Context effects in real life decision making

Erasmus School of Economics

Master's Thesis Marketing

Student: Stefan Sprangers
Student ID: 312101
Supervisor: Prof. dr. A.C.D. Donkers
December 2012

Abstract

Context effects are psychological effects that can influence customers' decisions to buy a certain product. Several researchers have discussed the presence of compromise, attraction and similarity effects in earlier papers. Previous researchers indicated that the middle option in the choice set will be more likely to be chosen (compromise effect). Furthermore, previous researchers indicated that alternatives that dominate one or more of their competitors should have a higher choice probability (attraction effect) and that alternatives are less likely to be chosen if there is a very similar alternative present in the choice set (similarity effect).

Until now, the presence of context effects has only been shown in experimental or other manipulated choice environments. My research attempted to investigate whether these context effects were present in real-life, non-manipulated decision making. A dataset of an online mortgage recommendations website was used to investigate whether these effects were presents in real-life. To do this, I created a multinomial choice model that included a context-free utility component and a context-dependent utility component. I found that compromise, attraction and similarity effects were present. Surprisingly, the direction of these effects was not always the same as previous researchers suggested. Possible explanations for these unexpected findings are discussed in this thesis.

Acknowledgements

This thesis marks the end of my life as a student in Rotterdam. With a lot of pleasure, I have studied at Erasmus University Rotterdam for more than five years. For my thesis and my education in general, I would like to thank a number of people.

First, I want to thank my parents and sister, who have always been very supportive. Back in the days when I was still living at my parents' houses they have often left the living room to allow me to work there. Second, I want to thank my supervisor Bas Donkers for providing me with a lot of help and directions. I had many questions regarding the data preparation and the statistical models. He always replied fast to my questions and provided me with useful answers. This was very helpful.

Table of Contents

1 Introduction

1.1 Introduction to context effects

Understanding how new products in a market can influence market shares of existing products is a relevant question for a lot of companies (Huber, Payne & Puto, 1982). To model choice behavior, different models can be used. Examples of choice models are multinomial logit (MNL) models and multinomial probit (MNP) models (McFadden, 1986). Most choice models assume that the addition of a choice alternative cannot increase the probability of choosing a different alternative. This assumption is called the regularity condition. In addition, most models assume that that the addition of a choice alternative decreases the market shares of all other products by the same percentage. We call this the proportionality principle (Huber, Payne & Puto, 1982).

However, situations are possible where these principles do not hold. In some cases, the attractiveness of a product does not only depend on the attributes of the product. It also depends on the relation of the product's attributes with its competitors' attributes. Context effects are psychological effects that can influence customers' decisions to buy a certain product. The presence of context effects entails that "consumer choices are partly driven by the context provided by the set of alternatives" (Rooderkerk et al, 2010).

Understanding these context effects can be very helpful to marketeers. Marketeers want to know how the introduction of a new product can influence the market shares of their existing products in the market. Cannibalization can be a big issue for companies; market shares of existing products can go down if a new product is added to the portfolio. However, if context effects truly exist in real life, then the market share of an existing product could actually be boosted when a new product enters the market in certain scenarios (Mishra et al, 1993). Aside from this strategic aspect, marketeers could adjust the advertising of their product to emphasize the position of their product in relation to the competitors. A thorough understanding of context effects could provide marketeers with directions on which characteristics of a product should be emphasized in an advertisement.

In the remainder of this section I will briefly introduce three commonly discussed context effects: the similarity effect, the attraction effect and the compromise effect. Some researchers state that companies with products that are similar to the new alternative are more likely to suffer a significant loss in market share than competitors with products that are dissimilar to the new alternative (Huber, Payne & Puto, 1982). This context effect is called the similarity effect. The proportionality principle is violated if the similarity effect occurs.

The regularity, similarity and proportionality conditions can all be violated when an attraction effect or asymmetric dominance effect occurs. Huber, Payne & Puto (1982) were the researchers that introduced the theory of asymmetric dominance. They stated that "an asymmetrically dominated alternative is dominated by one item in the set, but not by another. Adding such an alternative to a choice set can increase the probability of choosing the item that dominates it". The "attraction effect" is a term in literature that is used interchangeably with the term "asymmetric dominance effect". From now on, I will talk about the "attraction effect" for the sake of consistency, even if a researcher only mentioned the term "asymmetric dominance effect".

I will use an example to explain the attraction effect in further detail. Two products are available in a market: product A and product B. Both products are rated on a scale from 1 to 100 for both attributes (quality and price). Product A scores highest on quality, while product B scores best on price. In this situation, none of the alternatives is dominated. You can see this situation in the graph below:

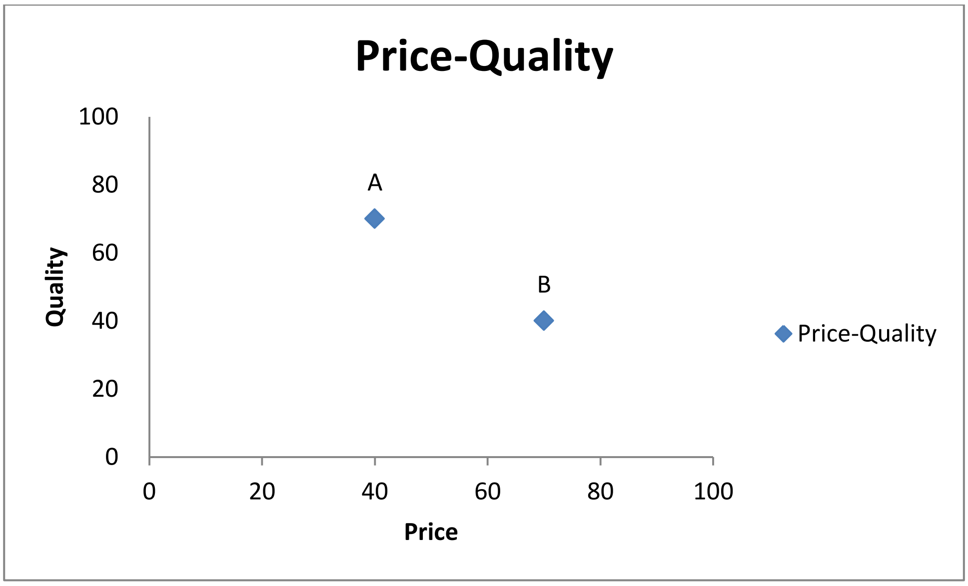


Figure 1: Choice set of two products

In this graph, there are no dominant alternatives. It is not clear whether product A or product B would be more appealing to the customer. However, an attraction effect can occur when a product is added to the choice set that is clearly inferior to product A or B. Such a situation can be displayed in the following graph:

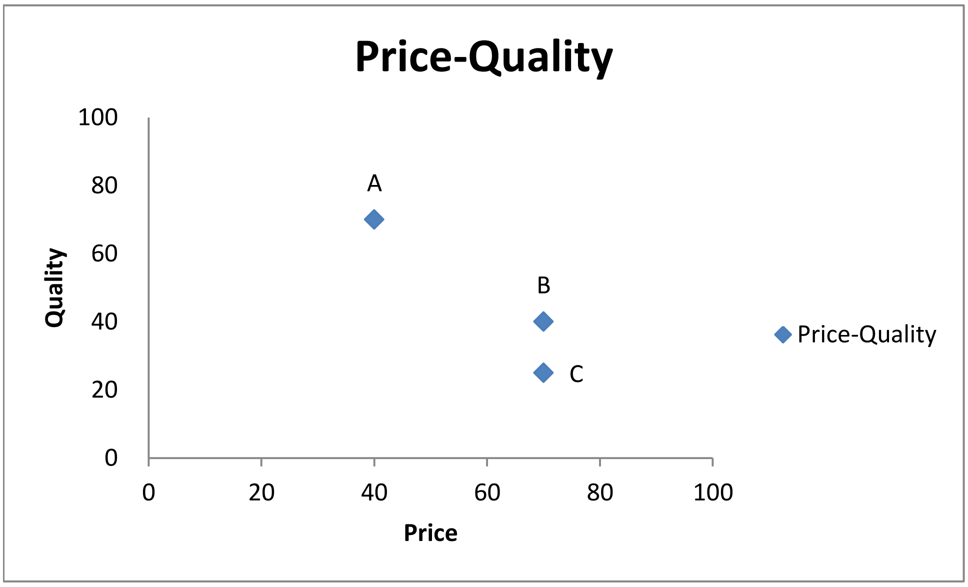


Figure 2: Illustration of the attraction effect

In this situation an inferior product has been added: product C. Product C has lower quality than product B, while both have exactly the same perception of price. Therefore, product B dominates product C. Product A does not dominate product C, because product C has a lower price (better price perception). In this situation, product B is called the target, product A is called the competitor and product C is defined as the decoy. When alternative C is added to the choice set, the market share of product B could go up, while the market share of product A is likely to go down. If the market share of product B would go up, then this is in conflict with the regularity, similarity and proportionality principles.

The similarity effect entails that alternatives that have a similar competitor have a lower market share (on average) than alternatives with no similar competitors (Huber, Payne & Puto, 1982). The introduction of a new product hurts products that are similar to the new product the most, if the similarity hypothesis holds. Suppose that the initial situation in the market is given in figure 1 and a new product enters the market that is similar to product B (product C). Such a situation is drawn in the figure below:

Figure 3: Illustration of the similarity effect.

When product C enters the market, the similarity effect could occur. This means that product B loses relatively more market share than product A. Customers could see product C as a substitute for product B, because the products are quite similar. This is less likely to occur for product A, because that product is less similar to product C.

The compromise effect entails that customers have a tendency to choose the “middle option” of a choice set (Huber & Puto, 1983). Again, figure 1 will be taken as a starting point for this explanation. Imagine that a product is added that has the highest rating for price and the lowest rating for quality (option C):

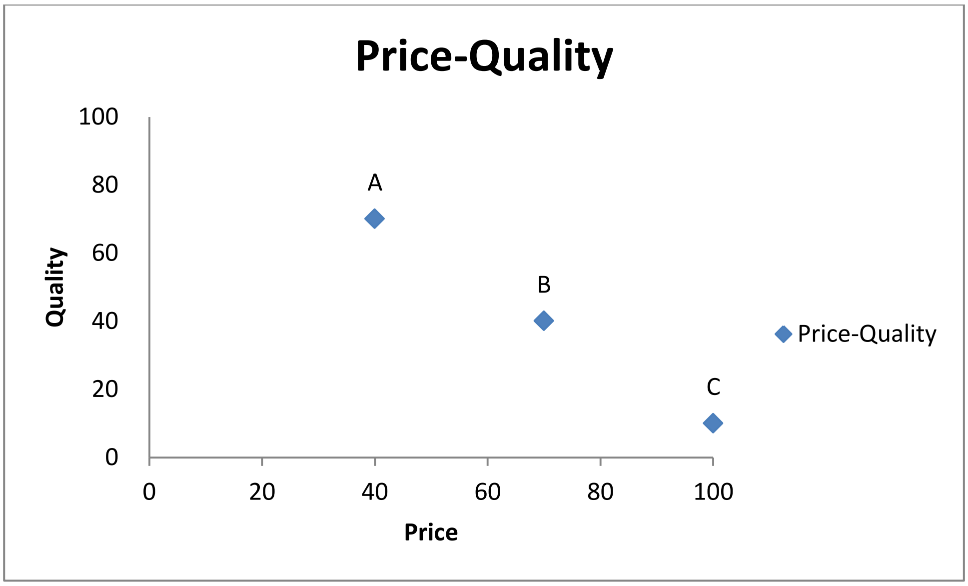


Figure 4: Illustration of the compromise effect

In this case, option B becomes the middle option in the choice set. Despite the fact that a new option is added to the choice set, the market share of option B could go up if the compromise effect occurs. Another, milder version of the compromise effect occurs when the market share of product A decreases relatively more than the market share of product B. Possibly, the cause for the compromise effect is that customers perceive the middle option to be a safer option (Huber & Puto, 1983). Possible explanations for the different context effects and factors that may enhance these effects are discussed in the literature review (chapter 2).

1.2 Research objective and relevance

Adding a new product to the portfolio can be very risky for a company. For this reason, knowledge about the consequences of adding new products to the portfolio is very valuable to companies. Understanding attraction, compromise and similarity effects will give companies more knowledge about the consequences of adding a new product to their portfolio. Companies may be able to boost sales of certain products by adding an alternative that is (nearly) dominated by the products that are currently part of the product portfolio, or by adjusting the degree to which products are considered to be middle products in the choice set. In addition, this research could generate knowledge on how to respond to new product introductions of other companies.

In addition to the relevance for companies, this paper is also relevant for researchers. According to Mishra et al (1993), "a better understanding of the attraction effect will enable consumer researchers to design consumer choice experiments more carefully by (1) controlling for influential factors in the choice task and (2) estimating the nature and magnitude of the effect of these factors on the choice task".

A lot of research has already been done about the occurrence of the attraction, compromise and similarity effects in experimental choice environments. This research attempts to show the occurrence (or non-occurrence) of context effects in a real, non-manipulated choice environment. Therefore, the main research question of this research will be

To what extent are attraction effects, compromise effects and similarity effects present in real life decision making?

2 Literature review

This literature review examines literature about context effects in more detail. Explanations will be given for the reason of the occurrence of context effects and possible moderators of context effects will be discussed. In the last section of this chapter, choice models will be examined that include attraction, compromise and similarity effects.

2.1 Attraction and similarity effects

The attraction effect has been discussed in multiple scientific papers. Attraction is a "gravitational metaphor that we use to describe the empirical finding that a new item can increase the favourable perceptions of similar items in the choice set" (Hubert & Puto, 1983). Hubert & Puto (1983) call attraction a positive similarity effect. This opposes the view of earlier researchers that talked about a negative similarity effect or substitution effect. According to those researchers, the market shares of similar products decrease by a larger percentage than the market shares of non-similar alternatives when a new alternative is added to the choice set. The reasoning behind that is that the choice of a certain item reflects the customers' preferences. If a new alternative shows up with similar characteristics, then this alternative fulfils partly the same needs as the similar alternative that was already present in the choice set. For this reason, some previous researchers assume that market shares of similar products should go down most (percentage-wise) when a new product enters the market, which is called the similarity effect.

Huber & Puto already demonstrated the presence of the attraction effect in 1982 (Huber, Payne & Puto, 1982) and did it once again in 1983 (Huber & Puto, 1983). However, there was an important difference in the research approach between these two papers. In 1983, they added decoys that were a lot less attractive than the other alternatives in the choice set, but the decoys were not entirely dominated by the targets. The attraction effect was still present in this experiment, despite the fact that there was no full domination by the target.

