PRESENTED AT THE 2005 ART FORUM

Using Parsimonious Conjoint and Choice Models to Improve the Accuracy of Out-of-Sample Share Predictions

Jay Magidson
Statistical Innovations

Tom Eagle
Eagle Analytics

Jeroen K. Vermunt
Tilburg University

Background

- By accounting for continuous respondent heterogeneity, traditional approaches such as hierarchical Bayes (HB) models for estimating unconstrained random effects regression models outperform *aggregate* models with respect to *in-sample* predictions.
- It has been suggested however, that HB over-fits sample data and thus *out-of-sample* predictions might be worse than indicated by in-sample performance and be more susceptible to invalid inferences (Natter et. al., 2002; Andrews et. al., 2002; Magidson et. Al, 2003).
- As a solution to the overfitting problem, a new class of continuous factor (C-Factor) models has been proposed, providing a more parsimonious alternative to HB for estimating random and mixed effects conjoint and choice models.

Outline of Presentation

- Overfitting – what is it?
  - Illustration based on a simple fixed effects regression simulation
- Introduction to CFactor models for random effects
- Comparison of CFactor and traditional HB approaches
  - Example 1: Analysis of sensory preferences for crackers
  - Example 2: Bank segmentation data re-analysis
- Conclusions

Overfitting – What is it?

- Models that overfit data contain too many free parameters. While overfitting improves upon in-sample predictions, such improvement may come *at the expense of* out-of-sample performance.

- In the presence of overfitting, in-sample performance measures such as R^2 and hit rate shrink when used to predict out-of-sample performance.

- With a moderate amount of overfitting, the shrinkage becomes so large that out-of-sample performance becomes substantially worse than that obtained with little or no overfitting.

Simple Simulated Example of Overfitting:

Including extraneous predictors in a regression model

- Use a linear regression to estimate fixed effects of K predictors Z_1, Z_2, ..., Z_K , on the dependent variable Y.

- Assume that only the first 2 predictors are important in determining Y; that is, the K-2 predictors Z_3, Z_4, ..., Z_K are extraneous.

- Use the true regression model to generate 2 sets of values for the dependent variable (Y_1 and Y_2) based on 2 exchangeable (matched) samples, one to be used for model estimation, the other for validation.

Simple Simulated Example of Overfitting

For matched* samples m = 1, 2, assume the model:

$Y_m = \alpha + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_1 + \ldots + \beta_K Z_K + \varepsilon_m$ is 'truth'

where $\alpha = 0$

$\beta_1 = .3$ $\quad\beta_3 = \beta_4 = \ldots = \beta_K = 0$

$\beta_2 = .4$

N = 100 cases are simulated for each sample as follows:

each Z_k , ε_m i.i.d. N(0,1), k=1,2,...,K m=1,2. This implies $R^2 = 0.2$

* The samples are matched (exchangeable) in the sense that each case in sample #2 has the same predictor values for Z_1, $Z_{2,}$ ... Z_K, as its matched counterpart in sample 1. The samples differ from each other only with respect to the random error term ε_1 and ε_2.

Results of simulation where Population R^2 = .2

What happens if extraneous predictors Z_3, Z_4, ... are included?

(K-2) # extraneous predictors	In-Sample Predictions: R^2 obtained from OLS regression — Sample 1	Sample 2	Out-of-Sample Predictions: R^2 obtained when model estimates from one sample are used to generate predictions for the other sample — Sample 2 used to predict Sample 1	Sample 1 used to predict Sample 2
0	0.148	0.267	0.146	0.264
3	0.181	0.281	0.166	0.257
6	0.200	0.333	0.143	0.239
9	0.212	0.340	0.152	0.245
12	0.261	0.344	0.154	0.203
32	0.480	0.476	0.104	0.104
52	0.660	0.740	0.047	0.053
97	1.000	1.000	0.034	0.034

- With no extraneous predictors, the sample R^2 values are unbiased estimates of the true R^2, and the corresponding shrinkage when applied out-of-sample is quite small.
- The in-sample R^2 increases well beyond .2 with 12 or more extraneous predictors, and the amount of shrinkage when applied out-of-sample increases substantially.
- Out-of-sample performance begins to deteriorate with a moderate amount of overfitting, as irrelevant variation dominates the predictions.

