

Loss Aversion Estimation in Brand Choice Using Markov Chain Monte Carlo

Haihao Guo*

May 23, 2020

Abstract

I estimate a multinomial logit reference-dependent model with a mixture of normals specification to study the prevalence of reference price effects based on panel data of individual-level purchases in cracker and yogurt market. Prospect Theory and the idea of Loss aversion have recently inspired choice modelers to think about the reference price effect. I use two scanner panel data sets to test for the presence of reference price effects and I conducted a random walk Metropolis MCMC sampler to capture the unobserved heterogeneity and the possible correlation between parameters using a mixture of normals. The result shows that:(1)the degree of loss aversion effect is small and it even disappears after considering the heterogeneity and possible correlations. (2)The promotion activities in the store could possibly make people less sensitive about the price change.

*Master of Computational Social Science, University of Chicago, email:haihao@uchicago.edu

1 Model Setup

1.1 Prospect Theory and Loss aversion

I first introduce the theory of reference-dependent risk-free choice (Tversky & Kahneman, 1979). Considering a choice alternative j with attribute x , then the utility of this choice alternative j could be evaluated from a response function of the deviation of a reference point:

$$U_j(x) = \begin{cases} u_j(x) - u_j(r) & \text{if } u_j \geq r_j \\ \lambda [u_j(x) - u_j(r)] & \text{if } u_j < r_j \end{cases} \quad (1)$$

Where $u_j(x) - u_j(r)$ represents the deviation from the reference point. $\lambda > 1$ specifies loss aversion in this process, that is, consumers are loss averse and that the response function is steeper in the domain of losses than in the domain of gains. Intuitively, one could empirically test the existence of loss aversion by estimating the value of λ to see if that is significantly greater than 1.

To combine this theory into the choice model, there are two specifications that I need to take care of. Firstly, the response function above assumes the reference point only differs between different brands and attributes. However, in the real world, different people have different preferences, and it is more reasonable to specify the heterogeneity for consumers regarding the reference price and the degree of loss aversion, i.e λ . For instance, consumers would have different reference price depends on how price-sensible they are.¹

Secondly, prospect theory only tells the existence but not the origin of the reference point, therefore I need to name a point 'reference point' if I want to empirically test the existence of loss aversion. There are lots of ways to define a reference point in the choice model, (Tversky & Kahneman, 1986) made a comparison between many of them and found that the best fit is the reference point depends on the price paid at last purchase occasion, which is called a memory-based measure.

¹A detailed discussion with an example could be found in (Bell & Lattin, 2000)

I would also use a stimulus-based measure, which measures the reference price as the current price of the last brand bought by the consumer. I use two reference price measures to show that the heterogeneity estimation is robust to the choice of reference price.²

1.2 Choice Model

In this section, I would adapt the reference-dependent model from (Bell & Lattin, 2000) to capture both the reference price effect and loss aversion. Similarly, I also specify the heterogeneity of different consumers and make a little change based on the availability of my data set. The model setting goes like follows.

I assume consumers are all utility maximization consumers, that is, consumers would choose the choice which could give her the maximum utility between alternatives. The reference price and loss aversion effect would be reflected in my framework. First let us develop some notations. U is for utility, i represents the individual i, j represents the brand j, and t represents the occasion, in my data, time. Then, by consuming a good of brand j at time t, consumer i would get the utility of U_{ijt} , that is:

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} \quad (2)$$

where V_{ijt} is the deterministic part of the utility model, which is the evaluation of the brand j for consumer i at time t. ε_{ijt} is the error term of the model, following TYPE-1-Extreme distribution³, which is from the setup of the utility maximization choice model. I implement the reference price and loss aversion effect in the deterministic part V_{ijt} :

$$V_{ijt} = \alpha_{ij} + \beta_{1i} \text{Feat}_{jt} + \beta_{2i} \text{Disp}_{jt} + \beta_{3i} \text{GAIN}_{ijt} + \beta_{4i} \text{LOSS}_{ijt} \quad (3)$$