Other researchers wrote papers about the attraction effect as well. Simonson (1989) did multiple experiments and showed the presence of the attraction effect in each experiment. Shafir et al. (1993) and Dhar (1997) incorporated a no-choice option in their experiments. They found that including an asymmetrically dominant alternative decreased the amount of choice deferral. Doyle et al. (1999) demonstrated the occurrence of the attraction effect for real, in-store purchases. However, it was still an experiment; the choice set in the store was manipulated for the research.

2.1.1 Two conflicting forces

Huber & Puto (1983) talked about two conflicting forces in the case where decoys are less desirable than target and competitor, but still viable options. In that case, the substitution and the attraction effect can be counteracting forces. The attraction effect entails that people switch from competitor to target due to the dominating position of the decoy. This increases the market share of the target. The substitution effect (or negative similarity effect) entails that a new alternative (e.g. the decoy) takes more share from similar items (e.g. the target). This decreases the probability of choosing the target.

The experiments of Huber & Puto (1983) showed that both effects were present. In study 1a, 1b and study 2 the attraction effects were stronger than the substitution effects, because the market shares of the targets were higher than the expected market shares under the model that assumes proportionality (Luce model). In study 1c, the market share of the target was exactly the same as expected under the Luce model. In this study, the decoy was quite strong when compared to studies 1a and 1b.

The local substitution effect gets stronger when the decoy becomes more attractive (Huber & Puto, 1983). However, Huber & Puto concluded that the positioning of the decoy did not change the magnitude of the attraction effect. Apparently, the strength of the local substitution effect was the reason why the market share of the target dropped when the attractiveness of the decoy was increased. The order of the attributes of the decoy, target and competitor is probably the main driver of the attraction effect, instead of the differences in values of the attributes.

In some cases, the Luce (proportionality) model can actually be a model that is relatively accurate. However, this happens in cases where the attraction effect and the substitution effect cancel each other out (Huber, Payne & Puto, 1983).

2.1.2 Causes and factors that can increase the attraction effect

Bettman et al. (1998) argued that customers can have multiple goals in decision making: minimizing the cognitive effort that is involved in making a choice, maximizing decision accuracy, minimizing the experience of negative emotions in decision making and maximizing the ease of justification of the decision. According to Bettman et al. (1998), using mental shortcuts like the attraction effect can reduce the required effort in decision making (because using a shortcut is faster than a thorough evaluation of all attributes of all the alternatives), it can decrease the probability of experiencing negative emotions (because using the attraction effect can reduce the need of thinking about trade-offs), and the attraction effect can provide people with a justification of their choice. Bettman et al. (1998) also predict that the occurrence of the attraction effect is more likely when people are exposed to emotionally difficult trade-offs.

Huber, Payne and Puto (1982) mentioned multiple possible reasons for the occurrence of the attraction effect. When a dominated alternative is present, it makes the choice easier for the customer; cost of thinking is less when a choice is being made between dominated pairs than when a choice is being made between nondominated pairs. This is one possible reason for the occurrence of the attraction effect.

Huber, Payne and Puto (1982) investigated the size of the attraction effect in various scenarios. Each scenario included the target (dominating alternative once the decoy is added), the competitor and the decoy (dominated alternative). The different scenarios are displayed graphically in Figure 3 below, where the circles with F, R, RF and R* display the position of the decoy in the various scenarios.

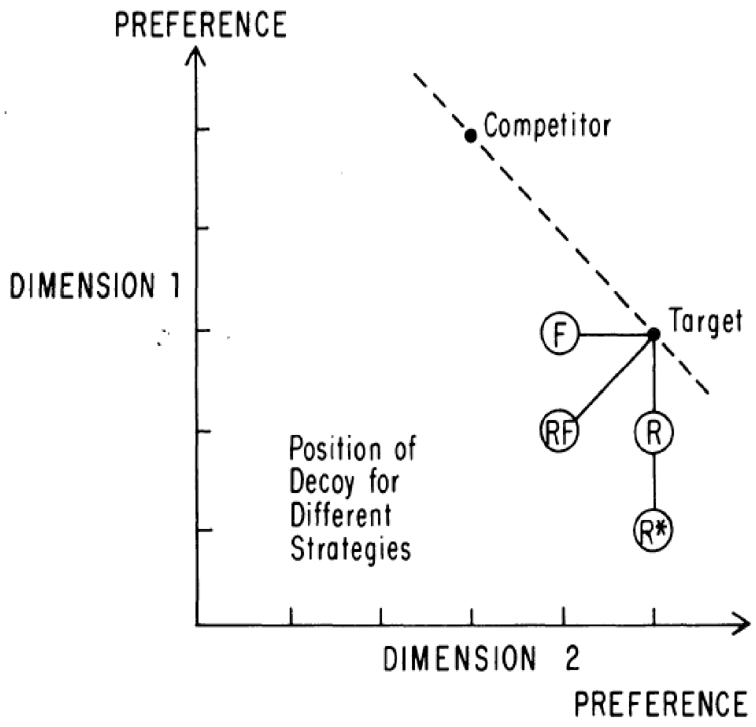


Figure 5: Different decoy placement strategies (Huber, Payne & Puto, 1982)

In the scenario where the decoy was worse than the target on the weakest attribute of the target (scenario R), the perceived difference between the competitor and the target on that attribute will be lower. By adding the decoy, the target becomes the middle option on that attribute, instead of being the worst option. The R* scenario was expected to strengthen this effect. Another scenario can be the scenario where the decoy is worse than the target at the target's strongest attribute, but still better than the competitor (scenario F). Instead of having the worst attribute value in a choice set of two items, the competitor has the worst attribute value in a choice set of three items. This can increase the perceived difference between the competitor and the target on dimension 2. The RF scenario creates superiority of the target on two dimensions when compared against the decoy. On the one hand, one could expect that this would increase the market share of the target when compared to the R and the F scenario. On the other hand, dominance may be easier to spot in the R and the F scenario than in the RF scenario. Huber, Payne & Puto (1982) checked if their expectations were correct by doing a survey. They found out that the R and R* scenario's increased market shares of the targets the most, followed by F, and RF turned out to be the least effective strategy. The fact that RF was the least effective strategy suggests that attraction effect is less strong when the target and decoy are very dissimilar (Huber, Payne & Puto, 1982).

Outside of the scenario-specific explanations, Huber, Payne & Puto (1982) discussed more possible drivers of the attraction effect. One reason is perceived popularity of the decoy. If respondents see the decoy as a popular product, then this may increase the attractiveness of the superior target.

Another possible evaluation method would be to compare all the alternatives with each other in a "round robin tournament", where the alternative with the most wins is chosen. The target would win against the decoy for sure, hence increasing the probability that the target wins the "tournament". Another explanation of Huber, Payne & Puto (1982) is related to the cost of thinking. The choice between a dominated pair (target vs. decoy) is a lot easier than the choice between a nondominated pair (competitor vs. target/decoy). Due to the difference in cost of thinking, the customer is more likely to make a choice between the target and the decoy, increasing the probability that the target is chosen. However, in the next study, Huber & Puto (1983) casted doubt on these process explanations. The experiments that they executed in 1983 contain inferior rather than dominated alternatives. Therefore, the likelihood of preferring the target above the decoy is lower than when the decoy is a truly dominated alternative. Huber, Payne & Puto (1982) also mentioned that the attraction effect might be weaker when there are more items present in the choice set, or when each alternative has more attributes/dimensions.

The popularity motive was also mentioned by Huber & Puto (1983). They indicated that the mere presence of an item in a choice set could let customers infer that some people must like that product, even if it's a decoy in an experiment. The presence of a superior alternative to that decoy then communicates quality to the customer.

The perceived uncertainty may be influencing the presence of an attraction effect. When there is not much information about the products in a choice set, people may have a higher tendency to choose the least risky product. Therefore, attraction effects are probably more often present in upcoming markets with not many established brands. When there is a lot of clarity involved in decision making, the substitution effect is probably more salient (Huber & Puto, 1983). When customers base their decisions on clear attributes of a product rather than brand image, then the substitution effect is expected to occur more often.

Similarly to Huber, Payne & Puto (1982), Pechtl (2009) examined the influence of the position of the decoy on the occurrence of the attraction effect. He found that, depending on the position of the decoy, the attraction, compromise or substitution effect can occur. In the figure below, alternative A represents the target, alternative B represents the competitor, and the rectangles indicate which effect was most likely to occur when they placed a decoy in the corresponding area. The decoy effect is a different term for the attraction or asymmetric dominance effect:

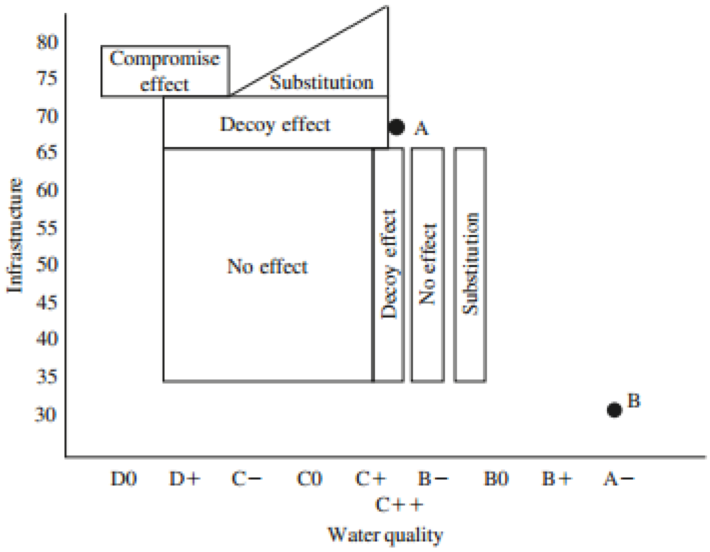


Figure 6: The influence of the position of the decoy on the presence of context effects (Pechtl, 2009)

Simonson (1989) adds another explanation in his paper. According to him, people often find it important to be able to justify their own choices to others. Possible explanations for this include "a desire to enhance one's self-esteem, anticipation of the possibility of regret, or cognitive dissonance, as well as people's perception of themselves as rational beings with reasons for preferring one option over others". If a dominant alternative is present, then the choice for the target is easier to be justified than the choice for the competitor. The attributes that important people in someone's life find important are not always clear. People are usually more certain about their personal preferences than about other's preferences. Therefore, people may be more likely to choose the dominant alternative, because people can give factual argumentation about choosing the target; argumentation that is valid, independent from other's preferences (Simonson, 1989). Experiments by Simonson (1989) showed that the size of the attraction effect interacts with the need for justification. In addition, he showed that the presence of asymmetrically dominant alternatives increases the ease of justification and decreases the likelihood of criticism.

Some researchers indicated that people base their choices on characteristics that other people find important. Even when the choice does not have to be explained to others, people still consider the opinions of others when making choices (Simonson, 1989). The last experiment of Simonson (1989) showed that the difficulty of the decision positively increases the likelihood of the occurrence of the attraction effect. Kivetz et al. (2004) argued that multiperson decision making is likely to increase the chance of being evaluated. People that are under the perception of being evaluated are more likely to be influenced by attraction effects and compromise effects.

Ratneshwar et al. (1987) expected the attraction effect to be insignificant when there is a meaningful stimulus description and a high level of product familiarity. Study 1 and study 2 of this paper suggested that meaningfulness of the stimuli decreases the probability of the occurrence of the attraction effect, while no significant effects were found for familiarity. A possible explanation for this could be that familiarity with a product category increases the meaningfulness of the stimuli. In addition, in study 3 of the paper of Ratneshwar et al. (1987), quality ratings of the target improved by a huge amount when a favourable decoy was present, as opposed to a situation where the decoy was in an unfavourable position.

Simonson & Tversky (1992) explain the asymmetric dominance effect by looking at tradeoff contrast. Consider the market with product x, product y and product z that is visible in the graph below:

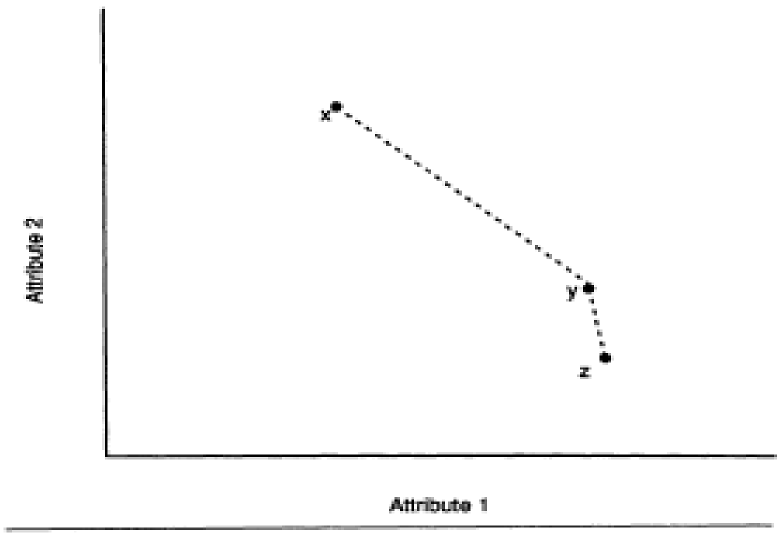


Figure 7: Asymmetric dominance effect with partial dominance (Simonson & Tversky, 1992)

In this picture, product y is relatively superior to product z. It is a lot better on attribute 2, while being only slightly worse on attribute 1. There is no dominance relationship between product x and product y. In the experiment of Simonson and Tversky (1992), which includes photo cameras, product y had a higher market share in a market with product x, y and z, than in a market with only product x and y. This difference can be explained by the fact that the price that customers pay for an extra unit of attribute 2 is very low in the case of a switch from product z to product y. The amount that customers pay per unit of attribute 2 (in terms of units of attribute 1) is a lot higher in the case where a switch is made from product y to product x. Tradeoff contrasts can be divided into two categories:

- Background contrast. In this case, the tradeoff contrast is present, due to the comparison between the current tradeoff(s) and past trade-offs
- Local contrast. In this case, the tradeoff contrast is present in the current choice set, in a choice set with three or more options.

Mishra et al. (1993) investigated the influence of several variables on the asymmetric dominance effect. People had to distribute points to different products for a number of choice sets. Three product categories were present in the research: beer, cars and TV-sets. The attraction effect was measured for every individual with the following formula:

$$AE_L = O(X) - E_L(X)$$

where AE_L is the attraction effect, $O(X)$ is the observed share of brand X and $E_L(X)$ is the expected share of brand X under the Luce model. The research of Mishra et al. (1993) identified seven variables that are possible causes of the attraction effect:

- Product class knowledge. More knowledge about the product class makes the decision less difficult. When people are not familiar with a product, people are more vulnerable to changes in the way that the alternatives are presented. As a result, they are more likely to be influenced by the attraction effect. Product class knowledge was significant for beer only.