Overfitting Leads to Invalid Inferences

Regression Results with 12 Extraneous Predictors

Sample 1: R^2 = 0.261

	Estimate	Std. Error	p-value
(Constant)	0.04	0.11	0.75
z1	0.27	0.13	0.04
z2	0.38	0.13	0.00
z3	0.11	0.13	0.40
z4	-0.08	0.11	0.48
z5	0.19	0.12	0.11
z6	0.00	0.12	0.98
z7	-0.10	0.12	0.39
z8	-0.13	0.12	0.28
z9	-0.06	0.11	0.58
z10	0.01	0.11	0.93
z11	-0.12	0.12	0.33
z12	0.10	0.10	0.31
z13	0.22	0.11	0.05
z14	0.11	0.13	0.39

The estimates for α, β_1 and β_2 are close to the correct values.

The out-of-sample prediction problem occurs by allowing irrelevant variance associated with the extraneous effects to contribute to the model. When the number of extraneous predictors is larger, the problem becomes worse as the irrelevant variance dominates the model.

This sample #1 model containing 12 extraneous predictors incorrectly implies that higher levels of Z_3, Z_5, Z_{12}-Z_{14} and lower levels of Z_4, Z_7, Z_8 and Z_{11} will yield a higher Y.

Summary and Implications of Results

- In this simulated example involving fixed effects regression, K regression coefficients were freely estimated although only 2 were non-zero. For moderate amounts of overfitting (K >15), out-of-sample performance begins to deteriorate.

- In practice, the true underlying structure is unknown. If a model is applied that fits the data and reduces the number of parameters, the likelihood of overfitting is reduced.

- In random effects regression, it is more important to impose a parsimonious model structure because there are many more parameters, and hence the potential for overfitting is a larger problem.

- Unlike fixed effects models, with *random* effects models it is *not* possible to apply the model parameters to a 'hold-out' sample without making strong assumptions about the nature of the random effects. Thus, there is no *direct* way to estimate the magnitude of overfitting.

How Many Model Parameters?

- In *fixed effects* linear regression, there are K+2 model parameters: K+1 regression coefficients α, β_1, β_2, ..., β_K and error variance σ^2.

- In *random effects* linear regression, *individual-level* α, and β coefficients are estimated for each case i:

$$Y_i = \alpha_i + \beta_{1i}Z_1 + \beta_{2i}Z_2 + \beta_{3i}Z_1 + \ldots + \beta_{Ki}Z_K + \varepsilon_i$$

Thus, in addition to K+2 parameters as in the fixed effects model (means of α and βs, now called 'hyper-parameters' plus σ^2), additional hyper-parameters must be estimated based on the assumed distribution of the random effects α and βs. Typically it is assumed that (α, β) follow the MVN distribution, which means that there are an additional (K+2)(K+1)/2 hyper-parameters in the variance-covariance matrix that must be estimated.

- The individual-level coefficients are derived from estimates of the hyper-parameters.

The Hierarchical Bayesian Approach

- Bayesian methods (HB) are often used to estimate random effects regression models. Typically, this is done without imposing constraints on the hyper-parameters. This can yield a very large number of parameters to be estimated. For example, for K=12, there are 105 parameters!

- One way to reduce the number of parameters, is to allow for the possibility that only *some* of the effects are random, the others fixed – i.e., mixed models. For example, a *random intercept model* is one where individual-level coefficients are obtained for α only, the βs being estimated as fixed effects. For K=12, the random intercept model contains only 14 parameters.

- When the βs are also treated as random effects, a parameter reduction can be achieved by superimposing a factor-analytic structure on the variance-covariance matrix (CFactor models). If the number of CFactors is sufficiently small, maximum likelihood estimation becomes feasible.

Continuous Factor Models

- Random (mixed) effects models can be re-parameterized in terms of continuous factors (CFactors) by superimposing a factor-analytic structure with $P \leq K+1$ CFactors on the random coefficients*.

- For example, for K=3 predictors and P=2 CFactors F_1 and F_2:

$$Y_i = \alpha_i + \beta_{i1} Z_1 + \beta_{i2} Z_2 + \beta_{i3} Z_3 + \varepsilon_i$$

Factor-Analytic Structure:

$$\alpha_i = \alpha_0 + \lambda_{10} F_{i1} + \lambda_{20} F_{i2}$$
$$\beta_{i1} = \beta_{01} + \lambda_{11} F_{i1} + \lambda_{21} F_{i2}$$
$$\beta_{i2} = \beta_{02} + \lambda_{12} F_{i1} + \lambda_{22} F_{i2}$$
$$\beta_{i3} = \beta_{03} + \lambda_{13} F_{i1} + \lambda_{23} F_{i2}$$

✓ $(F_{i1}, F_{i2}, F_{i3}) \sim MVN(0, I)$

✓ The means of the individual coefficients are represented by α_0 and β_0

✓ The random effects are represented by λs. Thus, for example, $VAR(\beta_{i1}) = \lambda_{11}^2 + \lambda_{21}^2$

✓ Note that $\lambda_{12} = \lambda_{22} = 0$ implies that $\beta_{i2} = \beta_{02}$ is a fixed effect.