²This idea is adapted from (Bell & Lattin, 2000)

³Intuitively I can think this way: I would assume people overall have a stable valuation of this good in a certain condition. Imagine this way, when people making a decision about a good with different states, the perception of one good in one state must be considered many times before he finally makes the decision, or in other words, finish the thinking process. Therefore the one with the extreme value would be chosen to be the utility evaluation, therefore I let ε_{ijt} follows TYPE-1-Extreme distribution

where

$$\begin{aligned} \text{GAIN}_{ijt} &= \begin{cases} \text{RP}_{it} - P_{jt} & \text{if } \text{RP}_{it} - P_{jt} > 0 \\ 0 & \text{otherwise} \end{cases} \\ \text{LOSS}_{ijt} &= \begin{cases} P_{jt} - \text{RP}_{it} & \text{if } \text{RP}_{it} - P_{jt} < 0 \\ 0 & \text{otherwise} \end{cases} \\ \text{Disp}_{jt} &= \begin{cases} 1 & \text{if the brand has a display in store} \\ 0 & \text{otherwise} \end{cases} \\ \text{Feat}_{jt} &= \begin{cases} 1 & \text{if the brand has a feature advertisement} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

α_{ij} , β_{1i} , β_{2i} , β_{3i} , β_{4i} are the parameters that I need to estimate. α_{ij} is the intercept for each brand, representing the initial value of brand j to individual i, β_{1i} , β_{2i} , β_{3i} , β_{4i} are the marginal utility of brand's feature, display, gain and loss. $\beta_{3i} < \beta_{4i}$ constitutes the evidence for loss aversion. GAIN and LOSS are the gain or the loss that consumers receive when the price is below or above the reference price. RP_{it} is the reference price for household i at occasion t and P_{jt} is the price of brand j at time t.

After the model is set up, I could use this model to represent the choice making process. Based on the multinomial logit choice model, the probability that consumer i choose brand j at time t can be represented by:

$$P_{it}(j) = \frac{\exp(V_{ijt})}{\sum_k \exp(V_{ikt})} \quad (4)$$

1.3 Heterogeneity Concerns

The literature in marketing proved that the empirical estimation of loss aversion in reference-dependent model has to take uncouncted-for heterogeneity in consumer price responsiveness

and the estimate of loss aversion will decrease after combine price-responsiveness heterogeneity into the reference-dependent model(Bell & Lattin, 2000). I will adapt the reasoning from Bell: A more price-responsive consumer (with a steeper response function) tends to have a lower price level as a reference point. This consumer faces a larger proportion of prices above his reference point, thus the response curve is steeper in the domain of losses. Similarly, the less price-responsive consumer sees a greater proportion of prices below his reference point, so the response curve is less steep within the domain of gains. As a result, any cross-sectional estimate of loss aversion that does not take this into account will be biased upward—researchers who do not control for heterogeneity in price responsiveness may arrive at incorrect substantive conclusions about the phenomenon(Bell & Lattin, 2000).

There are many methods to deal with different degrees of demand heterogeneity(?, ?). The key point here is to use a normal component mixture model of heterogeneity. Considering my need to estimate the parameter for each individual, I use a restricted version of the normal component mixture model: the well-known multinomial logit model and a careful check for convergence of my random walk Metropolis MCMC sampler.

I specify my heterogeneity model across consumers by assuming that $\theta_i = (\alpha_{ij}, \beta_{1i}, \beta_{2i}, \beta_{3i}, \beta_{4i})$ follows a multivariate normal distribution with the mean of $\bar{\theta}$ and the variance of V_θ :

$$\theta_i \sim N(\bar{\theta}, V_\theta)$$

This setting considers the heterogeneity and resolves the IIA-assumption of the model which does not allow for correlations between brand choices by using a full variance-covariance matrix.

