Chapter 2
A Random Regret Minimization-based Discrete Choice Model

Abstract This Chapter presents the RRM-model. First, the Random Regret-function is presented and explained (Sect. 2.1). Subsequently, this function is compared with the classical linear-additive Random Utility-function (Sect. 2.2). Finally, it is shown how the Random Regret-function translates into MNL-type choice probabilities for a particular distribution of the random error terms (Sect. 2.3).

2.1 A Random Regret-Function

The key behavioral notion on which the RRM-model is built is that people, when choosing, compare a considered alternative with each of the other available alternatives in terms of each characteristic (or from here on: attribute), and that they wish to avoid the situation where a chosen alternative is outperformed by one or more other alternatives on one or more attributes (which would cause regret).[1] Importantly, in contrast with other models and theories that are based on regret-minimization, the RRM-model postulates that anticipated regret is also a determinant of choices when there is no uncertainty about the performance of alternatives. The RRM-model postulates that as long as alternatives are characterized in terms of multiple attributes,

[1] Obviously, the level of anticipated regret that is associated with a particular alternative will vary between individuals. More specifically, different individuals may have different tastes and perceptions regarding alternatives and their attributes. Mathematically, this heterogeneity across individuals can be expressed by making relevant terms in the regret equation presented below (such as β and ε) individual-specific by means of an index (usually 'n'). In this tutorial, for reasons of readability, no such indices are used. As a result, equations refer to (the tastes and perceptions of) an average or 'representative' individual.

C. G. Chorus, *Random Regret-based Discrete Choice Modeling*,
SpringerBriefs in Business, DOI: 10.1007/978-3-642-29151-7_2,

which implies that trade-offs have to be made by the decision-maker, there will be regret in the sense that there will generally be at least one non-chosen alternative that outperforms a chosen one in terms of one or more attributes.

More specifically, the RRM-model is designed to incorporate the following seven behavioral intuitions relating to the anticipated regret associated with a considered alternative:

1. when a considered alternative *outperforms* another alternative in terms of a particular attribute, the comparison of the considered alternative with the other alternative on that attribute does not generate anticipated regret.
2. When a considered alternative *is outperformed* by another alternative in terms of a particular attribute, the comparison of the considered alternative with the other alternative on that attribute generates anticipated regret.
3. Anticipated regret increases with the importance of the attribute on which a considered alternative is outperformed by another alternative.
4. Anticipated regret increases with the magnitude of the extent to which a considered alternative is outperformed by another alternative on a particular attribute.
5. Anticipated regret increases with the number of attributes on which the considered alternative is outperformed by another alternative.
6. Anticipated regret increases with the number of alternatives that outperform a considered one on a particular attribute.
7. Anticipated regret is, from the perspective of the analyst, partially 'observable' (in the sense that it can be explicitly linked to observed variables) and partially 'unobservable'.

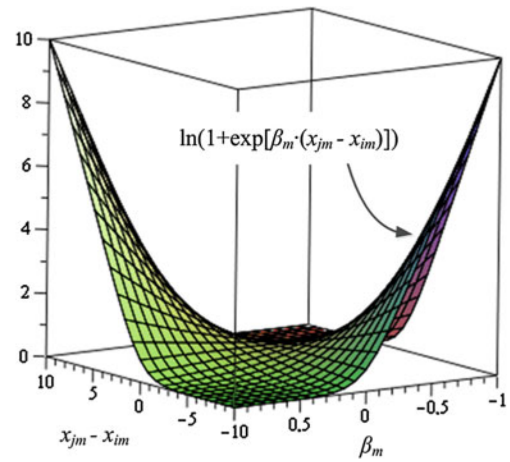
The following equation (Eq. 2.1) gives a formulation of regret that is consistent with these intuitions:

$$RR_i = R_i + \varepsilon_i = \sum_{j \neq i} \sum_m \ln\left(1 + \exp\left[\beta_m \cdot \left(x_{jm} - x_{im}\right)\right]\right) + \varepsilon_i \tag{2.1}$$

RR_i denotes the random (or: total) regret associated with a considered alternative i
R_i denotes the 'observed' regret associated with i
ε_i denotes the 'unobserved' regret associated with i
β_m denotes the estimable parameter associated with attribute x_m
x_{im}, x_{jm} denote the values associated with attribute x_m for, respectively, the considered alternative i and another alternative j.