* Skrondal and Rabe-Hesketh, 2004; Vermunt and Magidson, 2005

Example 1: Conjoint with Soft Attributes

- **Products**: T = 15 crackers
- **Consumers**: n=157 (category users)
  - evaluated all products over three days
  - Y = 9-point liking scale RATING (dislike extremely➔like extremely)
  - completely randomized block design balanced for the effects of day, serving position, and carry-over
- **Sensory attribute evaluations**: trained sensory panel (n=8)
  - 18 flavor, 20 texture, and 14 appearance attributes rated on 15-point intensity scales (low➔high)
  - reduced (via PCA) to four appearance, four flavor, and four texture factors – Z_1, Z_2, ..., Z_{12} (K=12)
  - For this presentation, the client requires that these attributes be named only as App1-4, Flv1-4 and Tex1-4.

CFactor Regression Data Layout

crackers4.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

1 : caseID 1101

	caseID	rating	App1	App2	App3	App4	Flv1	Flv2	Flv3	Flv4	Tex1	Tex2	Tex3	Tex4
1	1101	6	-.13	-.08	2.16	-.32	.53	-.71	1.2	.18	-.42	-1.28	-.79	-1.14
2	1101	7	-.44	-1.2	-1.71	-1.89	1.71	-.30	.6	.63	2.34	.66	.27	-1.94
3	1101	6	-.46	-.97	1.06	1.10	-.70	-.76	.2	.91	-1.13	1.10	1.95	-.89
4	1101	6	3.35	.31	.18	-.76	-.32	-.58	-.3	-.15	-1.18	-.98	-.56	.29
5	1101	6	-.20	.63	-.89	.71	.47	-.73	-.1	.81	.70	.98	.60	.18
6	1101	8	-.43	.81	-.96	-.67	.75	-.66	-1.8	-2.02	1.28	-1.28	.77	.86
7	1101	9	-.22	1.71	.13	.58	-.82	-.07	-1.3	-.64	.17	.77	.02	1.62
8	1101	9	-.27	-.60	-.21	1.06	.08	2.91	-.9	.93	-.24	1.01	-.57	.12
9	1101	7	-1.02	.15	.72	-1.58	-.34	.38	.9	-1.93	.69	-.47	-1.47	-.23
10	1101	8	-.28	-1.2	-.44	.92	.08	-.25	1.6	-.01	-.69	.24	-1.36	-.70
11	1101	6	.81	-.91	-.80	1.27	-2.75	-.51	-.1	.60	-1.22	-.51	1.50	-.52
12	1101	9	.06	-1.1	.21	-.11	-.21	.89	1.1	-.73	-.25	1.25	-.97	.56
13	1101	9	-.35	1.01	-.73	-.18	.92	-.57	-.8	.37	.68	-1.65	.83	.96
14	1101	8	-.19	-.27	1.25	-.81	.22	1.18	.5	-.33	-.71	-.51	-.17	-.67
15	1101	8	-.24	1.65	.04	.68	.38	-.22	-1.0	1.36	-.01	.68	-.05	1.47
16	1102	8	-.13	-.08	2.16	-.32	.53	-.71	1.2	.18	-.42	-1.28	-.79	-1.14
17	1102	7	-.44	-1.2	-1.71	-1.89	1.71	-.30	.6	.63	2.34	.66	.27	-1.94
18	1102	6	-.46	-.97	1.06	1.10	-.70	-.76	.2	.91	-1.13	1.10	1.95	-.89
19	1102	6	3.35	.31	.18	-.76	-.32	-.58	-.3	-.15	-1.18	-.98	-.56	.29
20	1102	9	-.20	.63	-.89	.71	.47	-.73	-.1	.81	.70	.98	.60	.18

Data View / Variable View

SPSS Processor is ready

There are 15 records per case. The repeated measures regression predicts the RATING for the 15 crackers as a linear function of the 12 sensory attributes.