- Information relevance. If respondents receive relevant information about the products that they have to choose from, then they will feel more certain about their choice. As a result, they are less likely to be influenced by the attraction effect. Information relevance was significant for all product classes.

- Task involvement. Mishra et al. (1993) hypothesized that task involvement would directly influence the attraction effect, because people that are not very involved with a choice are more likely to use heuristics such as attraction effects to make their decisions. This main effect was insignificant for all product classes. However, the interaction between task involvement and information relevance was significant for all product classes. Information about the products became more relevant if people were involved in the task. This effect probably occurred because the ability to understand the information was enhanced when people were involved in the choice task.

- Decoy-target similarity. This variable was significant and positive for the beer and car market. Higher similarity between the decoy and the target may make it easier to spot the dominance relationship between the two variables. The decoy-similarity variable was insignificant for the tv-sets market, but slightly negative. The negative coefficient may be due to the fact that the products in the tv-sets market were often too similar. When two products are too similar, the target may lose share to the decoy, because people may find it too hard to distinguish the target and the decoy.

- Preference strength. This variable indicates the degree to which the customer preferred the target in a two-item choice set with the target and the competitor. Low preference strength increased the probability of the occurrence of the attraction effect. Preference strength was the most important predictor of the attraction effect for all markets. If a customer had a strong preference for the target, then this preference was unlikely to change due to the addition of a decoy.

- Popularity of the decoy and decoy share are the last two antecedents of the attraction effect that were discussed in the research of Mishra et al. (1993). In the beer and the car market, decoy share had a significant, positive influence on the attraction effect, while decoy popularity only had impact on the attraction effect by an interaction with the decoy share. In the tv-sets market, there was only a positive, significant influence of decoy popularity on the attraction effect.

Heath & Chatterjee (1995) did a meta-analysis of past experiments about the attraction effect, as well as an experiment with traditional (MBA students) participants and an experiment with non-traditional participants (students from a rural, teaching-oriented university that focused on the steel and coal industries). In the meta-analysis, the attraction effect was only present for high-quality targets, when the target dominated the decoy entirely. In this analysis, several drivers of the asymmetric dominance effect were found:

- Dominated decoys increased shares of the target more than viable decoys (decoys that are slightly inferior but not considered irrelevant alternatives)
- Decoys increased shares of high-quality targets more than shares of low-quality targets.
- Decoys that extend the range on the least favourable attribute of the target are more likely to cause attraction effects (similar to Huber, Payne & Puto; 1983)

The experiment showed an interesting difference between the traditional and the non-traditional population. For the traditional population, the attraction effect was only present for high-quality targets. On the contrary, the attraction effect was only present for low-quality targets for the non-traditional population (Heath & Chatterjee, 1995).

There have been some other reasons for or drivers of the attraction effect that were mentioned by research. Below I will give a quick summation of these reasons/drivers:

- Heath & Chatterjee (1995) mentioned that higher priced goods may be more susceptible to the attraction effect, due to increased risk. However, they did not test this.
- Wa et al. (2009) found that the presence of a unique categorical variable attenuated the attraction effect, while not affecting the compromise effect.
- Kim & Hasher (2005) found that older adults were less likely to experience the attraction effect than younger adults.

2.2 Compromise effect

Huber & Puto (1983) talked about a compromise effect that influences choice behavior by consumers. The compromise effect entails that consumers are inclined to choose the middle item in a choice set. According to Simonson (1989), the compromise effect proposes that “an alternative would tend to gain market share following the addition of an adjacent competitor that makes the brand a compromise choice within the set”.

Simonson (1989) hypothesised the compromise effect to be stronger when people assume they have to explain their choice to others. Choosing the middle option has the smallest maximum error; it is the safest choice. Furthermore, people can use the argument that the middle alternative has reasonable attribute values for all attributes. However, the experiment of Simonson (1989) did not see a significant interaction between the compromise effect and the justification category (high or low justification to others).

Another experiment was done by Simonson (1989), involving a different group of students. He showed that a compromise brand decreases the perceived likelihood of criticism, but no significant influence of the compromise effect on perceived ease of justification was found. Finally, similar to the findings for the attraction effect, Simonson (1989) found that decision processes leading to the choice of a compromise alternative are, on average, related to more difficult decisions than decisions processes leading to the choice of a non-compromise alternative. The reason for that is that logical reasoning based on trade-offs between attributes did not provide individuals with a solution in these difficult decisions. Choosing a compromise (or asymmetrically dominant) alternative can then be used as a second choice strategy.

Kivetz (2004) mentioned that the compromise effect is more likely to be present in group decision making. There are two reasons for this. First, people feel like they are being evaluated and are therefore more likely to choose the middle option. Second, the opinions of multiple people may be weighed into the decision, which can cause a choice to be made that partly fulfils the (conflicting) wishes of different people.

Kivetz et al. (2004) also talked about enhancement and detraction. Consider the following figure:

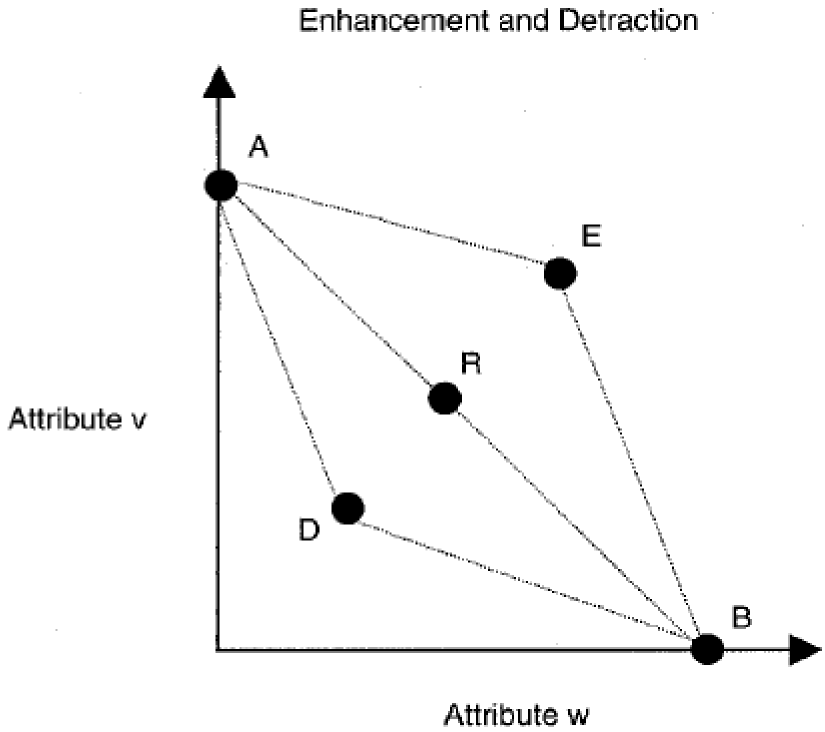


Figure 8: Enhancement and detraction (Kivetz et al, 2004)

In this figure, A, B, D and E are actual alternatives, while R is the average of the attributes of alternatives A and B. Alternative E was found to have a higher market share in a product set with {A, E, B} than in a product set with only {A,E} or {E,B}, because alternative E dominates R; E lies above the line that connects A and B (Kivetz et al, 2004; Simonson & Tversky, 1992). This effect is called enhancement and it is consistent with the compromise effect. The opposite of enhancement is called detraction; this effect is inconsistent with the compromise effect. Detraction entails that alternative D is likely to have a lower market share in a product set with {A, D, B} than in a product set with {A,D} or {D,B}, because R dominates alternative D; alternative D lies below the line that connects alternative A and B (Kivetz et al, 2004). These effects can be explained by tradeoff contrast (Simonson & Tversky, 1992). In the enhancement case, the cost of an extra unit v (in terms of units w) is much larger when a switch is made from E to A, compared to a switch between B to E. Enhancement and detraction can both be explained by differences in trade-offs.

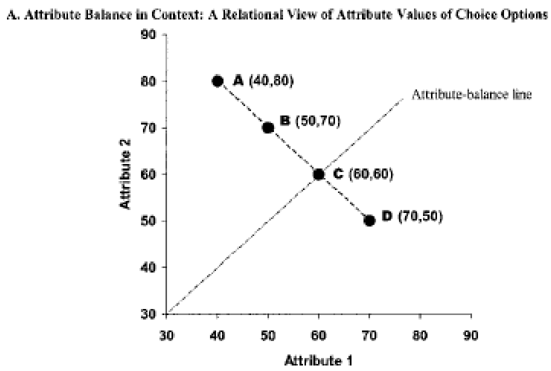
Simonson & Tversky (1992) explain the compromise effect by looking at extremeness aversion. Consider a market with product A, R and B (see figure above). In this market, there are no dominance relationships; all alternatives are on the same line. However, consumers are likely to prefer product R in this market. The reason for this is loss aversion. According to Simonson & Tversky, “outcomes that are below the reference point (losses) are weighted more heavily than outcomes that are above the reference point (gains)”. People are likely to choose the middle alternative as the reference point. If people would compare product R to product A or B, then the losses that occur when a switch would be made to one of these products may be considered worse than the gain on the other attribute when the switch is made, due to loss aversion. Simonson & Tversky (1992) also talk about polarization. In the case of polarization, extremeness aversion is only present for part of the attributes of the product. They argued that extremeness aversion is more likely to be present for quality-related attributes than for price attributes.

Sheng et al. (2005) identified several factors that can positively or negatively influence the compromise effect:

- People who were familiar with the product class were less likely to choose a compromise option.
- People who experienced a high amount of decision uncertainty were more likely to choose a compromise option.
- People were more likely to choose a compromise option if the decoy was further away from the target.
- People were more likely to choose a compromise option if the individual's attribute importances were asymmetric (unequal).
- Attribute asymmetry mediated the relationship between product familiarity and the likelihood of choosing a compromise option.

Lehmann & Pan (1994) looked at the difference between the compromise effects for brand entries in two-brand markets and eight-brand markets. They concluded that "a compromise option has a positive impact when the market is small and a negative one when the number of brands in the market increases to the extent that it is hard to stand out in the crowd". So the negative compromise effect for new entrants in eight brand markets can be explained by the fact that brands that are in the middle of a nine-item choice set will have a hard time distinguishing themselves from the other eight brands. In addition, Lehmann & Pan (1994) found that interaction effects between compromise and asymmetrically dominating options and interaction effects between compromise and asymmetrically dominated options can exist. He also found that the impact of the entry of compromise, extreme, dominating or dominated options on utility can differ greatly, depending on the position of the entry option in the choice set; the entry can be either in an desirable area (with more considered existing brands) or in a less desirable area (with less considered existing brands).

Chernev (2004) wrote an entire paper about the influence of attribute balance on consumer's choices. Consider the following example where we have three alternatives with each three attributes:



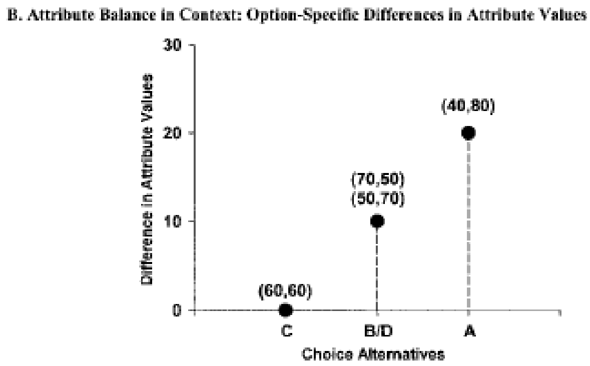


Figure 9: Compromise effects

Figure 10: Attribute balances

There are four alternatives: A, B, C and D. In a three-option choice set {A, B, C}, B can be seen as the compromise alternative. However, one could also argue that option C is the least extreme option, because it has decent attribute values for both attributes. In figure 8, you can see that option C is least extreme in terms of the difference between the two attributes. There could be two effects working together: a compromise effect and an attribute balance effect. Chernev (2004) tested this in an experiment. He found that the position and the balance of the attributes both affected the market share in the predicted way. Chernev (2004) did a similar experiment for the attraction effect in combination with attribute balance. He concluded that "the presence of a balanced alternative moderates the strength of the attraction effect".

Aside from the aforementioned reasons, researchers have mentioned other possible drivers of the compromise effect:

- Whether the middle option is also presented in the middle of the choice set (Chang & Liu, 2008)
- Whether the middle option is presented along with other options or separately (Chang & Liu, 2008)
- Familiarity of the middle option (Sinn et al, 2007)
- Heath & Chatterjee (1995) mentioned that loss-aversion may be greater for quality-related attributes than for price-related attributes. Loss aversion is a possible driver of the

compromise effect, so this suggests that higher quality items may be more susceptible to the compromise effect.

2.3 Modelling attraction and compromise effects

Most of the choice models that are used are utility-based models. In most of these models (like the multinomial logit model), dominated alternatives will still have a market share, because they only have to be inferior to the dominating alternative on one attribute. Both the dominating and the dominated alternative have utilities, giving them both a market share. This is inconsistent with the hypothesis of Huber, Payne and Puto (1982) that says that inferior alternatives should not have any market share. In this section I discuss several ways in which context effects can be incorporated into choice models to account for the relationships between alternatives.

Simonson (1989) incorporated the attraction effect in a multinomial logit model with just three independent variables:

- Utility of an alternative (U), which is computed by using the customer's preference weights and evaluations of the products on the two attributes
- Dummy which indicates if the alternative is a dominant alternative (DOM)
- HDOM: Dummy for the interaction between DOM and dummy variable for the choice condition (people were distributed to the high and low justification conditions. These are conditions that manipulate someone's need to explain their choices). This variable was added to the choice model to monitor the influence of the need for justification from others on the asymmetric dominance effect.

The following regression equation was obtained for the model (all variables were significant):

$$0.82U + 1.17DOM + 0.44HDOM$$

Rooderkerk et al. (2008) found a way to capture three context effects (attraction, compromise and similarity effects) in one choice model. His utility equation consists of a context-free utility component and of a context-dependent utility component:

$$z_{hti}^{A} = \underbrace{V_{hti}}_{\text{context-free}} + \underbrace{VC_{hti}^{A}}_{\text{context-dependent}} + \varepsilon_{hti}$$

where z_{hti}^{A} is the utility of individual h in choice set t for alternative i in the context A. The context-free component is a standard utility equation with beta's and attribute evaluations for each attribute q:

$$V_{hti} = \sum_{q=1}^{Q} \beta_{hq} \cdot x_{htiq}$$

The context-dependent utility depends on the size of the similarity, attraction and compromise effects, as well as their corresponding beta's:

$$VC_{hti}^{A} = \beta_{h}^{SIM} \cdot SIM_{hti}^{A} + \beta_{h}^{ATT} \cdot ATT_{hti}^{A} + \beta_{h}^{COM} \cdot COM_{hti}^{A}$$

The sizes of the three context effects are each calculated by using the distances between all the attributes in the choice set as an input. The distance between alternative i and j on attribute q in context A is given by the following formula:

$$d_{ij}^{A} = \sum_{q=1}^{Q} \frac{\left| x_{iq} - x_{jq} \right|}{L_{q}^{A}}$$

The size of the three context effects are all calculated in a different way (Rooderkerk et al, 2008). This model is totally different from the model of Mishra et al. (1993), where the attraction effect was found by calculating the difference between the observed market share and the expected market share of a product under the Luce model.