I choose to run a random walk Metropolis MCMC Sampler to make the estimation. To let the MCMC Sampler run, I need a prior. I choose to use a natural conjugate prior where the prior on $\bar{\theta}$ is normal and the prior on V_θ is inverted Wishart:

$$V_\theta \sim IW(v, V)$$

In this article, I use the package `bayesm` in R and the function `rhierMnlRwMixture`(MCMC Algorithm for Multinomial Logit with Mixture of Normals Heterogeneity) to make the estimation. Detail of this approach is discussed in (,). According to Rossi et al(2005), "rhierMnlRwMixture is an MCMC algorithm for a multinomial logit with a mixture of normals heterogeneity distribution. This is a hybrid Gibbs Sampler with an RW Metropolis step for the MNL coefficients for each panel unit." I run the model to draw 10,000 samples with the drop of first 1000 as the burn-in. Details of the empirical studies are shown in Section 2.

2 Initial Result

I separate the empirical studies part as follows. First I describe the data set and point out the potential shortcomings. Next, I present and compare my estimation results of the null model(no reference effect), reference-dependent model with homogeneous and with heterogeneity.

2.1 Data

I use two datasets of consumer panel data. The first panel data is provided by Information Resources Incorporated, and it is the optical scanner panel data for saltine cracker purchase in Rome, GA. The second data set is provided by A. C. Nielsen, and it is a scanner panel data for yogurt purchase in Springfield, MO. Both datasets have household consumption records on those product categories and last for 2 years. Therefore, the panel data provides a complete purchase data for individuals over time, which enables us to get both stimulus-based reference price and the memory-based reference price for each consumer. However, the panelists do not coincide therefore I cannot do a cross-category analysis.

The first data set of Cracker market contains 3292 purchase records of 137 households and 4 major brands (including private labels). The brand consist of 3 major brands, Sunshine,

Table 1: Descriptives of the Cracker market(136 panelists, 3,292 purchases)

Variable	Brand			
	Sunshine	Kleebler	Nabisco	Private Brand
Av.Display	0.129	0.106	0.340	0.099
Av.Feature	0.038	0.043	0.087	0.047
Av.Price	0.957	1.126	1.079	0.681
Market Share	0.073	0.069	0.544	0.314

Keebler, and Nabisco, and several local brands that are counted as 'private' in this data set. Their market shares are 7.3%, 6.9%, 54.4%, and 31.4% respectively. The data set also provides on each purchase occasion the price of 4 brands, whether there is a display or not in the store, whether there is an advertisement feature or not. Table 1 reports the market shares (choice shares among the panelists) of the 4 brands, the mean value of Display, Feature, and Price.

Table 2: Basic Descriptive Statistics: Yogurt (100 panelists, 2,412 purchases)

Variable	Brand			
	Yoplait	Dannon	Weight Watchers	Hiland
Av.Feature	.056	.038	.038	.037
Av.Price	.107	.082	.079	.054
Market Share	.339	.402	.230	.029

In the Yogurt market, the data set contains 2412 purchase records of 100 households and 4 major brands (including private labels) capture 4 brands in this market: Yoplait, Dannon, Weight Watchers, and Hiland. Their market share are 33.9%, 40.2%, 23%, and 2.9% respectively. The data set also provides on each purchase occasion the price of 4 brands, whether there is an advertisement feature or not. Table 2 reports the market shares (choice shares among the panelists) of the 4 brands, the mean value of Feature and Price.

2.2 Reference-dependent model assuming homogeneity

Table 3 and Table 4⁴ presents the model fits and parameter estimates for three models: (1) A null model with the main effect for the price, (2) a reference-dependent model that relies on stimulus-based reference points, and (3) a reference-dependent model that relies on memory-based reference points. I assume homogeneous consumers⁵ in the reference-dependent model, and I use a classical Maximum likelihood estimation method to do the estimation. In particular, I use the package `mlogit` in R and the function `mlogit` to get the result.

As shown in both Table 3 and Table 4, that the mean promotion estimation, i.e feature advertisement/display value is significantly positive, which indicates that brands can affect consumer choices by promotion in the Cracker and Yogurt Market. The price estimation of the null model with the main effect for price indicates that price has a significant effect on brand choice.