Estimating Random Effects Regression Models
Unconstrained HB Approach

- Allow for continuous heterogeneity*

$$Y_{i.t} = \alpha_i + \beta_{i1} Z_1 + \beta_{i2} Z_2 + \ldots + \beta_{iK} Z_K + \varepsilon_{i.t}$$

$Y_{i \cdot t}$ = case i's rating of product t

Unconstrained HB** approach:
- Assumes these random coefficients follow a MVN distribution
- Uses Bayesian methods for estimation

- In the example, the number of parameters is computed as follows:

13 means + 13 standard deviations + $\binom{13}{2}$ = 78 correlations + 1 error variance
= 105 free parameters to be estimated (a large number)

* For simplicity, we omit quadratic terms.
** We used Gauss to estimate the HB models presented here.

Estimating Random Effects Regression Models
General CFactor Approach

$$Y_{i.t} = \alpha_i + \beta_{i1} Z_1 + \beta_{i2} Z_2 + ... + \beta_{iK} Z_K + \varepsilon_{i.t}$$

$Y_{i.t}$ = case i's rating of product t

CFactor Model fitting strategy * :

- Assumes these coefficients follow a MVN distribution
- Superimposes a factor-analytic structure on the regression coefficients
- Uses maximum likelihood estimation
- Uses information criteria (e.g., BIC) to determine number of CFactors
- Uses Wald tests to determine whether to restrict loadings to 0

- In the example, the final model contained only 2 CFactors -- 1 CFactor for the (random) intercept, and a 2nd associated with the attribute effects.

The number of parameters were reduced from 105 to 19 !!!

✓ 13 means + 13 loadings +1 factor correlation + 1 error variance**

✓- 9 nonsignificant parameters set to 0 (1 mean and 8 loadings).

* We used Latent GOLD 4.0 (Vermunt and Magidson, 2005) to estimate the CFactor Model. Such models could also be estimated using the program gllamm (Skrondal and Rabe-Hesketh, 2004).

** Additional CFactor models allowing separate error variances for each product t were also evaluated.

Model Comparisons

BIC is Best for Our Final Restricted 2-CFactor Model

Model	LG Model Specification	Explanation	LL	BIC	npar
1	0-CFactor	fixed effects	-5153.6	10378.0	14
2	LG 1-CFactor for intercept only	random intercept (RI)	-5093.3	10262.4	15
3	LG unrestricted 1-CFactor		-5068.3	10273.0	27
4	LG restricted 2-CFactor (1 only for intercept)	1-CFactor + extra variance term for intercept	-5036.4	10214.3	28
5	LG unrestricted 2-CFactor		-5023.6	10249.4	40
6	LG unrestricted 3-Cfactor		-5004.3	10276.5	53
7	Final restricted 2-CFactor		-5042.6	10181.3	19
8	SPSS 1-CFactor homogenous	1-CFactor + equal extra variance term on all diagonals	-5040.5	10222.7	28
9	SPSS 1-CFactor heterogenous	1-CFactor + extra variance terms on diagonals	-5016.4	10235.1	40*
10	SPSS variance components	only diagonals = uncorrelated random effects	-5048.6	10233.8	27**
11	SPSS unstructured (= implicit HB model)	full matrix	N/A	N/A	105

*3 diagonal parameters inestimable and 1 zero **1 variance estimated to be 0 (eliminated) N/A: does not converge: nonsense results

BIC suggests that more than 2-CFactors and/or additional error variances were not required.

Results from 3 Regression Models

- While the HB Regression R^2 is quite high, since it is based on 105 parameters, a substantial amount of shrinkage is possible if results were to be applied to an 'exchangeable' sample.

	Aggregate Regression		2 CFactor Regression		HB Regression	
# Parameters	14		19		105	
R^2 =	0.13		0.39		0.67	
	Estimate	Significant?	Mean	Std Dev	Mean	Std Dev
Intercept	5.95	Yes	5.95	0.64	5.95	0.68
APP1	0.31	Yes	0.30	0.23	0.30	0.31
APP2	0.85	Yes	0.77	0.43	0.85	0.38
APP3	-0.46	Yes	-0.43	0	-0.48	0.23
APP4	0.70	Yes	0.68	0.13	0.69	0.14
FLV1	0.93	Yes	0.92	0	0.93	0.17
FLV2	0.73	Yes	0.71	0	0.73	0.24
FLV3	1.03	Yes	1.03	0	1.03	0.20
FLV4	-1.07	Yes	-1.04	0.18	-1.07	0.25
TEX1	-0.55	Yes	-0.51	0	-0.57	0.25
TEX2	0.19	Yes	0.17	0	0.19	0.17
TEX3	0.38	Yes	0.38	0	0.38	0.15
TEX4	-0.08	No	0	0	-0.09	0.31

By setting the corresponding λs to 0, the CFactor model omits heterogeneity associated with the sensory attributes TEX1-4, FLV1-3 and APP3 as irrelevant variation. These are indicated above by Std Dev = 0.