Rooderkerk et al. (2010) developed a slightly different model than the one developed in 2008. The distance variable is calculated by computing the Euclidean distance between different alternatives. Aside from the alternative distance formula, the attraction and similarity effects are calculated in a different way. In addition, an interaction effect between attraction and similarity was added to the model.

Another model that incorporates multiple context effects is the model by Tversky & Simonson (1993). Just like Rooderkerk et al. (2008), Tversky & Simonson (1993) split up the utility equation into a context-free and a context-dependent part:

$$V_B(x,S) = v(x) + \beta\, f_B(x) + \theta\, g(x,S)$$

$V_B(x,S)$ is the utility of alternative x, with choice set S and background context B. $v(x)$ is the context-free part of this formula, which consists of a simple utility equation:

$$v(x) = \sum_{i=1}^{n} v_i(x_i)$$

This equation simply adds the utility for each attribute value x_i up to get the total context-free utility of alternative x. The rest of the total utility equation is the context-dependent part of the utility of x. $f_B(x)$ contains the effect of the background (the influence of the observation of earlier choice sets), while $g(x,S)$ contains the impact of the different alternatives in the choice set. The effect of the background can be considered as a change in the relative weights of the attributes:

$$f_B(x) = \sum_{i=1}^{n} b_i v_i(x_i)$$

When the first two parts of the total utility equation are combined, the result is the following equation:

$$v(x) + \beta\, f_B(x) = \sum_{i=1}^{n} \beta_i v_i(x_i) \text{ where } \beta_i = 1 + \beta b_i$$

β_i can be seen as "a combination of the intrinsic weight of that attribute and the weight of the background" (Tversky & Simonson, 1993). The last component of the total utility equation is a bit harder to calculate. First, the advantage of product x over product y has to be calculated for each product y, and on each attribute i:

$$A_i(x,y)=\begin{cases} v_i(x_i)-v_i(y_i) & \text{if} \quad v_i(x_i)\geq v_i(y_i) \\ 0 & \text{otherwise} \end{cases}$$

Then the overall advantage of product x over product y can be calculated by summing up these attribute-related advantages. Similarly, the overall disadvantage of x w.r.t. y has to be calculated:

$$A(x,y)=\sum_{i=1}^{n} A_i(x,y)$$

$$D(x,y)=\sum_{i=1}^{n} D_i(x,y)$$

The relative advantage of option x w.r.t. y is then calculated with the following formula:

$$R(x,y)=\frac{A(x,y)}{A(x,y)+D(x,y)}$$

The last component, $g(x,S)$, of the utility equation, is equal to $R(x,y)$ if more than two alternatives are present in the choice set, and equal to 0 if not. Combining all these formulas, the formula below gives the "componential context model" of Tversky & Simonson (1993):

$$V_B(x,S)=\sum_{i=1}^{n} \beta_i v_i(x_i)+\theta\sum_{y\in S} R(x,y)$$

θ indicates the extent to which the relationship context influences the utility of an alternative. Attraction and compromise effects are indirectly incorporated in this model. The local context (the last part of the formula) makes sure that tradeoff contrast and extremeness aversion are present in the model. Tradeoff contrast is an important driver of the attraction effect (or asymmetric dominance effect), while extremeness aversion is an important driver of the compromise effect. In that way, both effects are present in the model.

Bettman et al. (1998) talked about possible determinants of θ. They hypothesized that θ is positively influenced by the weight that consumers place on effort and justification, while being negatively influence by the weight that consumers place on accuracy. In addition, θ should be positively related to transparency of the relationships between the different alternatives and θ should be lower when more attributes and alternatives are present in the choice set.

Chernev (2004) extended the componential context model of Tversky & Simonson (1993) by incorporating the attribute balance effect as the third part of the utility formula:

$$V_B(x,S) = \sum_{i=1}^{n} \beta_i v_i(x_i) + \theta \sum_{y \in S} R(x,y) + \delta \sum_{\substack{i,j=1 \\ i>j}}^{n} \gamma_{ij} \left| x_i - x_j \right|$$

where x_i and x_j are two different standardized attribute values of the same alternative and δ is the weight of the attribute balance effect.

Huang (2009) adjusted two choice models: prospect theory and the componential context model. Huang (2009) argued that prospect theory lacks the ability to incorporate the compromise effect, because the model does not incorporate direct comparisons between different alternatives. Prospect theory assumes that every alternative is evaluated is compared with a reference point for each attribute. The original prospect theory includes the following formulas to compute the utility of an alternative:

$$v(x_i) = \begin{cases} (x_i - r_i)^{\alpha_i} & \text{if } x_i \geq r_i \\ -\lambda_i (r_i - x_i)^{\beta_i} & \text{if } x_i < r_i \end{cases}$$

$$v(\mathbf{x}) = \sum_{i=1}^{n} w_i v(x_i)$$

where $v(x_i)$ is the value of attribute i for alternative x, $v(\mathbf{x})$ is the total utility of alternative x, w_i is the weight of attribute i, α_i and β_i are the convexity and concavity parameters for attribute i, r_i is the reference point for attribute i and λ_i is the loss aversion parameter for attribute i, which is assumed to be greater than 1. If α_i = β_i and λ_i >1, then loss aversion is incorporated in the model, because decreases in attribute value below the reference point will have a greater impact (>1) than decreases in attribute value above the reference point (1). However, Huang (2009) shows by a simple theoretical example that this model can still result in people choosing the extreme alternative. Therefore, the model was adjusted to incorporate comparisons between different alternatives:

$$v'(x_i) = \begin{cases} \sum_{s=1}^{m} (\Delta x_{is} - r_i)^{\alpha_i} & \text{if } \Delta x_{is} \geq r_i \\ -\lambda_i \sum_{s=1}^{m} (r_i - \Delta x_{is})^{\beta_i} & \text{if } \Delta x_{is} < r_i \end{cases}$$

where Δx_{is} is the difference in attribute value of attribute i between alternative x and s.

Huang (2009) also found that the componential context model (Tversky & Simonson, 1993) lacked the ability to incorporate reference-dependent choices in the model. For this reason, Huang (2009) argued that the model may not be sufficient when modelling asymmetric dominance effects. Therefore, Huang adjusted the componential context model to incorporate reference-dependent choices. To do this, he adjusted the formula to calculate the advantage of x over y with respect to the i th attribute:

$$A'(x_i, y_i) = \begin{cases} v_i(x_i - r_i) - v_i(y_i - r_i) & \text{if } x_i \geq y_i \geq r_i \\ v_i(x_i - r_i) - \gamma_i v_i(y_i - r_i) & \text{if } x_i \geq r_i \geq y_i \\ \gamma_i v_i(x_i - r_i) - \gamma_i v_i(y_i - r_i) & \text{if } r_i \geq x_i \geq y_i \end{cases}$$

where "γ_i denotes the coefficient of loss aversion on the i th attribute and r_i is the reference point corresponding to the i th attribute". $D'(x_i, y_i)$ is an increasing and convex function of the corresponding advantage $A'(y_i, x_i)$.

Kivetz et al. (2003) discussed several methods of incorporating the compromise effect in a choice model:

- The Contextual Concavity Model (CCM)
- The Normalized Contextual Concavity Model (NCCM)
- The Relative Advantage Model (RAM)
- The Loss Aversion Model (LAM)

CCM and NCCM explain the compromise effect by looking at diminishing returns. Their assumption is that utility as a function of attribute values is a concave function. Kivetz et al. (2003) did two studies to test the different models. In the first study, there were three choice alternatives, with each two attributes. In the second study, there were five choice alternatives with four attributes. In both applications the compromise effect was present, but the compromise effect was larger in the first study. This suggests that the compromise effect is less present in environments with more choice alternatives, and more attributes for each choice alternative.

Roe, Busemeyer & Townsend (2001) and Usher & McClelland (2004) proposed connectionist type models that could explain attraction, compromise and similarity effects. However, I do not consider their models to be relevant for my purpose for two reasons. First, they do not show any validation of the model by performing an experiment. They only used purely theoretical examples of certain situations, but they did not test their models in practice. Second, these neural network models do not provide a way to assess significance of variables.

3 Methodology and data description

This research investigates to what extent context effects are present in real life, non-manipulated decision making. I focus on three context effects:

- Attraction effect, which has been discussed in section 2.1
- Similarity effect, which has also been discussed in section 2.1
- Compromise effect, which has been discussed in section 2.2

To understand attraction, compromise and similarity effects, a literature review of papers that cover these topics has been done. This literature has focused on the definitions of the three effects, on drivers of the attraction and compromise effects and on choice models that can model these effects.

A dataset of a large mortgage recommendations website has been gathered to answer the main research question that has been described at the end of chapter 1. It is the same data as the data that was used in the research of Tsekouras et al. (2011). The dataset contains clickstream data of a website that lets consumers search for mortgages, based on certain parameters. The first part of the dataset contains data about the time that the consumer arrives at the website, and at which page the consumer arrives. The first step for a consumer is to enter the desired type of mortgage, the amount of the mortgage, the duration of the fixed interest rate period the mortgage (if any), the type of job (full-time, part-time, etcetera) of the consumer and whether the mortgage is requested for an existing house or for a house that still has to be built. These parameters are all visible in the database. More parameters have to be filled in by the respondent, but these were not visible in the database.

After filling in these parameters, the consumer gets a list of possible mortgages. This list is an ordered list of recommended mortgage products. Mortgages are sorted by the interest rate, starting with the lowest interest rate. The database displays the nine mortgages that are highest in the list. The brand of each mortgage is displayed, as well as the product name itself. In addition, the interest rate, quality, duration of the mortgage, gross and net payments per month and the broker's commission rate are visible for each of the nine mortgages in the list.

The consumer then has the choice to examine one or more of the suggested mortgages in more detail, to ask for an offer, to search again or to close the website. The action of the consumer is logged into the database.

My own model that is used for the analysis is inspired by the papers that Rooderkerk et al. wrote in 2008 and 2010. It is a model that takes context-free and context-dependent characteristics of the choice set into account to be able to predict consumer decision making. A combination of aspects from both of the models of Rooderkerk et al. has been used as the basis of my statistical model. The model was then slightly altered to incorporate importance weights in the calculation of the context effects and to incorporate more interactions in the model. A detailed description of my research model is given in chapter 5.

Before using the model, I had to prepare the data. Data about consumer’s input, data about the choice set and data of the actions of the consumer (asking for an offer or asking for more details) were in three different tables. These tables had to be joined, even though the “Action” table wasn’t connected to any of the other tables by a unique ID. This proved to be a tough job. A detailed description of the data preparation is given in the next chapter.

Scope

To narrow the scope, I’ve set a few boundaries for the research:

- This research focuses on individual decision making. Group decision making has not been considered in this research.
- No experimental research has been done. Instead of this, the basis for answering the research questions was a statistical analysis on a huge database that contains consumer's behaviour on a mortgage recommendations website.
- Due to the database, I only focused on true, real-life data. Possible moderators like “need for justification” or “familiarity with the product class” cannot be taken into account for this research, due to the nature of the data set.

4 Data preparation

Five tables were present in the Excel dataset of the online mortgage recommendation website. The first two tables were irrelevant, because they did not contain data that was directly related to customers' choices. The other three tables were relevant. I have given them the names CustomerInput, SearchResult and Action. The tables can be seen as steps that the customers took in a sequential order:

- The customers enter details about themselves and the type of mortgage they desire in an input form.
- The customers see the list with the characteristics of the possible mortgages on their screen. The mortgages that are on the screen are the direct result of the forms that were filled in by the customers.
- The customers choose to examine one or more mortgages in detail, or the customers choose to request an offer from one or more mortgage providers.

The relevant variables of the three tables are described in Appendix 1. The goal of the data preparation was to be able to connect customer preferences and characteristics of the choice set directly to the choices that customers made. This allowed me to create a multinomial logit model in EViews, because the model needs characteristics of the choice set to be connected to the actual choices to estimate the influences of the product's characteristics and the influences of the context variables on choice behavior. I chose to use SQL to prepare the dataset. I started my data preparation with importing the different tables in Microsoft Access. Each row in the three tables was given a unique ID.

The first step was to delete all the redundant rows. In some cases, customers pressed the same button multiple times, resulting in several identical rows that contained exactly the same data. First, I started with deleting all the redundant rows in the Action table. 10897 of the 43788 rows in Action were deleted, which means that 32891 rows were left. Second, I deleted all the duplicate rows of the CustomerInput table. 8912 of the 45950 rows were deleted and 37038 rows were left. No rows were deleted in the SearchResult table itself, but only relevant rows of SearchResult were used to join the different tables.

The next step was to connect the tables to each other. There were several conditions that needed to be fulfilled for joining the rows in the different tables. The CustomerInput table and the SearchResult table could easily be connected to each other. The variable Transaction2 in the SearchResult table directly referred to the Transaction variable of the CustomerInput table. Connecting these tables to the Action table was a lot harder. The Action table did not have a unique variable that connected a row in the Action table with a row in the CustomerInput table or the SearchResult table. Therefore, I had to find a more complex method of linking the Action table to the other two tables.

First, I connected the three tables to each other by connecting the SearchResult table to the Action table via the Session variable that was present in both tables. The Session value for Action had to be equal to the Session value for SearchResult. Of course using only this condition would result in incorrect data, because there can be multiple actions and multiple search results in one session. Therefore, I had to add more conditions to connect the tables. Each row in Action contained a variable called "Position", which is the position of the chosen item in the choice set. For joining the tables, I tested whether the brand name, product name, and interest rate of the chosen product was the same as the brand name, product name and interest rate of the product in the SearchResult table that corresponds to the Position number in the Action table. Furthermore, the variables Amount, Type and Period of the Action table had to be equal to the Amount, Type and Period variables of the CustomerInput table and the brand names of the 9 products in the choice set could not have an empty value. Joining the three tables resulted in a new table with 27748 rows, with each row containing the input of the customer, the resulting choice set and the action of the customer (requesting an offer or requesting more details of a specific mortgage product).

The next step was to separate the table with the offers from the table with the details. There was some inconsistent data in the choice set, meaning that requesting an Offer was indicated by either an "O", "z", " Z", "O_o_d", "O_o_t" or "O-o". These values were all labelled as "O" for the sake of consistency. I then created two separate tables for Offer and Details. The table for Offer included 4798 rows and the table for Details included 12023 rows. Both tables were exported to Excel and then imported back to Access, because I wanted to add a unique ID to each row in the Offer and Details tables.

