Sawtooth Software

RESEARCH PAPER SERIES

The Random Regret Minimization Choice Modeling Paradigm: An Introduction with Empirical Tests

Keith Chrzan and Jefferson Forkner
Sawtooth Software, Inc.

1457 E 840 N Orem, Utah
+1 801 477 4700
www.sawtoothsoftware.com

The Random Regret Minimization Choice Modeling Paradigm:
An Introduction with Empirical Tests

Keith Chrzan and Jefferson Forkner
Sawtooth Software, Inc.

Background

For 40 years choice modelers have relied upon the linear-in-parameters Random Utility paradigm for understanding human choice behavior. And for good reason:

- The Random Utility Model (RUM) conforms well with the conventional economic assumptions that consumers are rational, utility-maximizing decision makers
- It provides a mathematical framework describing those decisions and even for allowing those decision makers to be human enough to act inconsistently from time to time
- RUM's mathematical formulation is simple enough to appear in many general purpose statistical programs but flexible enough to cover a wide variety of decision making situations
- RUM has been found to work well in practice: applied marketing researchers, for example, use it to gain valuable insights about marketing scenarios – insights which often hold up well when compared to results in real markets

Such success notwithstanding, some of the assumptions RUM makes about human decision making have been called into question. Behavioral economists, for example, seem to delight in pointing out the irrational, non-utility maximizing decisions even smart humans make; they have cataloged scores of decision biases which produce systematic deviations between the way humans make judgments and the way calculating utility maximizers would judge (Kahneman, Slovick and Tversky 1982; Kahneman 2011). Behavioral economists have also done a good job of explaining their work entertainingly for non-scientists through the work of Thaler and Sunstein (2009) and Ariely (2010), among others.

More recently, neuroscientists' ability to monitor brain activity in real time using functional magnetic resonance imaging (fMRI) technology has opened the door to physiological understanding of choice behaviors. Some of this literature supports RUM (Webb et al 2013) while other studies confirm the insights of the behavioral economists. Take for example, loss aversion, the fact that we mourn the loss of a given amount of money, say, more than we celebrate a financial gain of the same size (Kahneman and Tversky 1979): neuroscientists have identified neurological loss aversion, brain activity that corresponds to behavioral loss aversion, even down to the neurological activity having a greater intensity for losses than for gains (Tom *et al* 2007). Other studies go further: Coricelli *et al* (2005) suggest that losses figure so strongly into our decision making that the avoidance of feeling regret about them powerfully motivates choice in its own right. Obviously the choices that maximize utility and those that minimize regret may not be the same.

A new choice paradigm available to academic and applied choice modelers is Random Regret Minimization (RRM). RRM posits a different psychological mechanism behind choice behavior, namely that a chooser selects the alternative that minimizes her chance for regretting her decision. Caspar Chorus, a scholar at the Delft University of Technology, is behind a growing body of work on RRM, including papers introducing and testing RRM (Chorus 2010, 2012a, 2012b).

It turns out that one can take an existing discrete choice experiment (choice-based conjoint analysis) data designed for analysis as a RUM model and recode it into an RRM model estimable using standard multinomial logit (MNL) software. One can thus run RRM using aggregate logit, latent class MNL or even hierarchical Bayesian MNL software. One could use Sawtooth Software's programs to do the work, or any other HB or logit program with similar capabilities.

Following an introduction of RRM and some of its interesting properties we will illustrate the variable recoding necessary to convert a standard choice-based conjoint model into an RRM model. We report results of some empirical comparisons of the predictive validity of CBC using both RRM and RUM. Finally we offer practical advice to applied researchers, noting the strengths and weaknesses of RRM relative to RUM.

Bringing RRM to Life

RRM can utilize existing data from a choice-based conjoint (CBC) model, though it may be that the optimal designs for RRM could differ from optimal designs for RUM-based CBC models.