Comparison of HB and CFactor Results

- Individual coefficients estimated for α based on the HB and CFactor models correlate almost perfectly with each other and with average ratings across all 15 products (all corrs >.99).

- The HB individual coefficients for TEX4 show substantial amounts of heterogeneity – the standard deviation is .31, which is third highest among all attributes. The CFactor model smooths away this heterogeneity as irrelevant variation.

- Generally, it is quite difficult to describe the heterogeneity obtained from an unconstrained HB model concisely. Using the factor structure helps. Here, since only CFactor #2 is used to describe the heterogeneity, this heterogeneity can be described quite simply in terms of differences in attribute and product preferences as shown on the following slides.

C-Factor Regression Results Indicate Fixed Effects on all but 4 Sensory Attributes

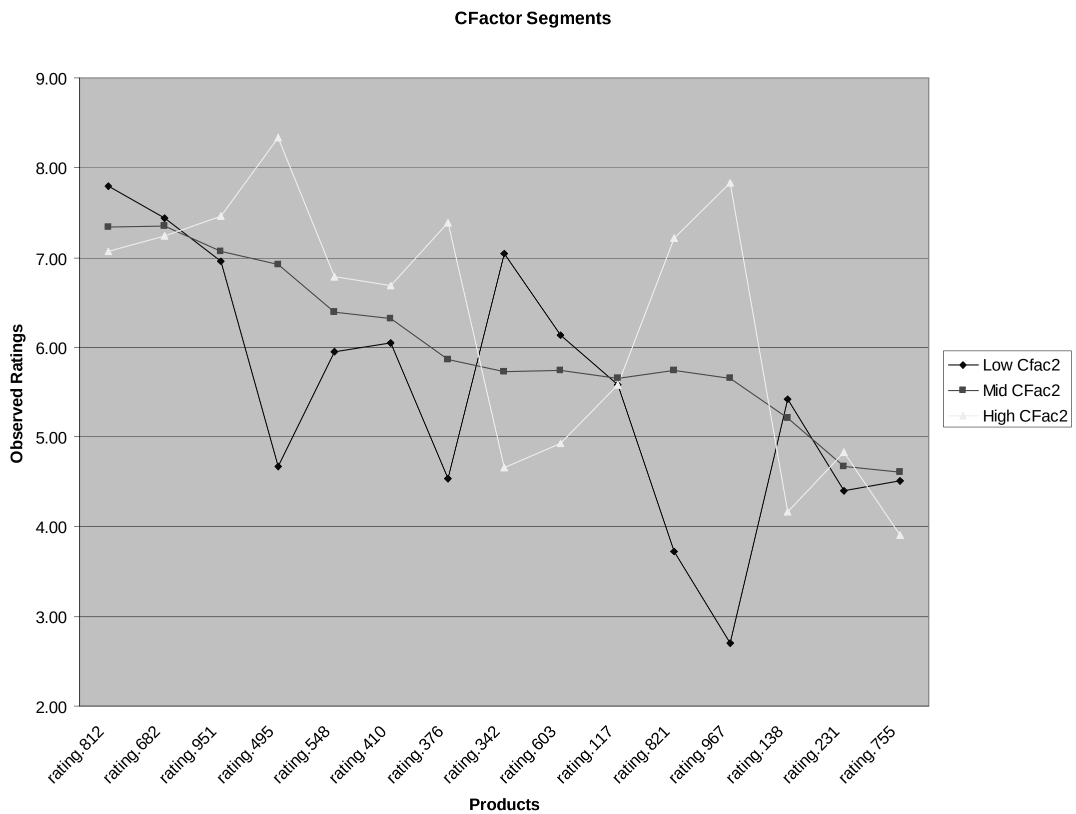
- The CFactor2 effects indicate that respondents scoring high on C-Factor2 were more favorably affected by higher values of APP2, APP4 and FLV4 and lower values of APP1 than those scoring low on C-Factor #2.

Attribute	**Fixed Component**	p-value	CFactor2 Effect	p-value
APP1	0.30	0.009	-.29	< 0.0001
APP2	0.77	< 0.0001	0.54	< 0.0001
APP3	-0.43	< 0.0001	0*	--
APP4	0.68	< 0.0001	0.16	0.0140
FLV1	0.92	< 0.0001	0*	--
FLV2	0.71	< 0.0001	0*	--
FLV3	1.03	< 0.0001	0*	--
FLV4	-1.04	< 0.0001	0.23	0.0003
TEX1	-0.51	0.007	0*	--
TEX2	0.17	0.004	0*	--
TEX3	0.38	0.002	0*	--
TEX4	0*	--	0*	--

Comparing Ratings by Products: Where is the Heterogeneity?