Deleting double rows in the Action and CustomerInput tables didn't result in Offer and Details tables without double rows. There were still double rows present. One large problem was that one specific action (with a unique ActionID) appeared more than once in the table. The cause of this problem was that there were multiple search results where one Action could be matched with. The reason for this was probably that customers entered two slightly different input forms, resulting in two similar choice sets that both matched with one specific Action. I solved this problem by checking if the

ActionID of a row had been seen earlier in the table. If this was the case, then this particular row was deleted. This decreased the number of rows of the Details table from 12023 to 9183 and the number of rows of the Offer table decreased from 4798 to 3796.

After this operation, almost all of the duplicate rows were removed. However, there was still one source of double rows present in the data. I found out that, for a small number of customer inputs, there were actually 2 search results connected to it. To speak in more official terms, there were some rows where the CustomerInput.Transaction variable connected to 2 rows in the SearchResult table that had the same SearchResult.Transaction2 value. I am not aware of the cause for this problem, but I managed to solve it by deleting rows every time that a second SearchResultID had been found that corresponded to the same CustomerInputID. After this operation, the amount of rows in the Details table decreased from 9183 to 9076 and the amount of rows in the Offer table decreased from from 3796 to 3785.

After this, I could not find any double rows anymore. However, there were some customers that asked for multiple offers from the same choice set, and customers that asked for details of multiple products in a choice set. This may influence the results of the model. Therefore, I added an extra binary variable called MultipleActions that was set at 1 if the same CustomerInputID has been found more than once in the table. 3727 of the 9076 rows of the Details table received a 1 for MultipleActions and 661 of the 3785 rows of the Offer table received a 1 for MultipleActions.

Somehow, the original dataset did not include Gross and Net payments for every row in the SearchResult table. Therefore, the final step was to only select rows that did not have any empty values for the Gross and Net payments. After this selection, 5193 of the 9076 rows of the Details table remained and 2237 of the 3785 rows of the Offer table remained.

The data preparation was done with SQL queries. All the queries that I used can be found in Appendix 2. After executing all of these queries, the tables were ready to be analysed statistically. The two resulting tables were exported to Excel and the analysis was done in EViews. Before describing the results of the analysis, the research model is described in the next chapter.

5 Research model

To investigate whether context effects are present in real-life, a model was necessary that has a context-free utility component and a context-dependent utility component. Furthermore, I wanted to incorporate compromise effects, attraction effects and similarity effects in one model. The significance of these effects had to be tested as well. Both of the models of Rooderkerk et al. (2008 and 2010) fulfilled these requirements. First, I will explain the relevant parts of both models and I will explain why I combined these two models to be the basis of my own model. Second, I will explain what adjustments I made to this “basis” that Rooderkerk et al. provided.

5.1 *Used components from existing literature*

The following utility equation was provided by Rooderkerk et al. (2008):

$$z_{hti}^{A} = \underbrace{V_{hti}}_{\text{context-free}} + \underbrace{VC_{hti}^{A}}_{\text{context-dependent}} + \varepsilon_{hti}$$

where z_{hti}^{A} is the utility of individual h in choice set t for alternative i in the context A. The context-free component is a standard utility equation with beta's and attribute evaluations for each attribute q:

$$V_{hti} = \sum_{q=1}^{Q} \beta_{hq} \cdot x_{htiq}$$

In this formula, each individual has its own Beta. However, for most individuals there has only been one choice set in my data file. Therefore, I have not used individual-specific Beta’s. In Rooderkerk et al. (2010), the contextual component was described with the following formula:

$$VC_{hi}^{S} = \beta_{h}^{COM} COM_{i}^{S} + \beta_{h}^{ATT} ATT_{i}^{S} + \beta_{h}^{SIM} SIM_{i}^{S} + \beta_{h}^{ATT*SIM} * ATT * SIM_{i}^{S}$$

This means that the context-dependent utility depends on the size of the attraction effect, the compromise effect, the similarity effect, the interaction between attraction and similarity and their corresponding Beta's. In this formula, S is the variable that describes the context (the choice set). The compromise effect was the easiest context variable to calculate in the paper of Rooderkerk et al (2010). First, a middle point had to be defined for each variable in this choice set. These middle points are the averages of the minimum and maximum values of each attribute in a particular choice set S:

$$x_{Mq}^{S} = \frac{\min_{k \in S} x_{kq} + \max_{k \in S} x_{kq}}{2}, \quad q = 1, \ldots, Q.$$

The compromise effect is then calculated by taking the negative (Euclidean) distance of an alternative to the middle point (Rooderkerk et al, 2010):

$$COM_i^S = -d_{iM}^S = -\sqrt{\sum_{q=1}^{Q} \left(x_{iq} - x_{Mq}^S\right)^2}$$

If a variable has only two dimensions, then the distance can be displayed in a two-dimensional graph (Rooderkerk et al, 2010):

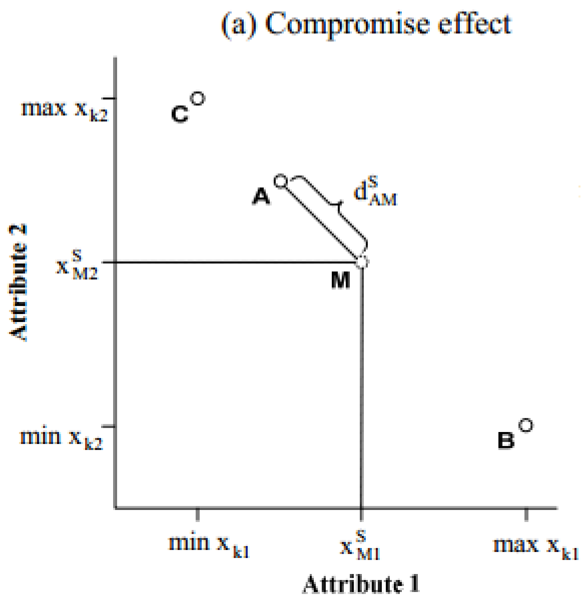


Figure 11: The compromise effect if each alternative has only two attributes

Most literature suggests that being the middle option in the choice set is an advantage; the compromise effect should be an effect that increases the choice probability of middle alternatives. The research of Lehmann & Pan (1994) was an exception; there was a negative compromise effect in one of their experiments. This was explained by the size of the choice set that was relatively large. However, since this was the only exception I still expected that the compromise effect in my experiment would be positive.

Calculating the other context effects is a bit more complicated. To calculate the similarity and the attraction effect, Rooderkerk et al. (2010) defined two vectors: the positioning vector and the preference vector. The preference vector is the vector that Rooderkerk et al. described with the following formula:

$$v^S_{preference} = \left[\left(\max_{k \in S} x_{k1} - \min_{k \in S} x_{k1} \right), \ldots\ldots, \left(\max_{k \in S} x_{kQ} - \min_{k \in S} x_{kQ} \right) \right]$$

If the maximum value is (utility-wise) the least attractive value, then the negative of that value may be taken as input for the model, to ensure that each maximum value is the most attractive value. The positioning vector is the vector orthogonal to the preference vector. Both vectors are displayed in a two-dimensional graph below in a choice set with three alternatives that all have 2 attributes:

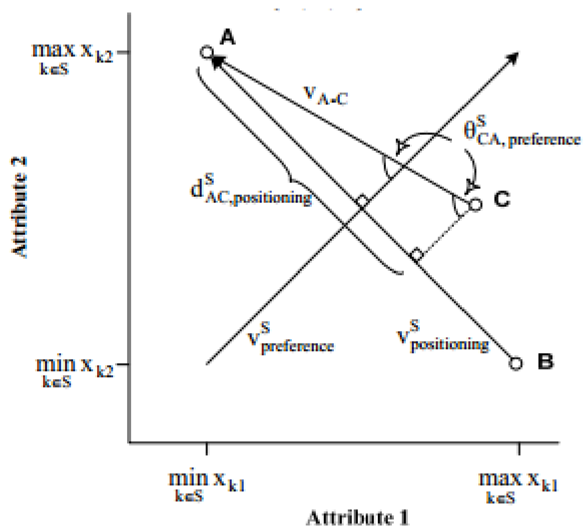


Figure 12: Positioning vector and preference vector

The positioning vector is used to calculate the similarity effect (SIM). The distance between two variables on the positioning vector indicates the degree to which two variables are considered similar. The distance on this vector can be computer by using a trigonometric formula: the sinus of the angle between the difference vector and the preference vector is multiplied by the distance between two alternatives (Rooderkerk et al, 2010):

$$d^{S}_{AC,\,positioning} = \sin\left(\theta^{S}_{CA,\,preference}\right) * d_{AC} = \sin\left(\theta^{S}_{CA,\,preference}\right) * \left\| v_{A-C} \right\|$$

where $\left\| v_{A-C} \right\| = \sqrt{\sum_{q=1}^{Q} \left(x_{Aq} - x_{Cq} \right)^{2}}$

The formula to calculate the angle is

$$\theta^{S}_{CA,\,preference} = \arccos\left(\frac{\left\langle v_{A-C}, v^{S}_{preference} \right\rangle}{\left\| v_{A-C} \right\| \left\| v^{S}_{preference} \right\|} \right)$$

$\left\| v^{S}_{preference} \right\|$ is the Euclidean norm of the preference vector, which can be calculated with almost the same formula as the Euclidean distance between variables A and C. The only difference is that the minimum of a certain attribute is subtracted from the maximum of that particular attribute, instead of subtracting the attribute values of two particular alternatives from each other:

$$\left\| v^{S}_{preference} \right\| = \sqrt{\sum_{q=1}^{Q} \left(\max_{k \in S} x_{kq} - \min_{k \in S} x_{kq} \right)^{2}}$$

Finally, the inner product of the difference vector and the preference vector has to be calculated. This is done by using the following formula:

$$\left\langle v_{A-C}, v^{S}_{preference} \right\rangle = \sum_{q=1}^{Q} \left(x_{Aq} - x_{Cq} \right) \left(\max_{k \in S} x_{kq} - \min_{k \in S} x_{kq} \right)$$

There was a situation possible in which the formula for calculating the angle could not be calculated by Excel, because the arccosine value can only be calculated if the input values are between -1 and 1.

In some cases, the formula gave an erroneous result in Excel. This happened when alternative A had the maximum value (most attractive value, utility-wise) for all attributes and C had the minimum value for all attributes, or the other way around. In that case, the two alternatives are at the opposite ends of the preference vector. I corrected this in Excel by setting the distance on the positioning vector as 0 whenever this occurred, because the distance on the positioning vector is 0 if both alternatives are on the preference vector. This can be clarified by looking at figure 10. In such a situation, both of the alternatives will be on the preference vector. The positioning vector is orthogonal to the preference vector, so the two alternatives will have a distance of zero on the positioning vector.

The formulas that were described above are the formulas to calculate the distance on the positioning vector between alternative A and C. The lower this value, the more similar these two alternatives are. However, there will always be more than 2 alternatives when you are using this formula. Therefore, the similarity effect is the minimum of the distances between alternative i and each of the other alternatives (Rooderkerk et al, 2010):

$$SIM_i^S = \min_{j \in S} d_{ij,\, positioning}^S$$

The similarity effect in my model only looks at differences in positioning. Therefore, if a SIM value is 0, then this doesn't have to mean that there are 2 alternatives in the choice set that are exactly the same in terms of attribute values. Alternatives are expected to have a lower choice probability when there is a lot of similarity with the closest alternative. Therefore, I expect the similarity effect to negatively influence the choice probability.

The difference on the preference vector can be calculated by using the following formula (Rooderkerk et al, 2010):

$$d_{AC,\, preference}^S = \frac{\left\langle v_{A-C}, v_{preference}^S \right\rangle}{\left\| v_{preference}^S \right\|}$$

The ATT value (size of the attraction effect) was calculated with the following formula in the paper of Rooderkerk et al (2010):

$$ATT_i^S = \begin{cases} d_{ij,preference}^S & \text{if item } i \text{ dominates item } j \text{ in set } S \\ -d_{ij,preference}^S & \text{if item } i \text{ is dominated by item } j \text{ in set } S \\ 0 & \text{if item } i \text{ is neither dominating or dominated in set S} \end{cases}$$

The problem with this formula is that is seems to assume that each alternative can only dominate one other alternative. In my dataset, each alternative can dominate multiple alternatives. Furthermore, even if I would adjust this by making the ATT variable the distance on the preference vector to the most dominant item, then the ATT formula is still a linear formula which is likely to have a high correlation with the Rate, Afsluitprovisie (broker's commission percentage), Quality and Net variables. Therefore, I chose to use the formula for ATT that Rooderkerk published in 2008.

Some other formulas of Rooderkerk et al. (2008) have to be discussed before giving the formula for ATT. First, Rooderkerk et al. (2008) defined the following formula for calculating the distance between two alternatives:

$$d_{ij}^A = \sum_{q=1}^{Q} \frac{|x_{iq} - x_{jq}|}{L_q^A}$$

where L_q^A is the range of attribute q in set A. The next formula calculates the advantage of option i over option j:

$$r_{ij}^A = \sum_{q=1}^{Q} \frac{\max\{x_{iq} - x_{jq}, 0\}}{L_q^A}$$

This formula calculates "the sum of attribute differences in favor of option i, scaled to the corresponding attribute range (Rooderkerk et al, 2008).The next formula gives the relative advantage of item i over item j, which can be defined as "the fraction of the distance between i and j that represents an advantage for item i":

$$R_{ij}^A = \frac{r_{ij}^A}{d_{ij}^A}$$

Finally, the attraction effect of alternative i in choice set A can be calculated by using the following formula:

$$ATT_i^A = \frac{2}{N_A - 1} * \sum_{j \in A \setminus \{i\}} R_{ij}^A$$

where N_A is the number of alternatives in choice set A. Obviously, the attraction effect is expected to be positive; if an alternative is dominant, then this increases the probability of being chosen.

The last component that Rooderkerk et al. (2010) included in their model was an interaction effect between attraction and similarity. They stated that literature explained that dominating items should have a higher choice probability when the decoy is close to the target. Therefore, Rooderkerk et al. specified an interaction variable for their model:

$$ATT * SIM_i^S = \begin{cases} d_{ij,preference}^S * d_{ij,positioning} & \text{if item i dominates item j in set S} \\ -d_{ij,preference}^S * d_{ij,positioning} & \text{if item i is dominated by item j in set S} \\ 0 & \text{if item i is neither dominating nor dominated in set S} \end{cases}$$

However, this formula is not suitable for my model for the same reason that I didn't use the formula for ATT that Rooderkerk published in 2010; an alternative has the possibility to dominate more than one alternative in my dataset. Therefore, my formula for the interaction effect is the ATT effect of Rooderkerk et al. from 2008, multiplied by the distance on the positioning vector between alternative i and the most dominant option in the choice set:

$$ATT_i^S * SIMHIGH_i^S = ATT_i^S * d_{ia,positioning}$$

where a is the index value of the most dominant option. Clearly, the modeled effect is very different from the effect modeled in the paper of Rooderkerk et al. (2010). I expect that, the greater the distance between an alternative i and the most attractive option, the greater the attraction effect of option i will be.