Before discussing this function in more depth, it should be noted that of course, constants can be added to regret-functions, to represent the mean unobserved regrets associated with particular alternatives. Also, note that attributes may take the form of continuous variables as well as variables of categorical measurement level. Furthermore note that socio-demographic variables (such as age, gender, income and education level) may enter the regret-function to express segmentations of the population in terms of preferences for alternatives and tastes for attributes. Finally,

Fig. 2.1 Attribute-level regret as a function of β_m and $(x_{jm} - x_{im})$



note that in this chapter the focus is on attributes that are common, or shared, or generic, across alternatives. See Chap. 4 for an in-depth discussion of how the RRM-model deals with constants, interactions with socio-demographic variables, non-continuous variables, and with attributes that are specific to particular alternatives.

Returning to the regret-equation presented above: the term $\ln(1 + \exp[\beta_m \cdot (x_{jm} - x_{im})])$ is the core of this equation: it forms a measure of the amount of regret that is associated with comparing a considered alternative i with another alternative j in terms of a particular attribute x_m. This attribute-level regret is computed for each of the bilateral comparisons with other alternatives, and for all available attributes; the summation of these attribute-level regret terms (totaling $M * (J{-}1)$ terms in a situation where there are M attributes and the choice set contains J alternatives) forms the observed regret that is associated with the considered alternative. In light of the important role of attribute-level regret, also when it comes to deriving and interpreting the properties of the RRM-model, it is worth paying additional attention to this regret-kernel.

First, let us investigate whether the third and fourth behavioral intuitions formulated above (which refer to attribute-level regret) are captured in the used measure. As $\ln(1 + \exp[\beta_m \cdot (x_{jm} - x_{im})])$ is a monotonically increasing function of both β_m and $(x_{jm} - x_{im})$ it is easily seen that attribute-level regret increases with the importance of the attribute on which a considered alternative is outperformed by another alternative, and with the magnitude of the extent to which the alternative is outperformed by another alternative on that attribute. Figure 2.1 provides a visual account of this argument, and illustrates how attribute-level regret emerges for the situation where higher attribute-values are preferred over lower ones (implied by a positive sign of β_m), respectively for the situation where lower attribute-values are preferred over higher ones (implied by a negative sign of β_m).

More specifically, when higher attribute-values are preferred over lower ones, regret emerges to the extent that x_{jm} becomes larger than x_{im} and to the extent that

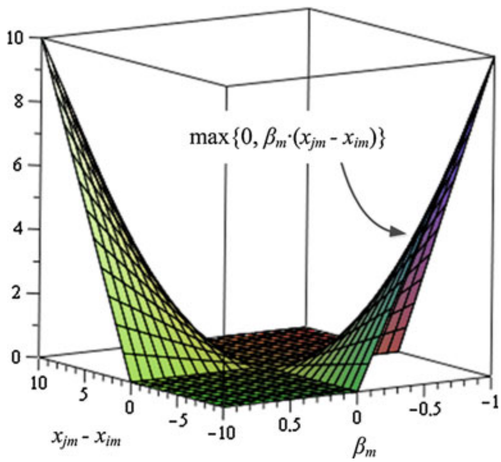


Fig. 2.2 Max {0, [$\beta_m \cdot (x_{jm} - x_{im})$]} as a function of β_m and $(x_{jm} - x_{im})$

β_m becomes more positive. In contrast, when lower attribute-values are preferred over higher ones, regret emerges to the extent that x_{jm} becomes smaller than x_{im} and to the extent that β_m becomes more negative. When the product $\beta_m \cdot (x_{jm} - x_{im})$ becomes negative (i.e., when the considered alternative outperforms the other alternative with which it is compared in terms of the attribute), attribute-regret starts to approach zero.[2]