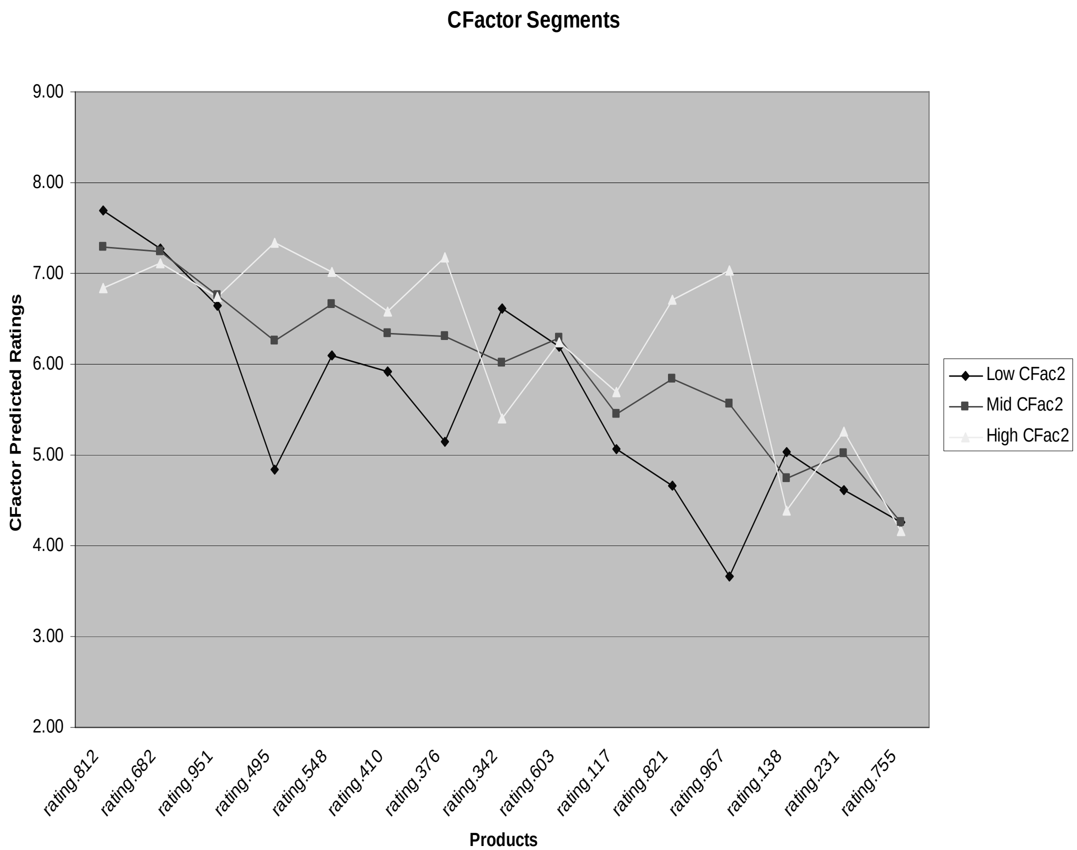
- Maximum likelihood estimates for the CFactor scores were obtained and used to classify respondents into 3 segments. Cutpoints of -.5 and .5 were used to discretize the estimated CFactor2 score to form these segments
- By plotting mean product ratings separately for each segment it can be seen how the CFactor2 heterogeneity translates into differences in relative preferences for one cracker product over another. In addition, such plots can be developed using observed as well as predicted ratings to see the extent to which the observed differences are captured in the model.
  - Note: Predicted ratings for those in the middle segment were similar to ratings predicted by the aggregate (fixed effects) model for all consumers.