5.2 *Own additions to the model*

First, I wanted to add three more interactions to my model. In SQL I created the MultipleActions variable that gives a 1 if there are multiple items chosen from one choice set. In my model, I included interactions between MultipleActions and the ATT, SIM and COM variables. Possibly, these context effects are weakened when multiple items are chosen from one choice set.

Second, when I tested the dataset with a simple multinomial logit model without context variables, I found that Rate was by far the most important variable. In addition, the Net payment per month hardly had any influence. However, in Rooderkerk's models, all these variables are weighed to be equally important in the calculation of the context effets. Therefore, I decided to do a two-step analysis. First, a simple multinomial logit model was run:

$$Utility_i = \beta^{Rate} * Rate_i + \beta^{Quality} * Quality_i + \beta^{BC} * BC_i + \beta^{Net} * Net_i$$

where BC_i is the broker's commission rate, labeled in the dataset as "afsluitprovisie". This step was taken to find the importances of the different product attributes. These importances are used as input for the calculation of the context effects. The gross payment per month is not included in the formula, because it's considered to be too similar to the Net variable.

Step two is to calculate the total model. All the context effects will be included in the total model. There is one important difference to the context variables of Rooderkerk et al. (2008, 2010) that should be taken into account. Each of the four variables will be multiplied with their corresponding Beta's, but only when these four variables are used as inputs for the formula's that calculate context effects.

This ensures that importance weights are taken into account in the calculation of the context effects. For the calculation of the context-independent utility, the original values of the four variables will be used. Only for calculating ATT, SIM, COM and the corresponding interactions, each of the four input variables will be multiplied with the Beta's that were found in step 1. The variable SIMHIGH in my model is the distance on the positioning vector between alternative i and the most attractive alternative.

The final model has the following formula:

$$
\begin{aligned}
Utility_i = {} & C_i + \beta^{Rate} * Rate_i + \beta^{Quality} * Quality_i + \beta^{BC} * BC_i + \beta^{Net} * Net_i + \beta^{COM} * COM_i \\
& + \beta^{SIM} * SIM_i + \beta^{ATT} * ATT_i + \beta^{ATT*SIM} * ATT_i * SIMHIGH_i \\
& + \beta^{MultipleActions*COM} * MultipleActions * COM_i + \beta^{MultipleActions*SIM} * MultipleActions * SIM_i \\
& + \beta^{MultipleActions*ATT} * MultipleActions * ATT_i
\end{aligned}
$$

In the next chapter, the results of this model will be presented. In addition, explanations are given for the (non-)occurrence of certain effects.

6 Results

In the analysis observations have only been included when the Amount of the mortgage was lower than one million euro's, to make sure that unrealistic amounts were not taken into account. This meant that 5165 of the 5193 rows were included for Details and 2214 of the 2237 rows were included for Offer. Furthermore, the amount variable was divided by 100000, because EViews didn't run the model without this modification.

In this chapter, the results of the simple multinomial logit model (only product characteristics) is described in section 6.1. In section 6.2, the results of the full model (including context effects and several interaction variables) are explained. In section 6.3, a comparison between the results for Offer and for Details will be made. Finally, in section 6.4 I will examine whether the presence of context effects actually leads to worse choices.

6.1 *Results of the simple multinomial logit model*

First I ran the simple multinomial logit model to find the four important weights. The model was run for Details and for Offer. Running the model in EViews gave us the following Beta's:

	Details		Offer	
	Coefficient	P-value	Coefficient	P-value
β^{Rate}	-9,145	0,000	-15,059	0,000
$\beta^{Quality}$	0,647	0,000	0,708	0,000
β^{BC}	-80,193	0,000	-26,377	0,000
β^{Net}	-0,007	0,001	-0,003	0,249

Table 1: Importance weights for the four product attributes

All of the variables have the expected direction and significance, except for the Net variable in the offer table. I would have expected this variable to be significant as well. The reason for the non-significance is probably that there is a high correlation between Rate and Net (around 42% on average)

6.2 *Results of the models that include context effects*

Three more models were calculated for Details and for Offer. Model 1 is a model with only constants (C_1 through C_8) and product attributes. Model 2 is a model that includes constants, product attributes and context effects (which use the Beta's of the previous table as inputs). Model 3 is the final model that also includes interactions between contexts effects and MultipleActions (the variable that describes if multiple items are chosen from the choice set). One row generated an error in the Details data set, because the SIM value of that particular row could not be calculated. This row was deleted, which leaves 5164 rows remaining for Details (and 2214 for Offer).

6.2.1 Descriptive statistics for the Offer table

Before showing the results of the analysis, some descriptive statistics are shown below. First, the means and standard deviations for all of the four product characteristics is given. These values have been calculated separately for each of the nine positions in the choice set; the average of Rate01 through Rate09 is given, as well as the average of Quality01 through Quality09, etcetera. In addition, in each row of the data, one item has been chosen and eight items have not been chosen. For both the chosen product (CP) and all the not-chosen products (NP), the means and standard deviations have been calculated as well:

	Rate		Quality		BC		Net	
Product	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
01	4,19	0,59	2,43	0,49	0,73%	0,42%	452	226
02	4,22	0,60	2,48	0,50	0,69%	0,45%	455	228
03	4,25	0,60	2,37	0,48	0,80%	0,38%	458	230
04	4,29	0,61	2,66	0,47	0,63%	0,45%	463	233
05	4,32	0,61	2,64	0,48	0,67%	0,45%	466	234
06	4,35	0,62	2,83	0,37	0,64%	0,48%	469	235
07	4,37	0,61	2,73	0,44	0,66%	0,45%	471	236
08	4,39	0,61	2,74	0,44	0,62%	0,44%	473	237
09	4,40	0,62	2,78	0,41	0,64%	0,45%	475	238
CP	4,23	0,59	2,57	0,49	0,69%	0,44%	456	228
NP	4,32	0,61	2,64	0,15	0,68%	0,13%	466	234

Table 2: Descriptive statistics for the context-independent attributes of the Offer table

From the table, it is clearly visible that the average Rate and Net values go up for each next position in the choice set. This is logical, because products with the lowest interest rate were always recommended first by the recommendation website. However, it is noteworthy that product 01 through product 03 have the worst average Quality and broker's commission rate (BC). On the first sight, it may seem remarkable that the chosen product (CP) has a lower average Quality value and a higher average BC than the products that are not chosen (NP). However, this can be clearly explained by the fact that people chose for product 01 or product 02 in 65% of the cases; these are the products that have a relatively low average Quality and a relatively high average BC. The low standard deviation for NP for Quality and for BC can be explained by the fact that averages of the 8 not-chosen products have been calculated for each of the rows; standard deviations of averages are usually lower than standard deviations of single variables. Below the descriptive statistics of the context variables of the Offer table can be found:

	ATT		COM		SIM		SIMHIGH	
Product	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
01	1,63	0,33	-1,67	1,02	0,20	0,22167	0,37	0,47
02	1,49	0,35	-1,28	1,00	0,12	0,13382	0,37	0,37
03	1,21	0,35	-1,01	0,75	0,13	0,14416	0,43	0,32
04	1,11	0,36	-0,74	0,44	0,13	0,12323	0,43	0,32
05	0,88	0,33	-0,77	0,48	0,09	0,11411	0,52	0,34
06	0,76	0,27	-0,99	0,61	0,07	0,11334	0,59	0,37
07	0,50	0,28	-1,22	0,78	0,08	0,11655	0,60	0,39
08	0,34	0,24	-1,46	0,93	0,08	0,12201	0,62	0,41
09	0,25	0,21	-1,67	1,01	0,11	0,12089	0,66	0,44
CP	1,41	0,50	-1,42	1,01	0,16	0,18158	0,37	0,41
NP	0,84	0,08	-1,17	0,66	0,11	0,04784	0,53	0,21

Table 3: Descriptive statistics for the context-dependent attributes of the Offer table

As expected, the average attraction value (ATT) was highest for product 01. In addition, the average compromise value (COM) was highest for the three middle products (in terms of interest rate), which was also expected.

6.2.2 Results of the models for the Offer table

In the table below, the results of running the three EViews models for the Offer table are displayed:

	Model 1		Model 2		Model 3	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
C_1	2,0715	0,000	1,9590	0,000	1,9261	0,000
C_2	1,1748	0,000	1,1128	0,000	1,0996	0,000
C_3	0,9016	0,000	0,8810	0,000	0,8698	0,000
C_4	0,6759	0,000	0,6820	0,002	0,6731	0,002
C_5	0,0542	0,758	0,0684	0,747	0,0619	0,771
C_6	0,2728	0,107	0,2750	0,144	0,2682	0,154
C_7	0,0125	0,946	0,0228	0,905	0,0158	0,934
C_8	0,1212	0,509	0,1291	0,483	0,1257	0,496
β^{Rate}	-5,3538	0,000	-4,7475	0,000	-4,7356	0,000
$\beta^{Quality}$	0,6244	0,000	0,6181	0,000	0,5946	0,000
β^{BC}	-15,3026	0,010	-10,4301	0,153	-10,1773	0,166
β^{Net}	-0,0038	0,159	-0,0033	0,220	-0,0029	0,324
β^{COM}			-0,1295	0,077	-0,2394	0,003
β^{ATT}			0,1448	0,402	0,2841	0,106
β^{SIM}			-0,0348	0,860	-0,0751	0,728
$\beta^{ATT*SIM}$			0,1071	0,061	0,0921	0,110
$\beta^{MultipleActions*COM}$					0,3998	0,000
$\beta^{MultipleActions*ATT}$					-0,5277	0,000
$\beta^{MultipleActions*SIM}$					-0,0030	0,995

Table 4: EViews results for Offer

Constants 1-4 are significant and positive for all three models for Offer. This was to be expected. The constants are significant due to the order of the nine items in the dataset. In the dataset, alternative 1 indicates the alternative with the lowest interest rate and alternative 9 is the alternative with the highest interest rate. β^{Rate} is negative and significant for all the models, as expected. The coefficient would probably have been more negative if the alternatives in the choice set had not been ordered by interest rate. $\beta^{Quality}$ is positive and significant, as expected. β^{BC} is only significant in the first model. The influence of the broker's commission has probably been moved to one or more of the context variables in Model 2 and 3. β^{Net} is insignificant in all models, probably due to the high correlation with the interest rate.

The size of the constants and the size of the Beta-values of the product attributes in models 2 and 3 are quite similar to the values in model 1. The non-significance (and less negative coefficient) of the broker's commission rate is the only big difference between these coefficients in model 1 versus these coefficients in models 2 and 3. This indicates that model 1 gives a pretty solid representation of the influences of the product attributes on utility.

Against expectation, β^{COM} is negative (and significant) for Model 2 and 3. This means that alternatives that were the middle options had a lower utility (on average). Although it was against expectation, it has been seen before. After their experiment, Lehmann & Pan (1994) concluded that "a compromise option has a positive impact when the market is small and a negative one when the number of brands in the market increases to the extent that it is hard to stand out in the crowd". In this case, the market is relatively large as well, because customers can choose between 9 alternatives. This is a possible explanation for the negative compromise effect. It may be hard to be distinct for an alternative that is a middle option in a large choice set.

Model 3 showed a positive interaction between MultipleActions and COM. A Wald test indicated that $\beta^{COM} + \beta^{MultipleActions*COM}$ doesn't significantly differ from 0 (p = 0,1315). This entails that a compromise effect is only present when only 1 item is chosen from the choice set (in 83% of the rows in my dataset, where MultipleActions = 0).

β^{ATT} was positive and insignificant in Model 2 (p = 0,402) and in Model 3 (p = 0,106). In Model 3, ATT is close to being significant on a 10% level. With a bit more observations, ATT might have been significant. The difference in p-values and coefficients between Model 2 and 3 can be explained by the addition of the interaction between ATT and MultipleActions. The coefficient of $\beta^{MultipleActions * ATT}$ was negative (and significant), which is the opposite of β^{ATT}. Even though $\beta^{MultipleActions * ATT}$ was significant, A Wald test that tested if $\beta^{ATT} + \beta^{MultipleActions * ATT}$ was equal to zero showed that there was also no negative attraction effect present when more than one item was chosen from the choice set (p = 0,2292). No similarity effect has been found for Model 2 and Model 3.

Model 2 showed a positive and significant interaction (p = 0,061) between attraction and similarity (ATT and SIMHIGH). This was expected. A high ATT value means that a product is dominant towards other products. An alternative can be dominating other alternatives, even if it's not the most dominant alternative in the choice set. A high SIMHIGH value means that an alternative is differently positioned than the most dominant alternative. This means that the alternative is less attractive than the most dominant option on at least 1 attribute, while also being more attractive than the most dominant option on at least 1 attribute. Therefore, a high SIMHIGH value entails that an alternative is not entirely dominated by the most dominant alternative. For these reasons, we already expected the positive interaction between attraction and similarity; the dominance/attraction of an alternative should become more visible if it is differently positioned than the most dominant alternative. For model 3, the interaction was not significant, although it was close to being significant on a 10% level (p = 0,110). However, it might have been significant if we had more observations.