Table 3: Cracker Market Estimation Results

	Null	Reference-Dependent Model	
		Memory-based	Stimulus-based
nabisco:(intercept)	1.962 (0.072)***	2.053 (0.074)***	1.681 (0.074)***
private:(intercept)	0.169 (0.117)	0.788 (0.117)***	0.397 (0.117)***
sunshine:(intercept)	-0.494 (0.101)***	-0.306 (0.102)**	-0.370 (0.101)***
price	-3.125 (0.209)***	-	-
disp	0.092 (0.062)	0.285 (0.071)***	0.277 (0.074)***
feat	0.496 (0.095)***	0.652 (0.112)***	0.679 (0.117)***
gain	-	-2.801 (0.279)***	-1.773 (0.270)***
loss	-	-6.585 (0.262)***	-8.078 (0.319)***
AIC	6707.427	5390.533	5024.146
Log Likelihood	-3347.713	-2688.267	-2505.073
Num. obs.	3292	3292	3292

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

The loss aversion effect specified by the gain and loss parameters exists significantly in

⁴I used the package called 'texreg' in R to get make the table used in this article.(?, ?)

⁵homogeneous consumers means the marginal effect of price, loss, gain, feature advertisement, and display on all the consumers in the dataset are the same.

both Cracker and Yogurt Market, which is consistent with the finding in the orange juice market that the loss aversion exists without considering the heterogeneity (Hardie, Johnson, & Fader, 1993). In Cracker Market, the loss aversion for the memory-based reference-dependent price is $((-8.078)/(-1.773) = 4.56)$ and the loss aversion for the stimulus-based reference-dependent price is $((-6.585)/(-2.801) = 2.35)$. In Yogurt Market, the loss aversion for the memory-based reference-dependent price is $((-1.024)/(-0.158) = 6.48)$ and the loss aversion for the stimulus-based reference-dependent price is $((-1.099)/(-0.266) = 4.13)$. They are all significantly above 1. This also confirms the fact that choices are affected by the deviation from the reference point. The results are robust to the choice of the reference price.

Table 4: Yogurt Market Estimation Results

	Null	Reference-Dependent Model	
		Memory-based	Stimulus-based
hiland:(intercept)	-3.716 (0.145)***	-3.414 (0.164)***	-2.713 (0.166)***
weight:(intercept)	-0.641 (0.054)***	-0.616 (0.055)***	-0.572 (0.056)***
yoplait:(intercept)	0.735 (0.081)***	1.246 (0.081)***	0.898 (0.088)***
price	-0.367 (0.024)***	-	-
feat	0.491 (0.120)***	0.606 (0.136)***	0.860 (0.149)***
gain	-	-0.158 (0.028)***	-0.266 (0.034)***
loss	-	-1.024 (0.037)***	-1.099 (0.045)***
AIC	5323.776	4230.342	3926.072
Log Likelihood	-2656.888	-2109.171	-1957.036
Num. obs.	2412	2412	2412

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

2.3 Reference-dependent model assuming heterogeneity

I now add heterogeneity into the reference-dependent model and use the random walk Metropolis MCMC sampler to do the estimation. I want to test if heterogeneity could make the estimation of loss aversion smaller in Cracker and Yogurt market like in the orange juice market (Bell & Lattin, 2000). I use the `bayesm` package in R to run a random walk Metropolis MCMC Sampler on the Multivariate Logit Model and make the estimation. I

draw 10,000 MCMC samples and drop the first 1,000 as the burn-in. Based on 9,000 valid draws, Table 5, Table 6 show the results of posterior mean of estimated parameters.

