Are We Overfitting Our Models with Too Many Price Parameters?

Prepared by Michael Smith

April 2022

What we will cover today

1 | Intro and Hypothesis

2 | Analysis

3 | Results and Conclusions

4 | Next Steps

Introduction and History

Introduction to Pricing Questions in Conjoint and a Short History of Price Testing

SKIM

Introduction

There are multiple different ways to estimate price parameters in conjoint models. Because of the continuous nature of price, there is a lot of flexibility on how it can be estimated from linear, log-linear, linear + quadratic, or part-worths. A solution leads to more consistent framework from a model fit and parsimony standpoint.

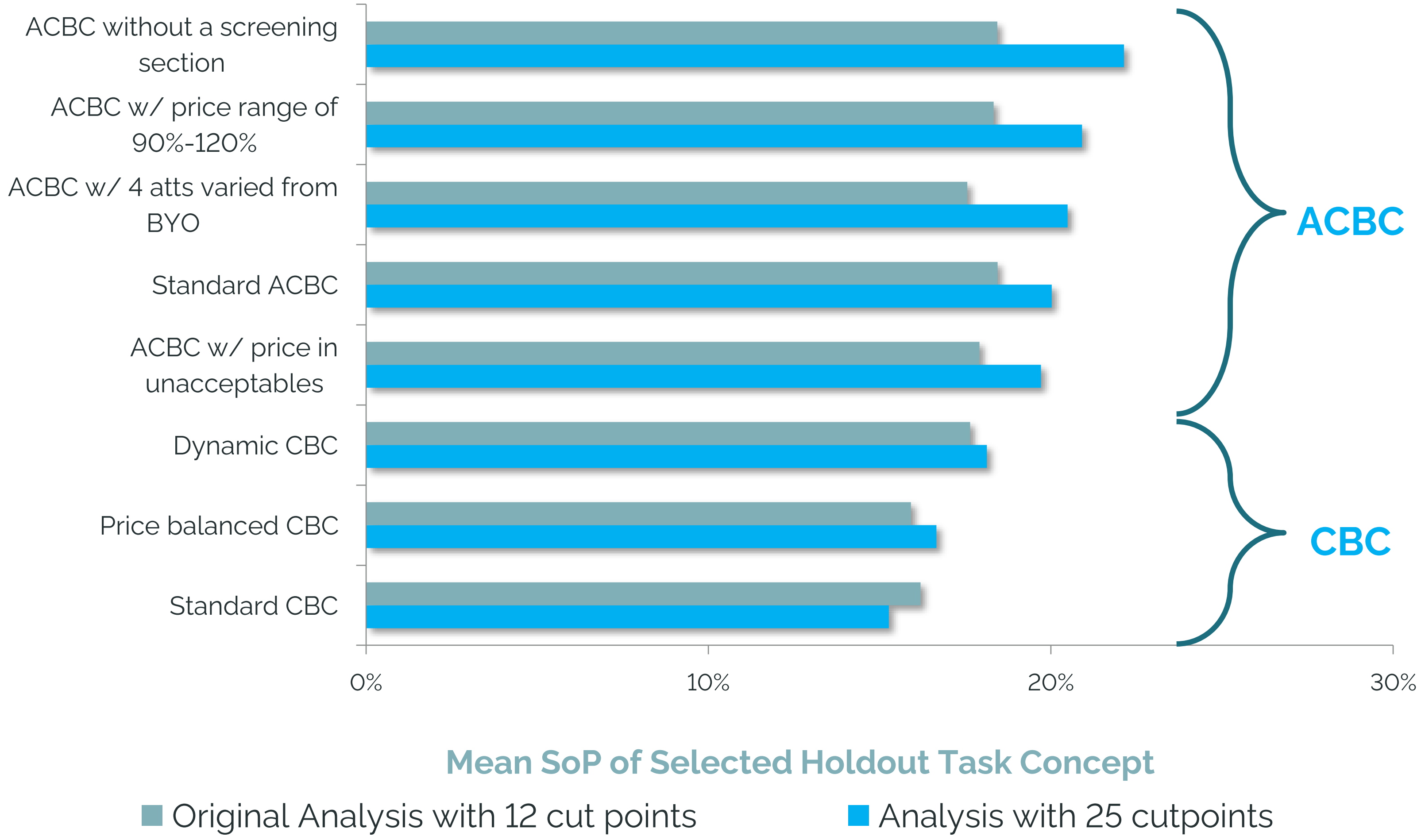
As part of this framework, we determine what the best pricing function is and how many cutpoints are needed.