Note that strictly speaking, the fact that attribute-regret *approaches* rather than *equals* zero when the considered alternative outperforms the other alternative implies that the first of the seven behavioral intuitions formulated at the beginning of this Chapter is violated. More generally speaking, although the measure $\ln(1 + \exp[\beta_m \cdot (x_{jm} - x_{im})])$ turns out to form a behaviorally intuitive measure of attribute-level regret, it seems a bit far-fetched at first sight. Indeed, a more intuitive and easy to interpret measure of attribute-level regret exists, in the form of max {0, [$\beta_m \cdot (x_{jm} - x_{im})$]}. This measure, introduced in Chorus et al. (2008),[3] actually represents attribute-level regret in a more direct way than does $\ln(1 + \exp[\beta_m \cdot (x_{jm} - x_{im})])$: when a considered alternative outperforms another alternative in terms of the attribute, regret equals zero (hence, the first behavioral intuition is not violated by this measure of attribute-regret). When the considered alternative is outperformed, regret equals the product of the importance of the attribute and the difference in attribute-values. Figure 2.2 plots max{0, [$\beta_m \cdot (x_{jm} - x_{im})$]} as a function of β_m and $(x_{jm} - x_{im})$, thus forming a counterpart of Fig. 2.1. The resemblance with Fig. 2.1 is very obvious: it is easily seen that $\ln(1 + \exp[\beta_m \cdot (x_{jm} - x_{im})])$

[2] It should be noted at this point that during the estimation process the sign of parameters is estimated together with their magnitude. That is, no a priori expectations need to be formulated by the analyst in terms of whether higher attribute-values are preferred by the decision-maker over lower ones, or vice versa.

[3] See Chorus et al. (2008, 2009) and Hess et al. (in press) for applications of this particular form of the RRM-model.

forms a close approximation of $\max\{0, [\beta_m \cdot (x_{jm} - x_{im})]\}$, especially when the absolute values of β_m and/or $(x_{jm} - x_{im})$ become larger.

The main difference between the two measures is that the former function is smooth while the latter (the one with the max-operator) is not. It turns out that it is exactly this 'non-smoothness' of $\max\{0, [\beta_m \cdot (x_{jm} - x_{im})]\}$ what makes this measure a less useful one for discrete choice modeling. More specifically, the fact that the function is discontinuous around zero (see Fig. 2.2) makes that the partial derivatives of the function with respect to βs and xs cannot be computed in that area. This causes non-trivial theoretical and practical difficulties in the process of Maximum Likelihood-based model estimation as well as in the process of computing elasticities. This non-smoothness also results in practical difficulties in the sense that conventional software-packages do not support this kind of non-smooth functions and hence the researcher has to rely on handwritten code for model estimation. This obviously greatly hampers the model's usability, especially among students and practitioners. These theoretical and practical issues are solved by adopting the smooth function plotted in Fig. 2.1, which was introduced in Chorus (2010) and which is used in the remainder of this tutorial.[4]

2.2 A Comparison with (Linear-Additive) Utility Maximization

Before deriving RRM-choice probability-formulations for the random regret-function presented in Sect. 2.1, it is instructive to compare the random regret-function with the random utility function that has dominated the field of DCM for decades. More specifically, the random regret-function is contrasted with its most natural counterpart: the so-called *linear-additive* random utility-function. The term 'linear-additive' refers to the fact that observed utility is a summation of terms that each consist of a product of a parameter and an attribute-value. It is this function that utility-maximizers aim to maximize by choosing between different alternatives. In notation (Eq. 2.2):

$$U_i = V_i + \varepsilon_i = \sum_m \beta_m \cdot x_{im} + \varepsilon_i \tag{2.2}$$

U_i denotes the random (or: total) utility associated with a considered alternative i
V_i denotes the 'observed' utility associated with i
ε_i denotes the 'unobserved' utility associated with i
β_m denotes the estimable parameter associated with attribute x_m
x_{im} denotes the value associated with attribute x_m for the considered alternative i.

[4] Note that, although the use of smooth attribute-regret-function instead of the non-smooth one is inspired mostly by pragmatic reasons as argued above, there is a deeper connection between the two functions as well. That is, when ignoring a constant, $\ln(1 + \exp[\beta_m \cdot (x_{jm} - x_{im})])$ gives the expectation of $\max\{0, [\beta_m \cdot (x_{jm} - x_{im})]\}$ when the two terms between curly brackets are considered i.i.d. random variables with Extreme Value Type I-distribution (having a variance of $\pi^2/6$). A reason for this stochasticity might be that the researcher is only able to assess these two terms up to a random error.