The mean gain and loss aversion parameters are consistent with reference-dependent choice behavior. Choices are affected by the deviation of the actual prices from the reference point. However, with heterogeneity, I find less loss aversion than without heterogeneity. In Cracker Market, the loss aversion for the memory-based reference-dependent price is $((-2.485)/(-2.625) = 0.95)$ and the loss aversion for the stimulus-based reference-dependent price is $((-12.004)/(-11.242) = 1.07)$. In Yogurt Market, the loss aversion for the memory-based reference-dependent price is $((-0.465)/(-0.802) = 0.58)$ and the loss aversion for the stimulus-based reference-dependent price is $((-1.685)/(-1.746) = 0.97)$. The values of λ in all cases are not significantly above 1. Compared with Table 3 and 4, I find that considering both heterogeneity and allowing for correlations between different parameters in the multinomial logit model, which breaks the I.I.An assumption, could yield a lower loss aversion estimation in both Cracker and Yogurt markets. With taking heterogeneity in consideration, I got the same result as previous literature(Bell & Lattin, 2000; Klapper, Ebling, & Temme, 2005), that the loss aversion effect is small or disappear, which shows that loss aversion may only be a phenomenon of not adequately accounting for consumer heterogeneity in the parameters.

Table 5: Posterior Means - Cracker Market

	Memory-based	Stimulus-based
feature	0.758 (0.03)	0.908 (0.04)
display	0.982 (0.04)	0.958 (0.04)
gain	-2.625 (0.15)	-11.242 (0.55)
loss	-2.485 (0.20)	-12.004 (0.27)
Num. obs.	3292	3292

Table 7, Table 8, Table 9, and Table 10 shows the posterior means of standard deviations and correlation matrix for both Cracker and Yogurt Market. One advantage of using MCMC sampler to solve this problem is that it breaks the I.I.A and enables us to detect

Table 6: Posterior Means - Yogurt Market

	Memory-based	Stimulus-based
feat	1.430 (0.04)	1.529 (0.04)
gain	-0.802 (0.02)	-1.746 (0.05)
loss	-0.465 (0.02)	-1.685 (0.06)
Num. obs.	2412	2412

the correlation between different parameters and I found some interesting results. Table 8, Table 9, and Table 10 shows a positive correlation between the marginal effect of gain and loss. This is consistent with the intuition that reference price effect and loss aversion are connected to the price sensitivity. Normally, if a consumer is more sensitive to the price level, she would be sensitive to both gain and loss regarding her reference point. Table 8, Table 9, and Table 10 also shows a negative correlation between the marginal effect of promotion, including display and feature advertising, and the reference price effect. Though I cannot directly get the causal relationship from this simple look, I could find a possible explanation for this. When a store gives promotion activity, consumers would pay more attention to the promotion activity, therefore the marginal effect of price on their decision goes down. This explanation is consistent with the salience theory and attention bias-related literature in both psychology and behavioral economics.

Table 7: Correlation Matrix-Cracker Market (Memory-based)

	Std Dev	feature	display	gain	loss
feature	0.97	1.00	0.23	-0.38	0.40
display	1.12	0.23	1.00	-0.40	0.71
gain	4.55	-0.38	-0.40	1.00	-0.71
loss	6.28	0.40	0.71	-0.71	1.00

Summarizing, the key results in the Yogurt and Cracker market are interesting and they

Table 8: Correlation Matrix-Cracker Market (Stimulus-based)

	Std Dev	feature	display	gain	loss
feature	1.2	1.000	0.35	-0.40	-0.014
display	1.2	0.346	1.00	-0.40	-0.131
gain	16.6	-0.399	-0.40	1.00	0.791
loss	8.0	-0.014	-0.13	0.79	1.000

Table 9: Correlation Matrix-Yogurt Market (Memory-based)

	Std Dev	feature	gain	loss
feature	1.33	1.000	-0.27	0.047
gain	0.70	-0.272	1.00	0.388
loss	0.65	0.047	0.39	1.000

extend the findings of Bell and Lattin in that (on average) loss aversion hardly exists into the market of Yogurt and Cracker when I allow for individual-specific parameters in the utility function using a Multinomial Logit model. Using the MCMC sampler, I also detect the correlation between different parameters, and I found out that overall, the marginal effect of promotion activity of a company(i.e feature advertisement or display) has a negative correlation with the marginal effect of gain or loss, which could bring some insight for both future research and managerial decision.

