt}	0.050	0.033	0.061	0.025
v_{it}	0.090	0.085	0.133	0.073
b_i	0.154	0.163	0.165	0.160
w_{it}	0.022	0.042	0.054	0.037
d_i	0.282	0.273	0.481	0.219

* Root Mean Squared Error; # Standard Deviation

Table A3: Means and Standard Deviations of the Hyper Priors

Parameters	Mean	Std.Dev.
τ_x^2	11.30	9.54
τ_b^2	0.13	0.03
τ_d^2	152.93	31.38
τ_v^2	17.24	2.47
τ_w^2	2.19	0.31

Table A4: A Comparison of Model Selection Criteria for Simulation II

		Calibration Data						Hold out Data	
Dimension	Model	Newton and Raftery	Harmonic Mean	BIC	RJMCMC	RMSE	VAF	RMSE	VAF
1	Vector only	-68491	-68465	138527	0	6.619	0.798	6.388	0.806
	Ideal Point only	-69102	-69058	139654	0	6.622	0.797	6.398	0.805
	Model 1	-68465	-68439	139291	0	6.617	0.798	6.387	0.807
	Model 2	-60220	-60185	126759	0	6.188	0.823	6.749	0.786
2	Vector only	-28001	-27945	58247	0	4.098	0.923	3.586	0.940
	Ideal Point only	-28959	-28921	60216	0	4.177	0.920	3.665	0.937
	Model 1	-25219	-25182	53583	0	3.867	0.931	3.311	0.948
	Model 2	**-4754**	**-4728**	16672	**1**	**1.147**	**0.994**	**1.047**	**0.995**
3	Vector only	-25316	-25229	53575	0	3.871	0.932	3.327	0.948
	Ideal Point only	-25787	-25598	54172	0	3.902	0.930	3.356	0.947
	Model 1	-25287	-25238	54462	0	3.871	0.931	3.326	0.948
	Model 2	-5646	-5544	19031	0	1.363	0.992	1.079	0.995

Table A5: RMSE and Standard Deviation of MCMC

Parameters	RMSE*	Average Std.Dev.	Maximum Std.Dev.	Minimum Std.Dev.**
x_{jt}	0.049	0.025	0.038	0.024
v_{it}	0.526	0.125	0.181	0.000
b_i	0.131	0.178	0.281	0.159
w_{it}	0.029	0.072	0.939	0.007
d_i	0.330	0.527	7.280	0.164

* Root Mean Squared Error; ** Standard Deviation

Table A6 : RMSE of v_{it} by the Number of Situations

Number of occasions	Count	Average RMSE
1	2	3.87
2	5	0.14
3	7	0.14
4	1	0.13
5	2	0.06
6	44	0.11

Table A7: Means and Standard Deviations of Hyper Priors

Parameters	Mean	Std.Dev.
τ_x^2	7.43	5.89
τ_b^2	0.13	0.03
τ_d^2	166.43	32.31
τ_v^2	17.91	1.83
τ_w^2	2.63	0.36