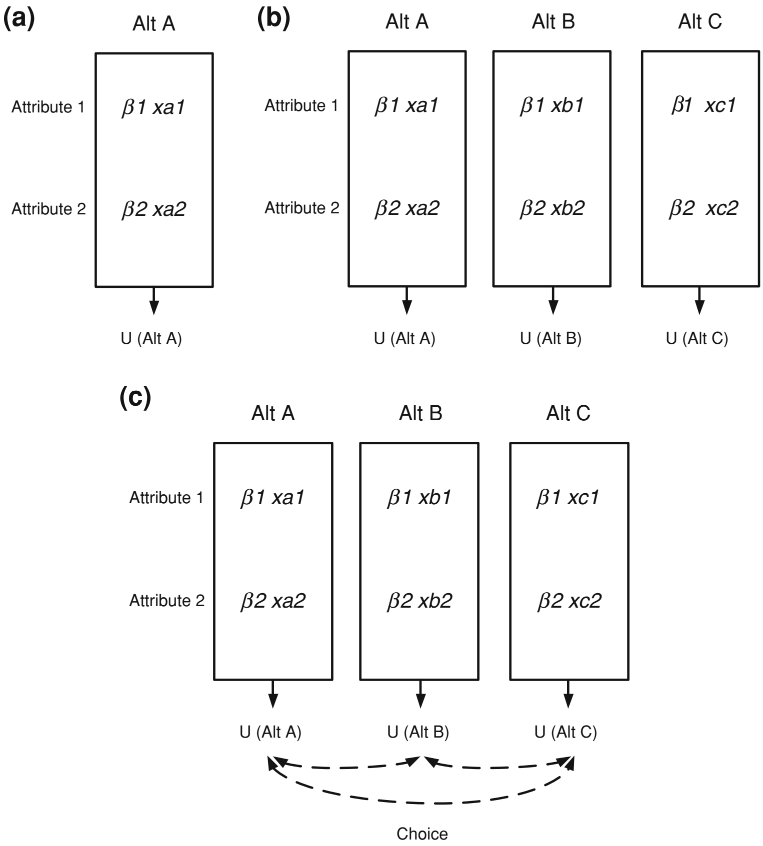


Fig. 2.3 A linear-additive utility maximization-based decision process (*solid arrows* represent summations, *dashed arrows* represent comparisons)

The conceptual differences between the utility-function presented directly above and the regret-function presented in Sect. 2.1 can be understood by inspecting Fig. 2.3: the Figure depicts the decision-process assumed in linear-additive RUM-models,[5] in the context of the following example: a decision-maker chooses between three alternatives A, B, and C (say, a train, car and bus-mode),

[5] Note that strictly speaking, DCMs do not really assume particular *processes* (in the sense that they do not postulate a particular order of decision-making steps). Rather, the mathematical formulation of the linear-additive RUM- model is in fact consistent with a range of underlying decision processes. Nonetheless, throughout the literature the linear-additive RUM-model form is generally considered to be the mathematical representation of the decision process described and visualized on this and the next page. It is instructive at this point to assume this particular order in decision-making steps as it highlights the ways in which RRM- and RUM-based decision rules differ in a conceptual sense.

and alternatives are evaluated in terms of two attributes (x1 and x2, say, travel time and travel cost). Linear-additive RUM-models assume that before a choice is made, the utility of each alternative is computed. Figure 2.3a depicts this process for alternative A (the train mode).

The decision-maker is assumed to 'compute' the utility of the train alternative by means of combining ('multiplying') his or her tastes (or: decision-weights) with the two mode-specific attribute-values. In other words: his or her taste for travel time is combined with the train's travel time and the same is done for the travel cost-attribute. Subsequently, the two combinations (one for time and one for cost) are summed together to form a measure of train-utility.[6] This process is repeated for the car and bus mode (Fig. 2.3b). After having this way 'computed' utilities for each of the three mode-options, the three utilities are compared and the highest utility alternative is chosen (Fig. 2.3c).

The RRM-model is based on quite a different assumed decision-making process (see Fig. 2.4): first, all alternatives are bilaterally compared on attribute x1 (travel time). The decision-maker does this by means of combining his or her taste for (decision weight attached to) travel time with travel time differences between train and car, train and bus, and car and bus, respectively. This process is done both ways (so: the car-option is compared with the bus-option in terms of travel time, and the bus-option is also compared with the car-option in terms of travel time, etc.); this results in $3*2 = 6$ attribute-regret terms. This process is repeated for the cost attribute (Fig. 2.4b), after which attribute level-regrets are summed together for each alternative (resulting in $3*2*2 = 12$ attribute-regrets in total, or $2*2 = 4$ attribute-regrets per alternative). This way, measures of overall regret are obtained for the train-, car- and bus-option; the regrets are compared and the mode with minimum regret is chosen.