History

From recent conferences and SKIM's own work, a piecewise function that uses from 2 to 6 breakpoints (aside from the endpoints) is recommended.

Research at the 2013 Sawtooth Software Conference presented by the SKIM Group suggests that even more breakpoints (potentially a dozen or more) is potentially useful for ACBC studies, as long as you constrain price to be negative. The previous work was based on only one study, so we would want to see additional evidence before making a broader recommendation.

Adding more cut points in the piecewise estimation (more granularity in price utilities) significantly increases the performance, especially of the ACBC legs!



SKIM

Hypothesis

How we execute the Pricing Test Investigation

Hypothesis

In a complicated pricing study (using summed pricing or large SKU pricing), we can end up with large variation of price levels tested. We will investigate whether a dozen or more breakpoints is an overfit and whether we are better off with a parsimonious approach.

We will test multiple price approaches on multiple different studies that we have run. We will use Log Likelihood, AIC, BIC, RLH, Holdout Hit Rate, and % of effects that don't need to be constrained.

Datasets Tested for Pricing Comparisons

All projects were run via online panels over the last two years. SKU Pricing normally only has SKU effect, price effect, and maybe promotion. Summed Pricing are normally multi-attribute studies with pricing based on features added.

Project	N=	Attributes	Levels	Unique Prices	Type of Model
Moisturizer	1096	2	30 + 5	113	SKU Pricing
Cleansers	1039	2	30 + 5	111	SKU Pricing
Batteries	1294	2	10 + 5	20	SKU Pricing
Cigarettes #1	3850	2	125 + 9	72	SKU Pricing
Cigarettes #2	2914	2	86 + 7	100	SKU Pricing
Appliances #1	827	8	47	30	Summed Pricing
Appliances #2	821	7	38	37	Summed Pricing