In standard RUM analysis of CBC experiments we use effects coding to represent the levels of categorical variables and (often) numerical coding to represent the linear or non-linear effects of quantitative variables. In RRM, however, the variables code the differences between an attribute level in a given product profile and the total regret of not being able to choose the more attractive levels for that attribute in the choice set. An example may help. Consider a choice set in which three products A, B and C have prices of $1, $3 and $6, respectively. Product C has a $5 higher price than product A ($5 worth of regret) and a price $3 higher than product B ($3 of regret). The total regret for Product C's price variable is (-$5+-$3) = -$8. Product B has $2 of regret relative to Product A but none relative to higher-priced Product C, so its regret is (-$2+$0) = -$2. Finally, Product A has no regret relative to the prices of either Product B or Product C, so its price regret is $0. Thus the coding of the price variable under a standard RUM-based CBC and RRM is as follows for products A-C above:

Price	RUM coding	RRM coding
$1	1	0
$3	3	-2
$6	6	-8

For categorical variables RRM resembles effects coding in standard CBC models – in fact RRM and RUM coding are perfectly correlated. This means that for studies that have only categorical variables, RRM and RUM will deliver equivalent results. Again an example will help: imagine the categorical variable packaging color with three levels orange, green and purple. Just as in effects or dummy coding we use two new variables to capture the three levels of our categorical variable. Let one be Orange Regret and the other Green Regret. Product A is orange, so it has no amount of regret for not being orange relative to Product B (which is green) or Product C (which is purple). Product A's Orange Regret is 0. When it comes to green, however, Product A has one unit of regret relative to Product B (because Product B is green and Product A isn't) but no regret relative to Product C (because neither is green). So Product A's value for Green Regret is (-1+0) = -1. By similar reasoning Product B has -1 Orange Regret and 0 Green Regret (it's green and the others aren't). Finally, Product C, which is purple, has an Orange Regret of -1 and a Green Regret of -1. RRM coding for the color attribute of the three products would be:

	Orange Regret	Green Regret
Product A (orange)	0	-1
Product B (green)	-1	0
Product C (purple)	-1	-1

Chorus (2010) shows that one can estimate RRM models coded as above using standard MNL software packages like SAS, SPSS or LIMDEP. Sawtooth Software users can estimate MNL, Latent Class MNL or hierarchical Bayesian (HB) MNL models for RRM by employing user-specified coding of the variables (where the researcher specifies the values to be taken into the design matrix rather than letting the software automatically code the design matrix for effects coding).

Interesting Properties of RRM Models

At first glance RRM may appear to be no more than a reworking of standard CBC coding, one which will give equivalent results (and this will happen if one has only categorical variables in the model). It turns out, however, that RRM's coding gives it a couple of interesting properties, ones that could prove advantageous over standard CBC coding, in some circumstances.

First, RRM is "semi-compensatory." For example, consider a choice set consisting of several automobiles. Automobile A has the best gas mileage but slightly worse trunk space than do the other automobiles. Automobile B has the worst MPG and a middling level of trunk space.

Small improvements in Automobile B's MPG can compensate for its deficit in trunk space (by reducing MPG-regret) but no amount of improvement in Automobile A's MPG can compensate for its trunk space deficit as it has no MPG-regret in the first place. This represents a major difference between standard CBC and RRM. RRM also has a compromise effect that can allow alternatives with attributes at intermediate levels of utility to perform better in some RRM simulations than in standard RUM CBC simulations. Whether this property helps or hurts the predictive validity of RRM relative to standard CBC and how often it does so we do not know.

An even more interesting property of RRM is that, uniquely, it separates the MNL model's independence of irrelevant alternatives (IIA) property from the red bus/blue bus problem (so though the red bus/blue bus case illustrates how IIA operates in RUM-CBC it does not serve this function in RRM) . IIA posits that the relative share between two alternatives does not change with the addition of a third product. The classic example has a bus company whose busses are all blue competing against a train system for customers. Perhaps the bus company has a 60% share and the train company 40%. Now a second bus company, one with red busses begins service and offers the same fares, pickup locations, pickup and transit times as the blue bus company. Under IIA, the 60/40 ratio of blue bus to train must be maintained, so we might predict 37.5%, 37.5% and 25% for blue bus, red bus and train respectively. This property is restrictive and often false but it is part of the standard RUM version of the logit model.