Comparisons of Product Ratings by Segment
Average Observed Ratings



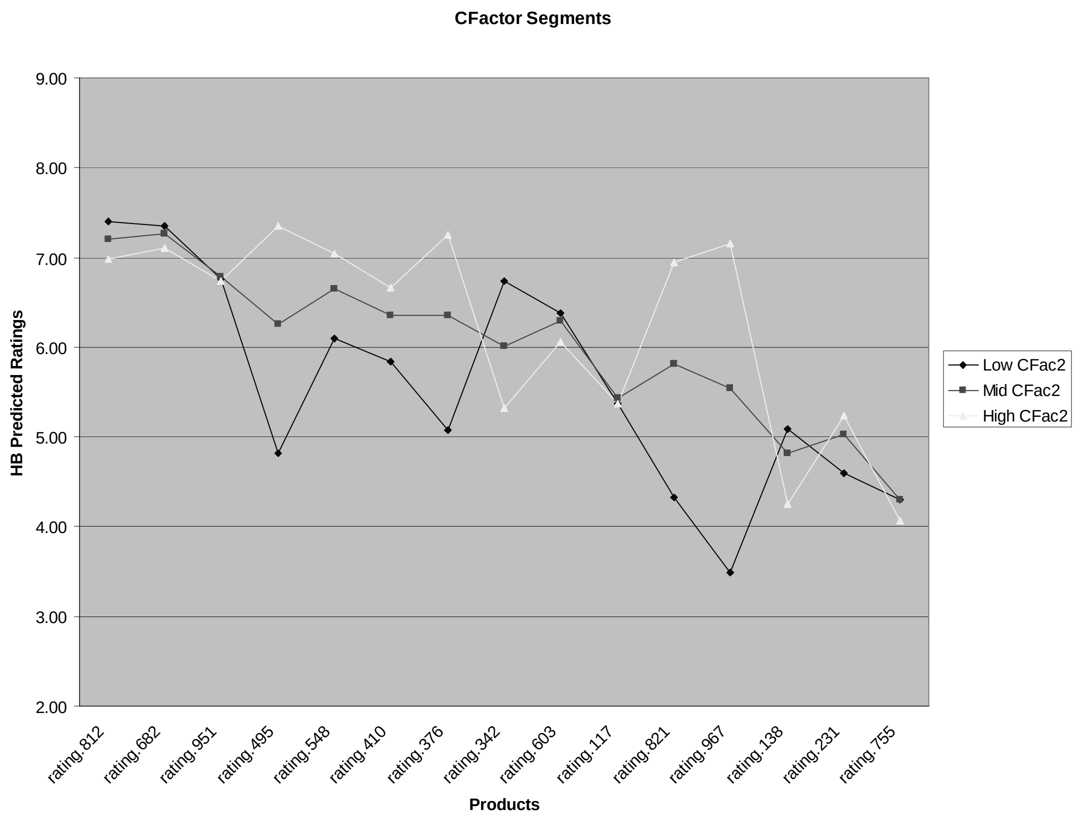
Consumers in the high CFactor2 segment were significantly more likely to prefer crackers #376, #495, #821, and #967 and significantly less likely to prefer cracker product #342 than those in the low CFactor2 segment.

Comparisons of Product Ratings by Segment
Average Predicted Ratings -- CFactor Model



These predicted ratings capture the major segment differences shown in the observed data

Comparisons of Product Ratings by Segment
Average Predicted Ratings -- HB Model



Average predicted HB ratings show similar patterns as the CFactor model predictions with respect to this CFactor2 type of heterogeneity.

Example 2: Checking Account Preferences:
Select most and least preferred from each set *

Set #	Profile #	MinBalance	Cost/Check	Fee	ATM
1	1	$0	$0	$6	No
1	2	$0	$0.15	none	Free
1	4	$500	$0	$3	Free
1	7	$1,000	$0	none	$.75 Fee
2	3	$0	$0.35	$3	$.75 Fee
2	5	$500	$0.15	$6	$.75 Fee
2	6	$500	$0.35	none	No
2	8	$1,000	$0.15	$3	No
2	9	$1,000	$0.35	$6	Free

Fewer parameters than example #1 -- only 5 beta parameters: 1 each for MinBalance, Cost/Check and Fee, and 2 for the nominal variable ATM

*** These data were restructured from a full ranking of all 9 alternatives (Kamkura, et. al., 1994).**

Restricted 2-CFactor Model is Preferred by BIC

Model				
	Npar	LL	BIC(LL)	Hit Rate*
1-Class (Aggregate)	5	-1122.03	2271.78	54.3%
1-CFactor	10	-1047.42	2150.30	69.5%
Restricted 2-CFactor	10	-1011.39	2078.23	73.4%
2-Class	11	-1041.28	2143.56	60.2%
2-CFactor	15	-1009.04	2101.27	76.2%
3-Class	17	-1002.41	2099.09	64.8%
HB	20			75.4%
3-CFactor	20	-1002.09	2115.08	91.0%
4-Class	23	-987.70	2102.94	67.2%
5-Class	29	-979.40	2119.61	80.1%
6-Class	35	-971.35	2136.77	75.0%

***Percent of first-choice hits in selection from *all 9* alternatives.**

Note: The restricted 2-CFactor model took 5 seconds to estimate using Latent GOLD 4.0 on a 2 gigahertz PC compared to over 1 hour for the HB model estimated using the standard software

C-Factors Structure the Heterogeneity

- All consumers prefer less minimum balance requirements, lower costs per check, lower monthly fees and inexpensive access to ATMs.