Price Effect Testing Options #1

1 No price

2 With conditional pricing coding without interactions

3 With conditional pricing coding including all interactions

4 With absolute pricing coding, one linear slope

5 With absolute pricing coding, all slopes

6 - #10: With absolute pricing coding, 2, 6, 12, 25, & 50 linear slopes

11 With absolute pricing coding, linear and quadratic effect

12 With absolute pricing coding, linear and quadric effect, all interactions

13 With absolute pricing coding, natural log

14 With absolute pricing coding, natural log, all interactions

SKIM

Price Effect Testing Options #2

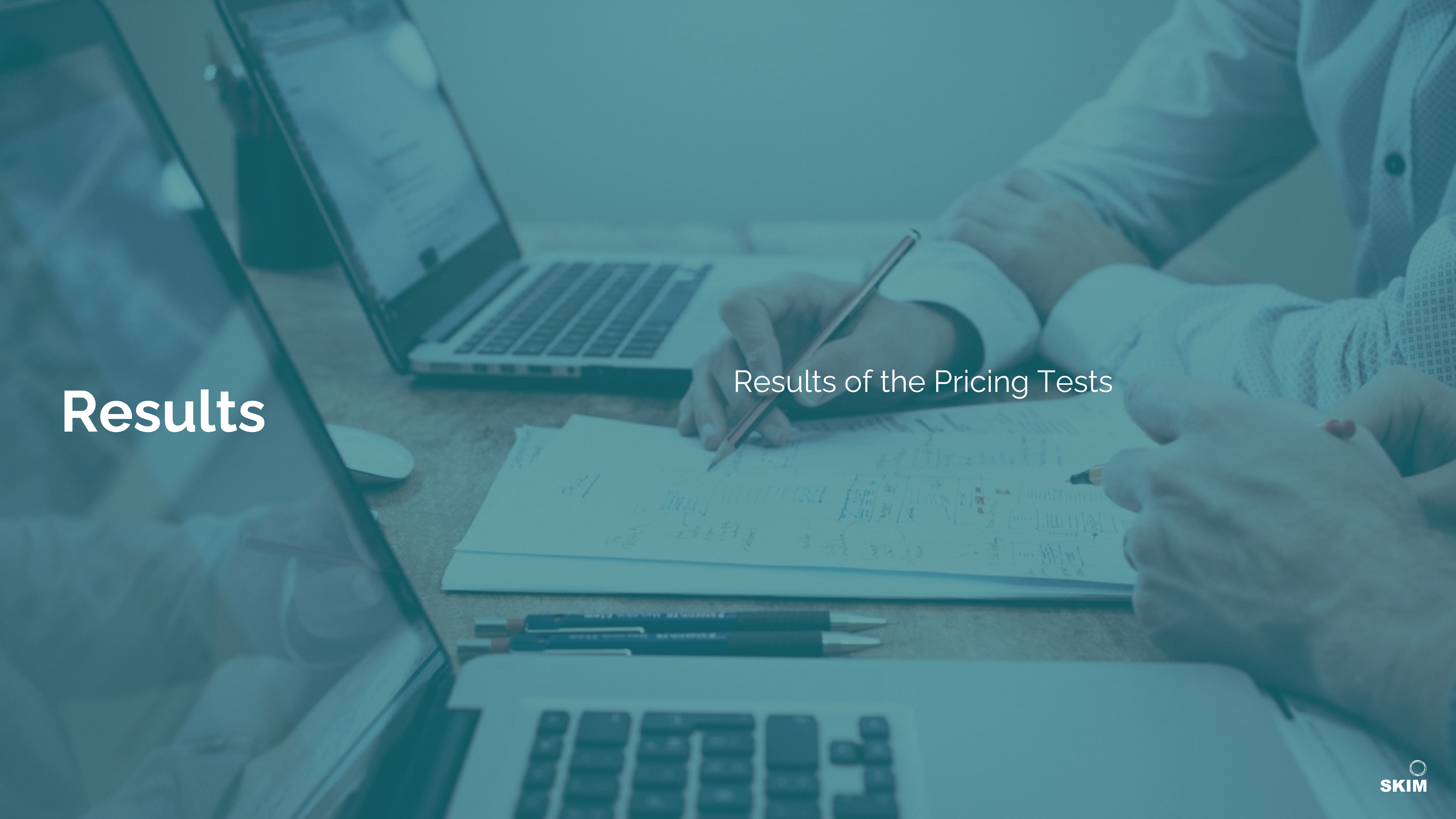
15-30 With absolute pricing coding, y linear slopes with z shock parameters for possible psychological price barriers

- Percentiles were used for all unique absolute prices.
- “Shock" points where handpicked. Shocks are like price-cliffs and are points where we felt were key psychological pricing barriers
- We took, 6, 12 and 25 parameters in total (slopes AND shocks) so you can compare performance with slopes only, given the same number of effects.
- For example, with 6 parameters, we tested "1 shock, 5 slopes", "3 shocks and 3 slopes" and "5 shocks and 1 slope" (that is, if we came up with 5 shocks)
- Same for 12 and 25

EACH OF THESE 30 PRICE TESTS WERE RUN CONSTRAINED AND UNCONSTRAINED

KPIs Used for Analysis

1. "in sample logL", Using the point estimates, the overall loglikelihood of the tasks used for estimation
2. “in sample RLH", Using the point estimates, the average RLH of the respondents based on tasks used for estimation
3. "out of sample logL", Using the point estimates, the overall loglikelihood of the tasks excluded from estimation (the holdout task(s))
4. “out of sample RLH", Using the point estimates, the average RLH of the respondents based on tasks excluded from estimation (the holdout task(s))
5. "summed AE SoP", summed absolute error of the predicted share using share of preference vs. the real chosen shares
6. "summed AE FC", summed absolute error of the predicted share using first choice vs. the real chosen shares
7. In-sample and Out of Sample Hit Rates