Of course the red bus company will likely take much more share from the blue bus company and little or none at all from the train company. As a result the blue bus/train relative shares of 60/40 might more realistically shrink more toward 30/40.

Under RRM, however, IIA does NOT hold. To illustrate, imagine a choice among transportation options. Decision maker Jones has these utilities for price, transit time and mode:

Attribute/level	Utility
Train	1
Bus	0
Trolley	0
Per $1	-1
Per 1 minute	-0.2

In one scenario Jones can choose a $1, 15 minute bus ride or a $2, 10 minute train ride:

- The utility for the bus is -1 (mode regret) + 0 (price regret) – 1 (5 minutes worth of time regret) = -2.
- For train we have 0 (mode regret) + -1 (price regret) + 0 (time regret) = -1
- So the likelihood that Jones chooses the bus is exp(-2)/[exp(-2)+exp(-1) = 26.9%
- The likelihood for train would be the complement, 73.1%.

Now we add in an 11 minute, $1.50 trolley. Now regrets and choice probabilities are as below:

	Bus	Train	Trolley
Price	$1.00	$2.00	$1.50
Time	15min	10 min	11 min
Bus regret	0	0	0
Train regret	-1	0	-1
Price regret	0	-1.5	-0.5
Time regret	-1.8	0	-0.1
Total regret	-2.8	-1.5	-1.6
Exponentiated regret	0.06	0.22	0.20
Choice probability	12.5%	45.9%	41.6%

Notice that the trolley takes disproportionately more share from buss than from train. As a result the proportion between bus and train in the bus/train scenario was .269/.731 = .368 but in the bus/train/trolley scenario the ratio is .125/.459 = .272. RRM thus does not require IIA.

Should a second bus line enter the market instead of the trolley, however, with the same transit time and price as the first bus line, shares for the two bus lines would be equal and the bus services will illogically take share from the train. Thus RRM still suffers from a version of the red bus/blue bus problem even if not from IIA.

Limitations of RRM Coding

While RRM has some interesting capabilities standard RUM/CBC lacks, it also suffers from some limitations. First, as noted above, if a study contains only categorical variables, RRM and RUM will produce equivalent results: RRM can only add value if it has quantitative attributes to model as linear functions.

Secondly, it turns out that if choice sets have only two alternatives, then RRM and RUM coding again produce equivalent results. So RRM only adds value if choice sets have three or more alternatives.

It could be that what drives decision-making isn't total regret as in RRM but rather maximum regret (the regret measured by the difference between the utility of an attribute level available in a given alternative and the most attractive level of that attribute available in a different alternative. Unfortunately, if we use a maximum regret coding rather than total regret coding, it turns out that RRM again returns the same results as standard RUM/CBC coding.

Of course ultimately we care about the value of the insights and predictions we can get from RRM and how they compare to a standard RUM/CBC model formulation. For this we turn to empirical comparisons.

Empirical Comparisons

<u>Study 1</u>
Research design

In March 2012, 1,204 owners/intenders of products in an unnamed category completed a multi-celled web-based CBC survey. One cell of 457 respondents completed a 3^7 x 2 CBC experiment in 18 choice sets of triples (three alternatives per set). The experimental design maximized D-efficiency and thus minimized level overlap. All 1,204 respondents also answered six holdout choice questions, also triples and also in a format shared by the CBC questions. Thus each holdout question had 457 in-sample and 747 out-of-sample observations.

Using both RRM and RUM/CBC coding and hierarchical Bayesian logit model estimation using the CBC/HB program produced two models we can compare in terms of their model fit and their predictions of holdouts.

Results

Differences in magnitudes (and in the case of price, the direction) of utilities owe to the distinct coding of the two models and thus do not reveal an advantage for one model over the other:

Table 1
HB Utilities for RUM/CBC and RRM Models in Study 1

	RUM	RRM
Brand 1	-0.561	-1.108
Brand 2	-0.545	-1.013
Attribute 1	0.094	0.114
Attribute 2	0.041	0.021
Price	-0.15	0.21
Attribute 3	0.061	0.119
Attribute 4	0.067	0.086
Attribute 5	0.084	0.099
Size1	-0.168	-0.358
Size2	0.104	0.216

The RUM/CBC model fits calibration data slightly better than does the RRM model with log likelihoods (LL) of -8,092 and -8,258, respectively, or root likelihoods (RLH) of 0.374 and 0.366 respectively.