- Some consumers are more willing than others to trade off the minimum balance requirement against the cost per check (CFactor 1), while others are more willing than others to trade off the minimum balance requirement against ATM availability (CFactor 2).

Attribute	**Fixed Component**	CFactor1 Effect	CFactor2 Effect
MIN BAL	-3.99	-2.07	1.43
Cost per Check	-7.74	6.05	0*
Monthly Fee	- 0.40	0*	0*
ATM:			
Fee = $.75 /use	- 0.39	0*	0.34
Free	1.19	0*	0.69
N/A	- 0.80	0*	-1.03

Conclusions

- In the examples considered here, the inclusion of CFactors provided a parsimonious alternative to the standard (unrestricted) use of HB with fewer than 3 C-Factors required.
- With a large number of parameters as in example #1, the results suggest that HB may provide a substantial amount of overfitting, which can be reduced considerably using CFactors to structure the heterogeneity.
- When the number of parameters is fairly small, as in example #2, results suggest that the amount of overfitting in HB may be minimal. However, even with few parameters, the structure added by CFactors was found to simplify the interpretation of the heterogeneity by a substantial margin.
- The use of CFactors in choice models may provide a substantial reduction in estimation time over HB *.

* This conclusion is based on comparisons between Latent GOLD 4.0 (for CFactor models) and Gauss (for HB). Estimation of CFactor models using gllamm was found to be much slower.

Acknowledgment

The authors wish to thank The Kellogg Company for providing the data for example #1, and Wagner Kamakura for the data used in example #2.

References

Andrews, R. L., A. Ainslie, and I. S. Currim (2002). "An empirical comparison of logit choice models with discrete vs. continuous representations of heterogeneity", Journal of Marketing Research, Vol. XXXIX (November), 479-487.

Kamakura, M. Wedel, and J. Agrawal (1994). “Concommitant variable latent class models for conjoint analysis”, *International Journal of Research in Marketing*,11, 451-464.

Natter, M. and M. Feurstein (2002). "Real world performance of choice-based conjoint models", European Journal of Operational Research, 137, 448-458.

Magidson, J., T. Eagle and J. Vermunt, “New Developments in Latent Class Choice Models," 2003 Sawtooth Software Conference Proceedings.

Skrondal, A. & Rabe-Hesketh, S. (2004). Generalized Latent Variable Modeling: Multilevel, Longitudinal and Structural Equation Models. London: Chapman & Hall/CRC.

Vermunt, J.K., and J. Magidson (2005). Technical Guide for Latent GOLD 4.0: Basic and Advanced. Belmont Massachusetts: Statistical Innovations Inc.

APPENDIX

CFactor Regression Model with Random Intercept and Continuous Random Product Attribute Effects

$$Y_{i.t} = \alpha_i + \beta_{i1} Z_1 + \beta_{i2} Z_2 + \ldots + \beta_{iK} Z_K + \varepsilon_{i.t}$$
$$\alpha_i = \alpha_0 + \lambda_{10} F_{i1} + \lambda_{20} F_{i2}$$
$$\beta_{ik} = \beta_{k0} + \lambda_{2k} F_{i2}$$

Thus,

$$E(\alpha_i) = \alpha_0$$
$$V(\alpha_i) = \lambda_{10}^{\ 2} + \lambda_{20}^{\ 2}$$
$$E(\beta_{ik}) = \beta_{k0}$$
$$V(\beta_{ik}) = \lambda_{2k}^{\ 2}$$
$$COV(\beta_{ik}, \beta_{ik'}) = \lambda_{2k} \lambda_{2k'}$$
$$COV(\alpha_i, \beta_{ik}) = \lambda_{20} \lambda_{2k}$$

where: $Y_{i.t}$ is case i’s rating for product t with attributes $Z_1, Z_2, \ldots, Z_K$

C-$Factor\ F_{i1}$ is associated with the intercept

C-$Factor\ F_{i2}$ is associated with the K product attribute effects

$$(F_{i1}, F_{i2}) \sim BVN(0, I)$$