6.2.3 Descriptive statistics for the Details table

In 6.2.4, the final results of the statistical analyses for the Details table will be shown. Before providing these results, I would like to discuss the descriptive statistics for the details table. These descriptive statistics are shown in the tables below:

	Rate		Quality		BC		Net	
Product	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
1	4,22	0,60	2,40	0,49	0,74%	0,42%	470	244
2	4,25	0,61	2,46	0,50	0,72%	0,44%	473	246
3	4,28	0,61	2,36	0,48	0,80%	0,39%	476	248
4	4,32	0,62	2,66	0,47	0,61%	0,46%	482	251
5	4,35	0,62	2,65	0,48	0,68%	0,45%	484	252
6	4,38	0,63	2,81	0,39	0,64%	0,47%	488	254
7	4,40	0,63	2,74	0,44	0,65%	0,46%	490	255
8	4,42	0,63	2,74	0,44	0,61%	0,44%	492	256
9	4,43	0,63	2,79	0,40	0,65%	0,45%	494	257
CP	4,29	0,61	2,65	0,48	0,56%	0,48%	477	248
NP	4,34	0,62	2,62	0,14	0,69%	0,13%	484	252

Table 5: Descriptive statistics for the context-independent attributes of the Details table

	ATT		COM		SIM		SIMHIGH	
Product	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
1	1,48	0,37	-1,12	0,58	0,27	0,28	0,62	0,54
2	1,38	0,42	-0,94	0,55	0,17	0,19	0,51	0,47
3	1,12	0,43	-0,80	0,39	0,17	0,18	0,55	0,37
4	1,10	0,47	-0,64	0,22	0,18	0,17	0,42	0,37
5	0,86	0,41	-0,67	0,24	0,13	0,17	0,56	0,37
6	0,78	0,37	-0,79	0,34	0,12	0,19	0,67	0,42
7	0,59	0,38	-0,89	0,41	0,13	0,19	0,67	0,42
8	0,48	0,35	-1,01	0,52	0,16	0,21	0,70	0,49
9	0,38	0,32	-1,13	0,57	0,17	0,19	0,69	0,49
CP	1,32	0,51	-0,95	0,54	0,21	0,23	0,50	0,51
NP	0,85	0,08	-0,88	0,37	0,16	0,07	0,61	0,22

Table 6: Descriptive statistics for the context-dependent attributes of the Details table

The means and standard deviations of the context-free attributes are very similar to the numbers that were provided for the Offer table. In addition, the distribution of the context variables (ATT, COM, SIM and SIMHIGH) among the 9 different positions in the choice set (product 01 through product 09) is quite similar between Offer and Details. Comparing the absolute values of the context effects between Offer and Details would not be useful, because both models used different importance weights for calculating the context effects. One interesting difference is that the average BC value for chosen products (CP) is a lot lower for the Details table than for the Offer table (56% vs. 69%). Apparently more weight was attached to the broker's commission rate when more Details were requested. This was also visible in the importance weight for BC in the simple multinomial model, which was three times higher for Details.

6.2.4 Results of the models for the Details table

The same models that were run for the Offer table have also been run for the Details table. Below the results are displayed of running Model 1, 2 and 3 for Details:

	Model 1		Model 2		Model 3	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
C_1	2,3530	0,000	2,0642	0,000	2,0451	0,000
C_2	1,5189	0,000	1,2574	0,000	1,2753	0,000
C_3	1,2112	0,000	1,0087	0,000	1,0373	0,000
C_4	0,8852	0,000	0,7586	0,000	0,7840	0,000
C_5	0,4914	0,000	0,4047	0,001	0,4199	0,000
C_6	0,4903	0,000	0,4312	0,000	0,4407	0,000
C_7	0,3021	0,003	0,2698	0,010	0,2709	0,010
C_8	0,3104	0,002	0,3052	0,002	0,3049	0,002
β^{Rate}	-1,1792	0,002	0,0919	0,845	0,1255	0,788
$\beta^{Quality}$	0,6154	0,000	0,3840	0,000	0,3960	0,000
β^{BC}	-70,7890	0,000	-47,2967	0,000	-48,6356	0,000
β^{Net}	-0,0086	0,000	-0,0072	0,002	-0,0081	0,000
β^{COM}			-0,1773	0,010	-0,6335	0,000
β^{ATT}			0,4660	0,000	0,8240	0,000
β^{SIM}			-0,2705	0,002	-0,2687	0,018
$\beta^{ATT*SIM}$			-0,0211	0,432	0,0001	0,997
$\beta^{MultipleActions*COM}$					0,8151	0,000
$\beta^{MultipleActions*ATT}$					-0,8019	0,000
$\beta^{MultipleActions*SIM}$					-0,0375	0,817

Table 7: EViews results for Details

All the constants were positive and significant for all the models, as expected. β^{Rate} was only significant in Model 1. It is likely that the influence of the Rate has partly been taken over by the influence of the attraction effect in model 2 and model 3. $\beta^{Quality}$ was positive and significant and

β^{BC} was negative and significant for all models, as expected. In contrast to the model that was run for Offer, β^{Net} was significant (and negative), as expected.

All the constants were quite similar across models 1, 2 and 3. Aside from β^{Net} , the size of the Beta's of model 2 and 3 were quite different from the Beta's in model 1. It is likely that part of the influences of the product attributes have moved to the influence of the attraction effect in model 2 and 3.

Again, β^{COM} was significant and negative for both models that include context effects. However, a Wald test shows that a positive compromise exists when MultipleActions = 1, which is the case for 41% of the rows for Details, because $\beta^{COM} + \beta^{MultipleActions*COM}$ = 0,181656 and the P-value was equal to 0,0260. This means that a positive compromise effect was present when people chose multiple items in the choice set. However, the size of the negative compromise effect when MultipleActions = 0 (β^{COM} = -0,6335) is more than three times bigger than the size of the positive compromise effect when MultipleActions = 1 ($\beta^{COM} + \beta^{MultipleActions*COM}$ = 0,181656).

It is not clear what the exact reason is for the fact that the compromise effect was reversed when details were requested for multiple alternatives in the choice set. Perhaps, the people that chose multiple alternatives did this to get more certainty. That could indicate that these people had a risk-averse attitude. People with a risk-averse attitude may be inclined to go for the safe, middle options sooner than people that are not risk-averse. However, this is only speculation; the true reason remains uncertain.

β^{ATT} was positive and significant for both models, as expected. $\beta^{MultipleActions*ATT}$ was negative and significant. A Wald test that tested if $\beta^{ATT} + \beta^{MultipleActions*ATT}$ was equal to 0 gave a P-value of 0,8521. This indicates that the attraction effect was only present when only 1 item was chosen from the choice set.

β^{SIM} was negative and significant for model 2 and 3. This indicates that a higher similarity gives a higher utility value for an alternative. This contradicts most of the theory that I read about the similarity effect, where researchers stated that brands that are close to a new market entrant are likely to suffer most from the entrance of a new competitor in the market. Perhaps choosing an alternative that has other similarly positioned alternatives in the dataset gives the customer a sense of certainty.

Glazer, Kahn & Moore (1991) talked about the lone alternative effect. This effect entails that a lone, distant brand that is far away from a cluster of similar brands, is less likely to be chosen. The reason for this could be that customers use a heuristic, which entails that they intuitively think that they have a higher probability of choosing the right brand when choosing from a large pool of similar options, as opposed to an extreme alternative. This is possibly the reason for the fact that similarity increases the utility in my model for Details.

6.3 Comparison between results of Offer and Details

The models that were run for Offer and Details gave some different results. In this paragraph I will zoom in on the differences between these two models and I will try to give explanations for the differences in coefficients and P-values. The focus of this paragraph will be model 3, since this was the complete model.

The first four constants were significant for Offer and Details in model 3, but the last four constants were only significant for Details. This is contrasted by the fact that Rate is significant for the Offer table, while Rate is insignificant (and has the wrong sign) for the Details table. In both of the models, the constants take away part of the influence of the rate, because the nine alternatives were ordered by interest rate in each choice set. However, all of the influence of the rate has been captured by the constants and by the attraction effect in the Details model, while only part of this influence has been captured in the Offer model. It is hard to find an explanation for this phenomenon, because both models are exactly the same, except for the underlying data. Possibly, the reason could be found by looking at browsing behavior of customers. People that are more involved in the mortgage selection process may be more inclined to ask for an offer, while people that are less involved may only look at details of a mortgage product. The first decision, to ask for an offer, may require a longer time to consider. Therefore, people that request an offer may be more inclined to look at all the attributes of a mortgage (interest rate, broker's commission rate, etcetera), while a person that is only exploring the website (with less involvement) may be inclined to quickly look at whatever the mortgage recommendation website recommends, without instantly looking at all the attributes of a product.

This is a possible explanation, but the cause of this difference may also be due to EViews. Perhaps this statistical program simply reaches a different optimal solution for both models.

The attraction effect had a P-value of 0,000 in the Details model, while it only had a P-value of 0,106 in the Offer model. This effect cannot entirely be explained by the difference in the number of observations between Offer and Details (2214 versus 5164). The P-value of Offer may become significant if more observations would be added to the dataset, but this does not explain this huge difference between the P-values of Offer and Details. In the Details model, the rate is insignificant. Perhaps the influence of the rate is reflected in the β^{ATT} value of the Details model. The coefficient of β^{ATT} is actually three times bigger in Details when compared to the Offer model. Therefore, part of the influence of the rate may have been moved to the β^{ATT} value in the Details model. However, examining the coefficient covariance matrices showed no negative covariances between β^{ATT} and β^{Rate}. A different explanation of the difference in P-values are the different Beta's that were used as importance weights for calculating the ATT variable. These were quite different between the model for the Offer table and the model for Details; the simple multinomial logit model gave different Beta's for Offer and Details. This may also be the reason for the significance of SIM in the Offer model and the insignificance of SIM in the Details model. However, these possible explanations cannot be backed up by statistical data. Therefore, more research is required if we would fully want to understand the interplay between different variables in creating a context-dependent choice model.

6.4 Influence of context effects on customer behaviour

Both of my models have shown that context effects actually exist in choice behavior. The question is: do customers make worse decisions when context effects are present? To investigate this, I looked at two scenarios:

1. The scenario where I assume that customers are not influenced by context effects (by using the context-independent probabilities)
2. The scenario where I assume that customers are influenced by context effects (by using the context-dependent probabilities)

To investigate the influence of the context effects on choice behavior, I took several steps. First, I entered the Beta's that Model 3 gave us in the Excel sheets of the Offer and Details datasets. Second, I computed the context-independent utility and the context-dependent utility by using the Beta's from Model 3 and the original data in the dataset. For the context-independent utility, the size of the context effects was set at zero to exclude these effects from the calculation.

Using these utilities, I computed the expected choice probability for each alternative, for the context-independent scenario and the context-dependent scenario. This was done by using the standard formula for calculating the probability of an alternative in a multinomial logit model.

The next step was to calculate the expected utility for the two scenarios. For both scenario's I calculated the sum of each probability multiplied with the corresponding context-independent utility. This was done, because the context-independent utility should reflect the true utility of a products; customers won't actually benefit from certain context effects when using the product, but they will benefit from low interest rates or a low broker's commission rate. For both Details and Offer I found that the average utility in the context-dependent calculation was slightly higher than the average utility for the context-independent calculation.

The reason for the higher utility is that the choice probability for choosing alternative 1 and 2 is higher in the model that included context effects. This is beneficial to the total expected utility, because due to the lower interest rate, alternatives 1 and 2 are usually the alternatives that have the most (context-independent) utility for the customer. The reason for the higher choice probability in the model that included context effects is the positive influence of the attraction effect; ATT values are very high for alternative 1 and 2.

From this data, I cannot conclude that people make worse decisions due to context effects. Further research could try to investigate this issue in further detail.

7 Conclusions and future research

This research has shown that context effects are present in day-to-day decision making. A negative compromise effect was present in both of my models. A positive attraction effect was present in the Details model. The reason for the positive attraction effect was obvious; dominant alternatives are more attractive to customers. The negative compromise effect contradicted most of the literature, but was most likely present because there were a lot of alternatives present in the choice set. Customers may prefer choosing a brand that clearly positions itself in large choice sets. They may prefer a brand that has several strong attributes and several weak attributes over a brand that has average values on all attributes, when there is a lot of choice between brands.

We can conclude that context effects will often be weaker when multiple choices can be made from one choice set. In the Offer dataset, the compromise effect disappeared when a customer had chosen multiple items in a choice set (in 17% of the cases). Although the positive attraction effect was insignificant in the Offer dataset (p = 0,106), this effect also disappeared for customers that had chosen multiple items. In the Details dataset, the (positive and significant) attraction effect also disappeared when a customer had chosen multiple items in a choice set (in 41% of the cases). However, the compromise effect was positive in that scenario, although more than 3 times weaker than the negative compromise effect that occurred when only 1 item was chosen. Future research should try to examine which factors can attenuate and possibly even reverse directions of context effects.

It was remarkable that the similarity variable in my model for Details was negative and significant; brands with higher similarities (lower SIM values) had higher utilities on average. Possibly this can be explained by the theory of the lone alternative effect of Glazer, Kahn & Moore (1991). They state that a lone, distant brand that is far away from a cluster of similar brands, is less likely to be chosen. A reason could be that customers feel that they have a higher probability of choosing the right product when choosing from a larger set of similar brands.

There are several areas that future researchers could try to explore further. First, it is not clear whether the variables of Rooderkerk et al. (2008, 2010) accurately measure the psychological processes that change behaviour of customers when there are dominant options, compromise options or two similar options present in the choice set. Future research could try to find better ways to measure context effects or to confirm that the variables of Rooderkerk et al. (2008, 2010) are correctly measuring context effects. There were a lot of differences in coefficients between the Offer and the Details dataset. Some (correlated) coefficients were significant for Offer and insignificant for Details and vice versa. Therefore, future research could also investigate how an optimal model can be built when the variables in the model are sometimes correlated with other variables (which is the case in a model with a context-free and a context-dependent component).

My dataset was very specific; it concerned online choice behaviour in the mortgages market. Furthermore, the dataset did not contain actual purchases by customers; only requests for offers and requests for more details were present in the dataset. Future research should explore the presence of context effects in other real-life businesses and it should focus on datasets that include (proof of) monetary transactions.

A negative compromise effect was present for my dataset, possibly because there were a lot of brands involved in the choice set. Future research could investigate the influence of the size of the choice set on the presence of context effects. Furthermore, future research could try to examine if the presence of context effects negatively influences the quality of customers' decisions.

Bibliography

1. S. Mishra, U.N. Umesh, D.E. Stem (1993), Antecedents of the Attraction Effect: An Information-Processing Approach. Journal of Marketing Research, Vol. 30, No.3, p. 331 – 449.

2. R. Kivetz, O. Netzer, V. Srinivasan (2004), Extending Compromise Effect Models to Complex Buying Situations and Other Context Effects. Journal of Marketing Research, Vol. 41, p. 262 – 268.

3. R. Dhar (1997), Consumer Preference for a No-Choice Option, Journal of Consumer Research, Vol. 24, No.2, p. 215 – 231.

4. E. Shafir, I. Simonson, A. Tversky (1993), Reason-based choice, Cognition, Vol. 49, p.11 – 36.

5. I. Simonson (1989), Choice Based on Reasons: The Case of Attraction and Compromise Effects, Journal of Consumer Research, Vol. 16, No.2, p. 158 - 174.

6. S. Rateshwar, A.D. Shocker, D.W. Stewart (1987), Toward Understanding the Attraction Effect: The Implications of Product Stimulus Meaningfulness and Familiarity. Journal of Consumer Research, Vol.13, No.4, p. 520 – 533.

7. J. Huber, C. Puto (1983), Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects, Journal of Consumer Research, Vol.10, No., p. 31 – 44.

8. J. Huber, J.W. Payne, C. Puto (1982), Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. Journal of Consumer Research, Vol. 9, No.1, p. 90 - 98.