The story reverses, however, for holdout predictions. The RRM model predicts the 18 holdout choice shares more accurately (mean absolute error, MAE, of 3.3%) than does the RUM/CBC model (MAE = 4.5%).

Table 2
Holdout Predictions for Study 1

Set	Alt	Actual	RUM Prediction	RRM Prediction		RUM Error	RRM Error
1	1	31.1%	30.5%	34.6%		0.6%	3.5%
1	2	56.7%	50.3%	48.1%		6.4%	8.6%
1	3	12.2%	19.2%	17.3%		7.0%	5.1%
2	1	26.7%	25.0%	26.7%		1.7%	0.0%
2	2	23.8%	23.8%	26.8%		0.0%	3.0%
2	3	49.5%	51.2%	46.5%		1.7%	3.0%
3	1	33.5%	40.7%	33.6%		7.2%	0.1%
3	2	33.8%	25.6%	29.8%		8.2%	4.0%
3	3	32.7%	33.7%	36.6%		1.0%	3.9%
4	1	57.1%	41.2%	45.2%		15.9%	11.9%
4	2	14.8%	23.0%	17.0%		8.2%	2.2%
4	3	28.2%	35.8%	37.7%		7.6%	9.5%
5	1	11.7%	13.7%	12.1%		2.0%	0.4%
5	2	26.3%	21.9%	24.7%		4.4%	1.6%
5	3	62.0%	64.4%	63.2%		2.4%	1.2%
6	1	8.8%	12.5%	9.2%		3.7%	0.4%
6	2	20.0%	17.5%	19.4%		2.5%	0.6%
6	3	71.2%	70.0%	71.4%		1.2%	0.2%
					MAE	4.5%	3.3%

The RRM model's predictions also correlate more highly with the actual holdout choice shares (r=0.967) than do the RUM/CBC model's predictions (r=0.945). This difference produces a t statistic of 3.81, significant at p < 0.002.

Thus in Study 1 in-sample model fit seems to favor the RUM/CBC model but this does not generalize well to (mostly) out-of-sample holdout predictions where RRM has a significant edge.

Study 2

Research design

A total of 1,247 tablet PC owners/intenders completed a web-based survey in May of 2013. Attributes and levels of the $5^2 \times 4^2 \times 3^2$ experiment were as follows:

Table 3
Attribute and Levels for Study 2

Attribute	Level 1	Level 2	Level3	Level 4	Level 5
Brand	Google Nexus	Amazon Kindle Fire	Apple iPad	Samsung Galaxy	Microsoft Surface
Price	$169	$199	$299	$399	$499
Screen Size	7 inch	8 inch	9 inch	10 inch	
Storage	16 GB	32 GB	64 GB	128 GB	
RAM	1 GB	2 GB	4 GB		
Battery Life	6 hours	7 hours	8 hours	9 hours	10 hours

Each of 931 respondents answered a calibration set of 12 CBC triples plus 3 holdout triples not used in utility estimation. Half of respondents received choice sets constructed with a minimum overlap and the other half had a design that allowed level overlap, an aspect of our experiment that turned out to make no difference to the results reported below. A second cell of 316 respondents completed 15 triples, constructed completely at random, with plenty of level overlap that would make them difficult to predict. This cell serves as a holdout for assessing the out-of-sample validity of RRM and RUM/CBC.