Results

Results of the Pricing Tests

Across all 7 studies, % of wrong sign increased as number of parameters increased when comparing unconstrained models

Wrong Sign is defined as a positive coefficient for a price parameter. Each increasing price should have a progressive negative coefficient

But it drastically increases between 5 and 20 and then tapers off

Looking within similar # of parameters, Linear approach is consistently lower % of wrong signs than Linear + Quadratic or Log-Linear

Handpicked shock parameters do lead to lower % of wrong signs

SKIM

SKU PRICING RESULTS

SKU Pricing BIC Analysis

BIC	STUDY 1	STUDY 2	STUDY 3	STUDY 4	STUDY 5
# 1 No price	34816.50111	32550.661	33823.44459	74085.23133	63322.52793
# 2 With conditional pricing coding without interactions	32913.20794	30993.37034	30081.48542	44949.5309	48002.13538
# 3 With conditional pricing coding including all interactions	30414.36734	28181.61699	27678.26973	39106.45173	39620.97861
# 4 With absolute pricing coding, one linear slope	34148.02932	31881.90355	30471.12822	57257.04508	59521.1238
# 5 With absolute pricing coding, all slopes	31556.00731	29535.49151	28615.94477	43155.66977	45912.21145
# 11 With absolute pricing coding, linear and quadratic effect	33973.83674	31909.26746	30447.18054	57315.30117	59439.63524
# 12 With absolute pricing coding, linear and quadric effect, all interactions	32713.83007	30549.43692	29945.49015	64490.53273	55099.02784
# 13 With absolute pricing coding, natural log	34425.97063	32321.11286	32262.39322	63957.90585	60302.65965
# 14 With absolute pricing coding, natural log, all interactions	33999.12572	31859.3169	32305.67771	71151.85998	61436.09619

We calculated Log Likelihood, AIC, and BIC for in-sample fit for all studies (only constrained models illustrated here). A lower AIC or BIC corresponds to a better model fit.

BIC is the best measure that provides the highest penalty for additional price parameters

Linear is better than Linear + Quadratic and Log-Linear

SKIM

SKU Pricing BIC Analysis

BIC improves as the number of slopes increase.

A low to medium number of shocks is best

BIC	STUDY 1	STUDY 2	STUDY 3	STUDY 4	STUDY 5
# 1 No price	34816.50111	32550.661	33823.44459	74085.23133	63322.52793
# 6 With absolute pricing coding, 2 slopes	33708.94279	31611.14224	29738.45603	48019.51994	51377.07941
# 7 With absolute pricing coding, 6 slopes	33244.00387	31069.29533	29284.92017	45043.05066	48867.08377
# 8 With absolute pricing coding, 12 slopes	32938.76903	30631.87215	28895.15848	43963.49248	47747.36288
# 9 With absolute pricing coding, 25 slopes	32436.26868	30298.60426		43376.42013	46929.60731
#10 With absolute pricing coding, 50 slopes	31921.12628	29879.58491		43324.25132	46604.64534
#15 With absolute pricing code, low # of slopes low # of shocks	33919.30346	31454.5354	30245.98561		
#16 With absolute pricing code, med1 # of slopes low # of shocks	33307.60076	31100.47706	29336.67277		
#17 With absolute pricing code, med2 # of slopes low # of shocks	32803.49303	30790.2807	28985.91927		
#18 With absolute pricing code, high # of slopes low # of shocks	32371.03777	30398.64288	28610.03977		
#19 With absolute pricing code, low # of slopes med1 # of shocks	33771.29648	31313.90519	29763.51427		
#20 With absolute pricing code, med1 # of slopes med1 # of shocks	33400.9857	31097.65392	29292.099		
#21 With absolute pricing code, med2 # of slopes med1 # of shocks	32943.53819	30704.26035	28931.50265		
#22 With absolute pricing code, high # of slopes med1 # of shocks	32524.04334	30293.32953	28735.3026		
#23 With absolute pricing code, med1 # of slopes med2 # of shocks	33674.0509	31269.64248		44954.6141	58527.96799
#24 With absolute pricing code, med2 # of slopes med2 # of shocks	33084.00057	30839.19801		44874.78967	48373.2674
#25 With absolute pricing code, high # of slopes med2 # of shocks	32559.06296	30323.56141		43609.4919	47025.36378
#26 With absolute pricing code, med2 # of slopes high # of shocks	32761.8958	30491.65407		43893.14922	57623.28761
#27 With absolute pricing code, high # of slopes high # of shocks	32635.57008	30339.89557			50517.1311