Figure A1: Comparison of Log Likelihoods between Model 1 and Model 2

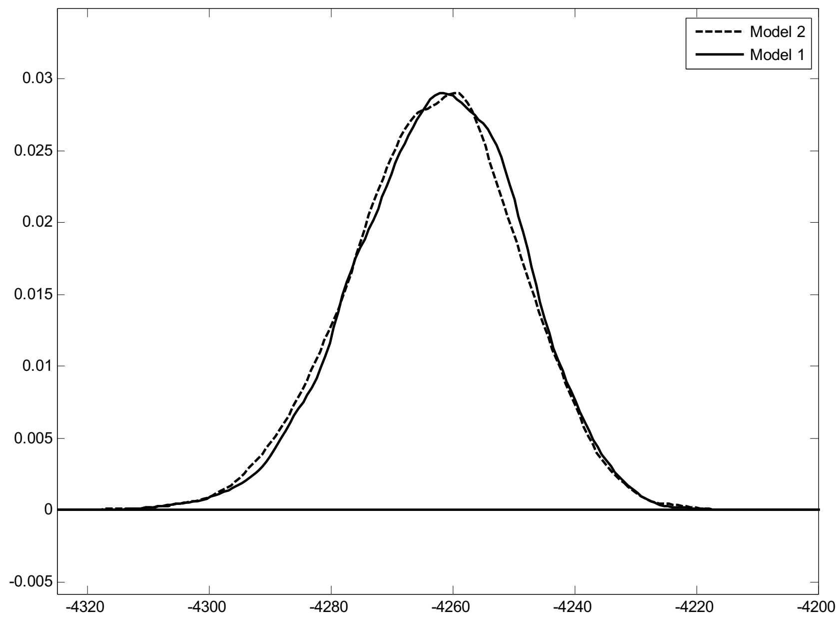
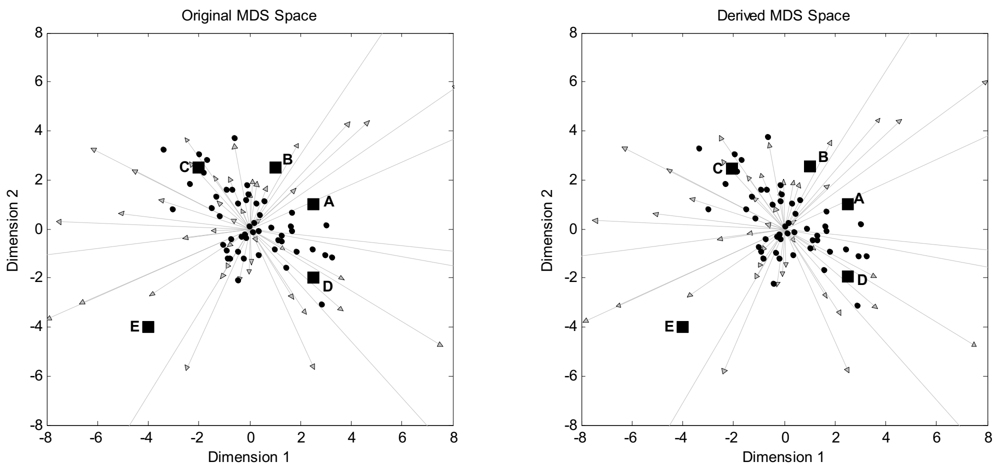
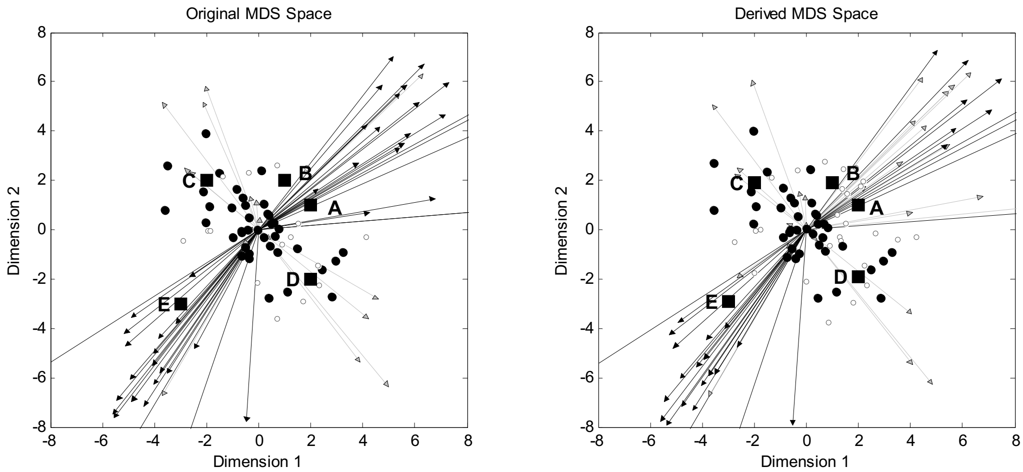


Figure A2: Comparison between the True Joint Space and the Estimated Joint Space for Model I



Note: individuals with only vector utility are represented as gray vectors, and individuals with only ideal point utility are represented as black circles.

Figure A3: Comparison between the True Joint Space and the Estimated Joint Space for Model II



Note: individuals with only vector utility are represented as solid black vectors, and individuals with only ideal point utility are represented as black circles, while individuals with both vector and ideal point utility are represented as gray vectors and circles.