Results

As in Study 1 the magnitude (and for price the direction) of utilities depends on the differences in the two models' coding and does not by itself tell us about their relative quality:

Table 4
MNL and HB Utilities for RUM/CBC and RRM Models in Study 2

	RUM (MNL)	RRM (MNL)		RUM (HB)	RRM (HB)
Google Nexus	-0.31	-0.62		-0.60	-1.09
Amazon Kindle Fire	-0.10	-0.17		-0.19	-0.36
Apple iPad	0.58	1.15		1.18	2.19
Samsung Galaxy	0.03	0.06		0.07	0.12
Microsoft Surface	-0.20	-0.42		-0.46	-0.87
7 inch screen	-0.14	-0.14		-0.37	-0.63

8 inch screen	-0.06	0.13		-0.10	-0.19
9 inch screen	0.07	0.27		0.18	0.31
10 inch screen	0.13	1.00		0.28	0.52
Storage	0.48	0.73		1.12	1.66
RAM	0.14	0.22		0.31	0.47
Battery Life	0.05	0.07		0.13	0.17
Price	-0.41	0.51		-1.09	1.48

Again the RUM/CBC model fits the calibration data better (LL = -12,880, RLH = 0.660) than does the RRM model (LL = -12,979, RLH = 0.643).

In-sample predictions, measured by hit rates, favor the RUM/CBC model (57.9%) over the RRM model (56.3%). A dependent t-test for this difference in hit rates shows a t statistic of 2.91, significant at $p < 0.004$).

For out-of-sample validity, the correlation of RUM/CBC model predictions with actual holdout shares is 0.9355, which exceeds that for the RRM model (0.9203) though to an extent that is not significant ($t = 1.70$).

With respect to predictive validity we see a reversal in Study 2: RUM/CBC predicts holdouts better than does RRM (significantly so in the case of in-sample holdouts, not significantly in the case of out-of-sample holdouts).

Conclusions and Recommendations for Practitioners

RRM offers a new choice paradigm, one conveniently estimable with standard logit or hierarchical Bayesian MNL software (e.g. CBC/HB). RRM requires a bit of additional effort in coding the regret functions, extra work that pays off to some extent in Study 1 but not in Study 2. These conflicting findings reflect the mixed results reported in the transportation literature (Chorus 2012b).

The bulk of previous research on RRM has occurred in the transportation literature and the current study extends evaluation of RRM to the marketing research domain.

While our empirical comparisons put RRM at rough parity with RUM/CBC, RRM's interesting foundation in regret theory makes it something worth keeping in mind: should neuroscience come to favor the role of regret over that of utility maximization in human decision making, the day could come when RRM could be a more preferred choice paradigm.

Also, it might be worthwhile investigating the interpretation and impact of using continuous non-linear functions for quantitative variables in RRM.

Finally, researchers may want to identify the best ways of making efficient and informative designs for RRM.

References

Ariely, Dan (2009) *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York: Harper Collins.

Chorus, Caspar G. (2010) "A New Model of Random Regret Minimization." European Journal of Transport and Infrastructure Research, **10**:181–196.

Chorus, Caspar G. (2012a) *Random Regret-based Discrete Choice Modeling: A Tutorial*. Springer Briefs in Business (e-book).

Chorus, Caspar G. (2012b) "Random Regret Minimization: An Overview of Model Properties and Empirical Evidence," *Transport Reviews*, **32**, 75-92.

Coricelli, Giorgio, Hugo D Critchley, Mateus Joffily, John P O'Doherty, and Raymond J. Dolan (2005) "Regret and Its Avoidance: A Neuroimaging Study of Choice Behavior," *Nature Neuroscience*, **8**: 1255-1262.

Kahneman, Daniel (2011) *Thinking Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, Daniel and Amos Tversky (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, **47**(2), pp. 263-291.

Kahneman, Daniel, Paul Slovick and Amos Tversky (1982) *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University.

Thaler, Richard H. and Cass Sunstein (2009) *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York: Penguin.

Tom, Sabrina M., Craig R. Fox, Christopher Trepel, and Russell A. Poldrack (2007) "The Neural Basis of Loss Aversion in Decision-Making under Risk," *Science*, 315: 515-518 (downloaded from www.sciencemag.org on April 14, 2011).

Webb, Ryan, Paul W. Glimcher, Ifat Levy, Stephanie C. Lazarro and Robb B. Rutledge, "Neural Random Utility and Measured Value," (August 1, 2013). Available at SSRN: http://ssrn.com/abstract=2143215 or http://dx.doi.org/10.2139/ssrn.2143215.