SKU Pricing Out of Sample Hit Rate

OUT OF SAMPLE HIT RATE	STUDY 1	STUDY 2	STUDY 3	STUDY 4	STUDY 5
# 1 No price	39.1%	40.6%	47.6%	65.0%	53.6%
# 2 With conditional pricing coding without interactions	39.7%	40.2%	48.1%	65.4%	54.2%
# 3 With conditional pricing coding including all interactions	40.0%	40.9%	49.7%	64.6%	55.1%
# 4 With absolute pricing coding, one linear slope	39.8%	40.7%	48.5%	65.5%	53.2%
# 5 With absolute pricing coding, all slopes	41.0%	41.3%	49.5%	65.0%	54.6%
# 11 With absolute pricing coding, linear and quadratic effect	39.8%	40.7%	48.5%	65.6%	53.0%
# 12 With absolute pricing coding, linear and quadric effect, all inter	39.8%	40.7%	49.9%	65.3%	53.6%
# 13 With absolute pricing coding, natural log	40.0%	41.0%	47.8%	65.6%	53.5%
# 14 With absolute pricing coding, natural log, all interactions	39.4%	40.1%	47.8%	65.1%	53.6%

We calculated Out of Sample Hit Rate using a Holdout Task (usually connected to a base case configuration

Linear is better than Linear + Quadratic and Log-Linear

Strong Hit Rates just based on SKU parameter estimation. Makes sense if there are a lot of SKU parameters we are estimating

SKU Pricing Out of Sample Hit Rate

Out of Sample Hit Rate seems to peak at 12 to 25 slopes and the addition of shock parameters do not consistently make improvements

OUT OF SAMPLE HIT RATE	STUDY 1	STUDY 2	STUDY 3	STUDY 4	STUDY 5
# 1 No price	39.1%	40.6%	47.6%	65.0%	53.6%
# 6 With absolute pricing coding, 2 slopes	40.0%	40.6%	49.3%	65.5%	54.2%
# 7 With absolute pricing coding, 6 slopes	39.7%	40.5%	49.5%	65.0%	54.7%
# 8 With absolute pricing coding, 12 slopes	39.9%	40.4%	50.5%	65.4%	54.8%
# 9 With absolute pricing coding, 25 slopes	40.3%	41.0%		65.4%	54.9%
#10 With absolute pricing coding, 50 slopes	40.2%	40.7%		65.4%	54.8%
#15 With absolute pricing code, low # of slopes low # of shocks	39.6%	41.5%	48.3%		
#16 With absolute pricing code, med1 # of slopes low # of shocks	39.8%	41.3%	49.8%		
#17 With absolute pricing code, med2 # of slopes low # of shocks	40.1%	41.1%	50.2%		
#18 With absolute pricing code, high # of slopes low # of shocks	39.8%	41.2%	49.9%		
#19 With absolute pricing code, low # of slopes med1 # of shocks	39.3%	40.2%	49.0%		
#20 With absolute pricing code, med1 # of slopes med1 # of shocks	39.8%	41.2%	50.2%		
#21 With absolute pricing code, med2 # of slopes med1 # of shocks	39.6%	41.2%	50.5%		
#22 With absolute pricing code, high # of slopes med1 # of shocks	40.1%	40.5%	50.2%		
#23 With absolute pricing code, med1 # of slopes med2 # of shocks	39.8%	40.9%		65.9%	53.5%
#24 With absolute pricing code, med2 # of slopes med2 # of shocks	40.1%	41.5%		65.8%	54.5%
#25 With absolute pricing code, high # of slopes med2 # of shocks	39.8%	40.4%		65.4%	54.8%
#26 With absolute pricing code, med2 # of slopes high # of shocks	40.3%	41.9%		65.5%	53.6%
#27 With absolute pricing code, high # of slopes high # of shocks	39.1%	41.1%			53.9%