9. D. McFadden (1986), *The Choice Theory Approach to Market Research*, Marketing Science, Vol.5, No.4, p. 275 – 297.

10. F. Martinez, E. Cascetta, M. Bierlaire, K.W. Axhausen (2008), *An application of the Constrained Multinomial Logit (CMNL) for modelling dominated choice alternatives.* 8th Swiss Transport Research Conference.

11. R.P. Rooderkerk, H.J. Van Heerde, T.H.A. Bijmolt (2008), *The Impact of Context Effects on Optimal Assortments.*

12. D. Tsekouras, B. Dellaert, B. Donkers (2011), *Disentangling the effects of online shopping decision time on website conversion.*

13. I. Simonson, A. Tversky (1992), *Choice in Context: Tradeoff Contrast and Extremeness Aversion.* Journal of Marketing Research, Vol. 29, p. 281 – 295.

14. A. Tversky, I. Simonson (1993), *Context-dependent preferences*, Management Science, Vol. 39, No.10, p. 1179 – 1189.

15. J.R. Bettman, M.R. Luce, J.W. Payne (1998), *Constructive Consumer Choice Processes*, Journal of Consumer Research, Vol. 25, No.3, p. 187 - 210.

16. B. Wernerfelt (1995), *A Rational Reconstruction of the Compromise Effect: Using Market Data to Infer Utilities*, Journal of Consumer Research, Vol.21, p. 627-633.

17. D.R. Lehmann, Y. Pan (1994), *Context Effects, New Brand Entry, and Consideration Sets*, Journal of Marketing Research, Vol.331, p. 364 - 374.

18. R. Kivetz, O. Netzer, V. Srinivasan (2003), *Alternative models for capturing the compromise effect.*

19. F. Sinn, S.J. Milberg, L.D. Epstein, R.C. Goodstein (2007), *Compromising the compromise effect: Brands matter*, Springer Science + Business Media.

20. C.C. Chang, H.H. Liu (2008), *Information Format-Option Characteristics Compatibility and the Compromise Effect*, Psychology and Marketing, Vol. 25, No.9, p. 881 - 900.

21. H. Pechtl (2009), *Value Structures in a Decoy and Compromise Effect Experiment*, Vol. 28, No. 8, p. 736 - 759.

22. T.B. Heath, S. Chatterjee (1995), *Asymmetric Decoy Effects on Lower-Quality versus Higher-Quality Brands: Meta-analytic and Experimental Evidence*, Journal of Consumer Research, Vol. 22, p. 268 - 283.

23. R. Glazer, B.E. Kahn, W.L. Moore (1991), *The Influence of External Constraints on Brand Choice: The Lone-Alternative Effect*, Journal of Consumer Research, Vol. 18, No.1, p. 119 - 127

24. A. Chernev (2004), *Context Effects without a Context: Attribute Balance as a Reason for Choice*, Journal of Consumer Research, Vol. 32, p. 213 - 223

25. J-J Huang (2009), *Revised behavioural models for riskless consumer choice*, Journal of the Operational Research Society, Vol.60, p. 1237 - 1243.

26. Y-W Ha, S. Park, H-K Ahn (2009), *The Influence of Categorical Attributes on Choice Context Effects*, Journal of Consumer Research, Vol. 36, p. 463-476

27. S. Kim, L. Hasher (2005), *The attraction effect in decision making: Superior performance by older adults,* The Quarterly Journal of Experimental Psychology, Vol. 58, p. 120 - 133

28. R. Meyer, E.J. Johnson (1995), *Empirical Generalizations in the Modeling of Consumer Choice*, Marketing Science, Vol. 14, No. 3

29. R. M. Roe, J.R. Busemeyer, J.T. Townsend (2001), *Multialternative Decision Field Theory: A Dynamic Connectionist Model of Decision Making*, Vol. 108, No.2, p. 370 - 392

30. M. Usher, J.L. McClelland (2004), *Loss Aversion and Inhibition in Dynamical Models of Multialternative Choice*, Psychological Review, Vol. 111, No.3, p. 757 - 769

31. B. Fasolo, R. Misuraca, G.H. McClelland, M. Cardaci (2006), *Animation Attracts: The Attraction Effect in an On-line Shopping Environment*, Psychology & Marketing, Vol. 23, No.10, p. 799 - 811

32. A. Felfernig, B. Gula, G. Leitner, M. Maier, R. Melcher, S. Schippel, E.C. Teppan (2008), *A Dominance Model for the Calculation of Decoy Products in Recommendation Environments*, Persuasive Technology, AISB 2008 Proceeding, Vol.3, p. 43 - 50

33. E.C. Teppan, A. Felfernig (2009), *Asymmetric Dominance- and Compromise Effects in the Financial Services Domain*, 2009 IEEE Conference on Commerce and Enterprise Computing, p. 57 - 63

34. E.C. Teppan, A. Felfernig, K. Isak (2011), *Decoy Effects in Financial Service E-Sales Systems*

35. A. Felfernig, B. Gula, G. Leitner, M. Maier, R. Melcher, E. Teppan (2008), *Persuasion in Knowledge-Based Recommendation*

36. J.R. Doyle (1999), *The Robustness of the Asymmetrically Dominated Effect: Buying Frames, Phantom Alternatives, and In-Store Purchases*, Psychology and Marketing, Vol. 16, No.3, p. 225 - 243

37. S. Sheng, A.M. Parker, K. Nakamoto (2005), *Understanding the Mechanism and Determinants of Compromise Effects*, Psychology and Marketing, Vol. 23, No.7, p. 591 - 609

38. R.P. Rooderkerk, H.J. van Heerde, T.H.A. Bijmolt (2010), *Incorporating Context Effects into a Choice Model*, Forthcoming in the Journal of Marketing Research

Appendix 1 – Description of the three relevant tables of the original dataset

CustomerInput

Variable name	Description
Transaction	Unique transaction ID that every input form gets
Session	Session ID that attaches a unique number to each visit of a customer on the website
Set/Session	This number equals n if it is the nth time that a customer completed the form at a certain session. If a certain transaction is the 4th transaction of a customer in a session, set/session equals 4.
Job Type	The main source of income for a customer (fulltime job, parttime job, pension fund, etcetera)
Amount	The amount of the requested mortgage (in euro's)
Type	The type of mortgage that is requested by the customer
Period	Duration of the fixed interest rate of the mortgage (if any)
House	The type of house

SearchResult

Variable name	Description
Transaction	Unique transaction ID that every search result gets
Session	Session ID that attaches a unique number to each visit of a customer on the website
Date	Date that the page with the search results was displayed
Time	Time that the page with the search results was displayed
Brand01 – Brand09	Brand names of each of the nine mortgages in the choice set
Product01 – Product09	Product names of each of the nine mortgages in the choice set
Afsluitprovisie01 – Afsluitprovisie09	The broker’s commission rate of each of the nine mortgages in the choice set
Gross01 – Gross09	The gross payment per month of each of the nine mortgages in the choice set
Net01 – Net09	The net payment per month of each of the nine mortgages in the choice set
Quality01 – Quality09	Quality of each of the nine mortgages in the choice set
Rate01 – Rate09	Interest rate of each of the nine mortgages in the choice set
Transaction2	Unique transaction ID of the submitted customer input form. This ID has been used to directly connect the CustomerInput table to the SearchResult table

Action

Transaction	Unique transaction ID that every action gets
Session	Session ID that attaches a unique number to each visit of a customer on the website
Date	Date that the offer or details were requested
Brand	Name of the mortgage brand that was chosen
Product	Name of the mortgage product that was chosen
Action	The type of action. “O” means an offer was requested. “D” means that more details about the mortgage were requested.
Position	Position of the mortgage in the choice set. For example: position is 4 if Brand04 in the choice set has been chosen
Rate	Interest rate of the chosen mortgage
Type	The type of mortgage that is requested by the customer
Period	Duration of the fixed interest rate of the mortgage (if any)
Amount	The amount of the requested mortgage (in euro’s)

Appendix 2: SQL queries

Query 1

Purpose: deleting all the double, redundant rows in the Action table

```
DELETE *
FROM [Action]
WHERE ID IN(

SELECT A.ID
FROM Action as A, Action as B
WHERE A.Session = B.Session AND A.Product = B.Product AND A.Rate = B.Rate AND A.Action =
B.Action AND A.Brand = B.Brand AND  A.Period = B.Period AND A.Position = B.Position AND A.Type =
B.Type AND A.Amount = B.Amount
AND A.ID > B.ID
);
```

Query 2

Purpose: deleting all the double, redundant rows in the CustomerInput table

```
DELETE *
FROM CustomerInput
WHERE ID IN(

SELECT A.ID
FROM CustomerInput as A, CustomerInput as B
WHERE A.Session = B.Session AND A.[Job Type] = B.[Job Type] AND A.Amount = B.Amount AND
A.Type = B.Type AND A.Period = B.Period AND A.House = B.House
AND A.ID > B.ID
);
```

Query 3

Purpose: connecting the CustomerInput, SearchResult and Action tables.

```
SELECT *
FROM CustomerInput, SearchResult, [Action]
WHERE CustomerInput.Transaction=SearchResult.Transaction2 And
[SearchResult].[Session]=[Action].[Session] And(

(Action.Brand =SearchResult.Brand01 AND Action.Product =SearchResult.Product01 AND
Action.Position=1 AND Action.Rate = Rate01) OR
(Action.Brand =SearchResult.Brand02 AND Action.Product =SearchResult.Product02 AND
Action.Position=2 AND Action.Rate = Rate02) OR
(Action.Brand =SearchResult.Brand03 AND Action.Product =SearchResult.Product03 AND
Action.Position=3 AND Action.Rate = Rate03) OR
(Action.Brand =SearchResult.Brand04 AND Action.Product =SearchResult.Product04 AND
Action.Position=4 AND Action.Rate = Rate04) OR
(Action.Brand =SearchResult.Brand05 AND Action.Product =SearchResult.Product05 AND
Action.Position=5 AND Action.Rate = Rate05) OR
(Action.Brand =SearchResult.Brand06 AND Action.Product =SearchResult.Product06 AND
Action.Position=6 AND Action.Rate = Rate06) OR
(Action.Brand =SearchResult.Brand07 AND Action.Product =SearchResult.Product07 AND
Action.Position=7 AND Action.Rate = Rate07) OR
(Action.Brand =SearchResult.Brand08 AND Action.Product =SearchResult.Product08 AND
Action.Position=8 AND Action.Rate = Rate08) OR
(Action.Brand =SearchResult.Brand09 AND Action.Product =SearchResult.Product09 AND
Action.Position=9)AND Action.Rate = Rate09)

AND Brand01<>"" AND Brand02<>"" AND Brand03<>"" AND Brand04<>"" AND Brand05<>"" AND
Brand06<>"" AND Brand07<>"" AND Brand08<>"" AND Brand09<>""

AND Action.Amount = CustomerInput.Amount

AND Action.Type = CustomerInput.Type AND Action.Period = CustomerInput.Period
```

Query 4a

```
SELECT *

FROM Query2COPY

WHERE Action = "D";
```

(Query2COPY is the table that included the results from Query 3 that is displayed above)

Query 4b

```
SELECT *

FROM Query2COPY

WHERE Action = “O";
```

(Query2COPY is the table that included the results from Query 3 that is displayed above)

Query 5a

Purpose: making sure that each ActionID only appears in the table once. DetailsCopy is simply a copy of the Details table that I made to ensure that the original Details table could not be deleted.

```
DELETE *
FROM DetailsCopy
WHERE ID IN(

SELECT A.ID
FROM DetailsCopy as A, DetailsCopy as B
WHERE A.ActionID =
B.ActionID
AND A.ID > B.ID
);
```

Query 5b

Purpose: making sure that each ActionID only appears in the table once. OfferCopy is simply a copy of the Offer table that I made to ensure that the original Offer table could not be deleted.

```
DELETE *
FROM OfferCopy
WHERE ID IN(

SELECT A.ID
FROM OfferCopy as A, OfferCopy as B
WHERE A.ActionID =
B.ActionID
AND A.ID > B.ID
);
```

Query 6a

```
DELETE *
FROM DetailsCopy2
WHERE ID IN(
```

```
SELECT A.ID
FROM DetailsCopy2 as A, DetailsCopy2 as B
WHERE A.CustomerInputID = B.CustomerInputID
AND A.SearchResultID> B.SearchResultID
ORDER BY A.SearchResultID
);
```

Query 6b

```
DELETE *
FROM OfferCopy2
WHERE ID IN(

SELECT A.ID
FROM OfferCopy2 as A, OfferCopy2 as B
WHERE A.CustomerInputID = B.CustomerInputID
AND A.SearchResultID> B.SearchResultID
);
```

Query 7a and 7b

7a

```
ALTER TABLE DetailsCopy2
ADD MultipleActions int;
```

7b

```
ALTER TABLE OfferCopy2
ADD MultipleActions int;
```

Query 8a and 8b

);

8a

```
UPDATE DetailsCopy3
SET MultipleActions = 0;
```

8b

```
UPDATE OfferCopy3
SET MultipleActions = 0;
```

Query 9a

```
UPDATE DetailsCopy3 SET MultipleActions = 1
WHERE ID IN(

SELECT A.ID
FROM DetailsCopy3 as A, DetailsCopy3 as B
WHERE A.CustomerInputID =
B.CustomerInputID
AND (A.ID > B.ID OR A.ID < B.ID)
);
```

Query 9b

```
UPDATE OfferCopy3 SET MultipleActions = 1
WHERE ID IN(

SELECT A.ID
FROM OfferCopy3 as A, OfferCopy3 as B
WHERE A.CustomerInputID =
B.CustomerInputID
AND (A.ID > B.ID OR A.ID < B.ID)
);
```

Query 10a

```
SELECT *
FROM DetailsCopy3
WHERE Gross01 <> NULL AND Gross02 <> NULL AND Gross03 <> NULL AND Gross04 <> NULL AND
Gross05 <> NULL AND Gross06 <> NULL AND Gross07 <> NULL AND Gross08 <> NULL AND Gross09 <>
NULL AND Net01 <> NULL AND Net02 <> NULL AND Net03 <> NULL AND Net04 <> NULL AND Net05
<> NULL AND Net06 <> NULL AND Net07 <> NULL AND Net08 <> NULL AND Net09<>NULL;
```

Query 10b

```
SELECT *
FROM OfferCopy3
WHERE Gross01 <> NULL AND Gross02 <> NULL AND Gross03 <> NULL AND Gross04 <> NULL AND
Gross05 <> NULL AND Gross06 <> NULL AND Gross07 <> NULL AND Gross08 <> NULL AND Gross09 <>
NULL AND Net01 <> NULL AND Net02 <> NULL AND Net03 <> NULL AND Net04 <> NULL AND Net05
<> NULL AND Net06 <> NULL AND Net07 <> NULL AND Net08 <> NULL AND Net09<>NULL;
```