In sum, in terms of the assumed underlying decision-process the (linear-additive) utility maximization-rule and the regret minimization-rule differ in the sense that, while the utility maximization-rule assumes that comparisons between alternatives are only made at the level of aggregated utilities, the regret minimization-rule postulates that comparisons between alternatives are made at the level of each attribute, as well as at the level of aggregated regrets. Phrased differently: whereas a utility maximizer is focused on the performance of a considered alternative itself (or: in isolation), a regret minimizer is focused on how a considered alternative compares with other alternatives in terms of every conceivable aspect.

However, it should at this point again be noted that to a considerable extent the differences highlighted above are artificial: strictly speaking, choice models do not assumed particular decision processes, and the same mathematical model-formulation is generally consistent with a range of underlying processes. The only

[6] Note that the random error is ignored in this example as it refers to the *analyst's* lack of knowledge and as such is irrelevant from an individual decision-maker's point of view.

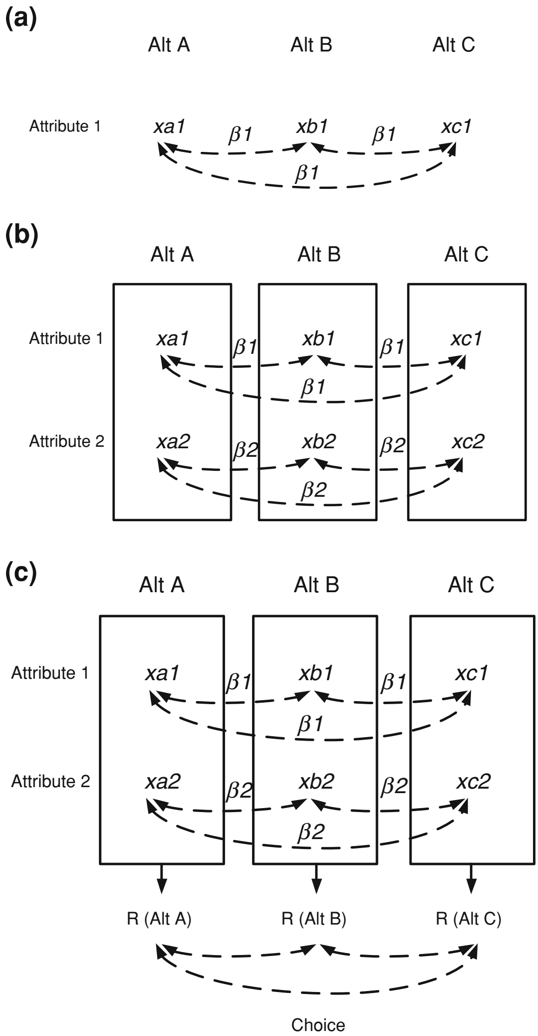


Fig. 2.4 A regret minimization-based decision process (*solid arrows* represent summations, *dashed arrows* represent comparisons)

formal difference between the RRM-model and a linear-additive RUM-model is that in the former, attributes of other alternatives codetermine the utility (called regret) of a considered alternative, and that they do so in an asymmetric, non-linear way.

As a consequence of the slightly fuzzy nature of the conceptual differences between the two models in terms of their assumed decision processes, it is more important to discuss how the two models differ in terms of their predictions (i.e., in terms of the choice probabilities they assign to different alternatives). The next chapter will provide such a comparison. However, before this comparison can be made, choice probabilities have to be derived for the RRM-model.

2.3 Regret-based Choice Probabilities and a RRM-based MNL-Model

Having established and discussed in-depth the random regret-function (as well as how it contrasts conceptually with its natural RUM-counterpart, the linear-additive random utility-function), the next step is to present a regret-based choice probability-formulation which gives the probability that a regret-minimizer chooses a particular alternative from a choice set. Obviously, should a researcher know the total (or: random) regret associated with each alternative, then he or she can readily determine the chosen alternative (which is the one with minimum regret). However, since part of the regret that is associated with a particular alternative is 'unobserved' by the analyst as is represented by the random error term, he or she can only predict choices up to a probability. Like is the case for RUM-models, this formulation of this probability depends on the particular distribution assumed for the error terms. For RUM-models it has been found (McFadden 1974) that the most convenient (because: closed form) formulation of choice probabilities is obtained when errors are assumed to be i.i.d. Extreme Value Type I-distributed.[7] This result (leading to the RUM-based MNL-or Multinomial Logit-model of discrete choice) can be used to obtain the same kind of closed form-probability formulation for regret-minimizers. This is most easily understood by first briefly revisiting the derivation of RUM-choice probabilities. More specifically, the probability that a utility-maximizer chooses alternative i from the set of J available alternatives, given i.i.d. Extreme Value Type I-errors, can be put as follows (Eq. 2.3):