Table 10: Correlation Matrix-Yogurt Market (Stimulus-based)

	Std Dev	feature	gain	loss
feature	1.3	1.00	-0.24	-0.29
gain	1.9	-0.24	1.00	0.85
loss	1.7	-0.29	0.85	1.00

References

- Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, 63(6), 1657–1672.
- Barseghyan, L., Prince, J., & Teitelbaum, J. C. (2011). Are risk preferences stable across contexts? evidence from insurance data. *American Economic Review*, 101(2), 591–631.
- Bell, D. R., & Lattin, J. M. (2000). Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity. *Marketing Science*, 19(2), 185–200.
- Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics*, 112(2), 407–441.
- Chang, K., Siddarth, S., & Weinberg, C. B. (1999). The impact of heterogeneity in purchase timing and price responsiveness on estimates of sticker shock effects. *Marketing Science*, 18(2), 178–192.
- Engström, P., Nordblom, K., Ohlsson, H., & Persson, A. (2015). Tax compliance and loss aversion. *American Economic Journal: Economic Policy*, 7(4), 132–64.
- Fibich, G., Gavious, A., & Lowengart, O. (2003). Explicit solutions of optimization models and differential games with nonsmooth (asymmetric) reference-price effects. *Operations Research*, 51(5), 721–734.
- Genesove, D., & Mayer, C. (2001). Loss aversion and seller behavior: Evidence from the housing market. *The quarterly journal of economics*, 116(4), 1233–1260.
- Hardie, B. G., Johnson, E. J., & Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing science*, 12(4), 378–394.
- Homonoff, T. A. (2018). Can small incentives have large effects? the impact of taxes versus bonuses on disposable bag use. *American Economic Journal: Economic Policy*, 10(4), 177–210.
- Jain, D. C., Vilcassim, N. J., & Chintagunta, P. K. (1994). A random-coefficients logit brand-choice model applied to panel data. *Journal of Business & Economic Statistics*, 12(3), 317–328.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of political Economy*, 98(6), 1325–1348.
- Kalwani, M. U., Yim, C. K., Rinne, H. J., & Sugita, Y. (1990). A price expectations model of customer brand choice. *Journal of Marketing research*, 27(3), 251–262.
- Klapper, D., Ebling, C., & Temme, J. (2005). Another look at loss aversion in brand choice data: Can we characterize the loss averse consumer? *International Journal of Research*

- in Marketing*, 22(3), 239–254.
- Kopalle, P. K., Rao, A. G., & Assuncao, J. L. (1996). Asymmetric reference price effects and dynamic pricing policies. *Marketing Science*, 15(1), 60–85.
- Lattin, J. M., & Bucklin, R. E. (1989). Reference effects of price and promotion on brand choice behavior. *Journal of Marketing research*, 26(3), 299–310.
- McFadden, D., et al. (1973). Conditional logit analysis of qualitative choice behavior.
- Paap, R., & Franses, P. H. (2000). A dynamic multinomial probit model for brand choice with different long-run and short-run effects of marketing-mix variables. *Journal of Applied Econometrics*, 15(6), 717–744.
- Pope, D. G., & Schweitzer, M. E. (2011). Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, 101(1), 129–57.
- Post, E., Pedersen, C., Wilmers, C. C., & Forchhammer, M. C. (2008). Warming, plant phenology and the spatial dimension of trophic mismatch for large herbivores. *Proceedings of the Royal Society B: Biological Sciences*, 275(1646), 2005–2013.
- Savage, L. J. (1954). The foundations of statistics. *John Wiley, New York*. [4] Schmeidler, D. (1989) "Subjective Probability and Expected Utility without Additivity", *Econometrica*, 57, 571–587.
- Tereyağoglu, N., Fader, P. S., & Veeraraghavan, S. (2018). Multiattribute loss aversion and reference dependence: Evidence from the performing arts industry. *Management Science*, 64(1), 421–436.
- Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, 106(4), 1039–1061.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4), 297–323.
- Von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior, 2nd rev.
- Winer, R. S. (1986). A reference price model of brand choice for frequently purchased products. *Journal of consumer research*, 13(2), 250–256.