SKIM

CONCLUSIONS ON SKU PRICING

- Very strong model fit from just the SKU parameters and no price parameters. Addition of price parameters only adds a raw 3-4% in hit rate
- 12 to 25 price cutpoints seem to optimize the BIC and out of sample hit rate improvements
- Using a linear effect or a set of linear effects have strong out of sample hit rates for the number of parameters we are estimating
- For SKU pricing models, we would recommend against using too many price parameters (over 15-20) because they don’t add too much value. Use a linear price effect or a set of nested linear price effects for parsimony


SUMMED PRICING STUDIES

Summed Pricing BIC Analysis

BIC is the best measure that provides the highest penalty for additional price parameters. A lower BIC corresponds to a better model fit.

BIC improves as the number of slopes increase.

A medium number of shocks is best. Price cuts do better than personalized shocks

Linear is better than Linear + Quadratic and Log-Linear

BIC	Study 6	Study 7
No price	12513.341	10630.374
With absolute pricing coding, one linear slope (this is one extreme)	11376.802	10436.235
With absolute pricing coding, all slopes	8758.5368	8201.4671
With absolute pricing coding, 2 linear slopes	11114.943	10039.528
With absolute pricing coding, 6 linear slopes	10487.976	9731.181
With absolute pricing coding, 12 linear slopes	9792.3508	9394.2811
With absolute pricing coding, 25 linear slopes	8950.5053	8775.7282
With absolute pricing coding, linear and quadratic effect	11447.958	10339.347
With absolute pricing coding, natural log, all interactions	12264.955	10510.661
With absolute pricing code, med1 # of slopes med # of shocks	10592.375	9748.6222
With absolute pricing code, med2 # of slopes med # of shocks	9898.7171	9420.3286
With absolute pricing code, high # of slopes med # of shocks	8999.1941	8852.9423

Summed Pricing Out of Sample Hit Rate

We calculated Out of Sample Hit Rate using a random Hold Out Task

Out of Sample Hit Rate seems to peak at 12 to 25 slopes and the addition of shock parameters do not consistently make improvements

Linear is better than Linear + Quadratic and Log-Linear but .Linear Slopes are significantly better

Compared to SKU Pricing, Summed Pricing shows significantly higher increase in hit rates

Holdout Hit Rate	Study 6	Study 7
no price	49.9%	48.8%
With absolute pricing coding, one linear slope (this is one extreme)	51.9%	48.7%
With absolute pricing coding, all slopes	53.1%	49.2%
With absolute pricing coding, 2 linear slopes	53.0%	49.8%
With absolute pricing coding, 6 linear slopes	52.9%	49.6%
With absolute pricing coding, 12 linear slopes	53.1%	50.1%
With absolute pricing coding, 25 linear slopes	53.3%	49.1%
With absolute pricing coding, linear and quadratic effect	52.0%	48.6%
With absolute pricing coding, natural log, all interactions	50.4%	47.6%
With absolute pricing code, med1 # of slopes med # of shocks	52.8%	50.2%
With absolute pricing code, med2 # of slopes med # of shocks	53.0%	48.7%
With absolute pricing code, high # of slopes med # of shocks	53.5%	49.3%

SKIM

CONCLUSIONS ON SUMMED PRICING

- Addition of price parameters adds 5-8% in hit rate (double what we see in SKU Pricing)
- 15 to 20 price cutpoints seem to optimize the BIC and out of sample hit rate improvements
- Using a linear effect has strong out of sample hit rates for the number of parameters we are estimating
- Contrary to SKU Pricing, we would recommend using 15 to 20 price cutpoints. The increase in hit rates brings more value than the parsimony of a simple linear effect

Thank you

Michael Smith

VP of Methodology

m.smith@skimgroup.com