$$\begin{aligned} P(i) &= P\left(U_i > U_j, \forall j \neq i\right) \\ &= P\left(V_i + \varepsilon_i > V_j + \varepsilon_j, \forall j \neq i\right) \\ &= \frac{\exp(V_i)}{\sum_{j=1..J} \exp(V_j)} \end{aligned} \tag{2.3}$$

It can now be easily seen (Eq. 2.4) that in a regret-minimization setting, the assumption that *the negative of* the random errors is i.i.d. Extreme Value Type I-distributed[8] results in a very similar, MNL-formulation of choice probabilities:

---

[7] The term i.i.d. stands for identically and independently distributed. This means that errors assigned to different alternatives are uncorrelated, and are drawn from the same distribution (with the same variance). This variance is usually fixed to $\pi^2/6$, which indirectly implies a normalization of systematic utility. In this tutorial, the scale of the utility or regret is always normalized this way, and is therefore not explicitly mentioned in equations.

[8] See the Appendix for a discussion of the validity of the assumption of i.i.d. errors in the context of RRM-models.

$$
\begin{aligned}
P(i) &= P\left(RR_i < RR_j, \forall j \neq i\right) \\
&= P\left(-RR_i > -RR_j, \forall j \neq i\right) \\
&= P\left(-(R_i + \varepsilon_i) > -(R_j + \varepsilon_j), \forall j \neq i\right) \\
&= P\left(-R_i - \varepsilon_i > -R_j - \varepsilon_j, \forall j \neq i\right) \\
&= \frac{\exp(-R_i)}{\sum_{j=1..J} \exp(-R_j)}
\end{aligned} \tag{2.4}
$$

The fact that the RRM-model features MNL-type choice probabilities (in combination with the fact that it has a smooth regret-function) comes with many benefits. Particularly it implies that, although the underlying behavioral premises of the RRM-model are fundamentally different from those of conventional RUM-models (see Sect. 2.2), the RRM-based MNL-model can use many of the econometric tools contained in the very comprehensive and well-understood 'toolbox' that has been developed over the past three decades in the context of the RUM-based MNL-model. This includes the use of estimation routines embedded in standard software packages.

It should be noted that, although in the remainder of this tutorial the focus will be on the MNL-model form presented above, extension towards so-called Mixed Logit-model forms is straightforward. That is, by adding error terms or by allowing parameters or the scale factor to vary randomly across individuals, so-called nesting and panel effects as well as random taste- and/or scale-heterogeneity can be accommodated in RRM-based Mixed Logit models. Translation of RRM-based MNL-models towards RRM-based Mixed Logit models is equivalent to the translation of RUM-based MNL-models towards RUM-based Mixed Logit models and hence will not be covered in this tutorial (see for example Train (2003) for an excellent treatment of RUM-based Mixed Logit models).

Finally, it should be noted that in binary choice situations (containing only two alternatives), the RRM-based MNL and the RUM-based MNL result in the same choice probabilities. For the interested reader, a formal proof is provided in the appendix of Chorus (2010).

References

Chorus CG (2010) A new model of random regret minimization. Eur J Transp Infrastruct 10:181–196

Chorus CG, Arentze TA, Timmermans HJP (2008) A random regret minimization model of travel choice. Transp Res B 42:1–18

Chorus CG, Arentze TA, Timmermans HJP (2009) Spatial choice: a matter of utility or regret? Environ Plan B 36:538–551

Hess S, Stathopoulos A, Daly A (in press) Mixing of behavioral processes: a modeling framework and three case-studies. Transportation. doi:10.1007/s11116-011-9365-6

McFadden D (1974) Conditional logit analysis of qualitative choice-behavior. In: Zarembka P (ed) Frontiers in econometrics. Academic Press, New York

Train KE (2003) Discrete choice methods with simulations. Cambridge University Press, Cambridge

http://www.springer.com/978-3-642-29150-0

Random Regret-based Discrete Choice Modeling
A Tutorial
Chorus, C.G.
2012, VIII, 55 p. 11 illus., Softcover
ISBN: 978-3-642-29150-0