ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Master Thesis Marketing

An Insight into Consumer Behavior and Market Structures for the Chocolate Industry in the Netherlands and Belgium.

Name student: Diederik de Loë
Student ID number: 454606

Supervisor: Prof. Luis Eduardo Pilli
Second assessor: Prof. Nuno Almeida Camacho

Date final version: 04-08-2023

Abstract

This thesis investigates the consequences of strategic managerial decisions on the market structure in the Dutch and Belgian chocolate industry. By mapping the dynamics of the competitive landscape and filling the gaps in predicting consumer demand, the route to profit maximization can become more clear. Based on former research in the field of consumer behavior that explored choice process heterogeneity in addition to heterogeneity in consumers' tastes, this thesis has adopted an econometric specification of a choice model that incorporates both. The analysis shows that market structures, identified by consumers' latent choice sets, are influenced by marketing managers' decisions to some extent and that a deep understanding of the industry's competition and consumers is crucial to using the exogenous shock to your advantage. This demonstrates that including choice process heterogeneity in choice modeling offers more accurate market estimations and avoids the misattribution of differences in demand solely to taste heterogeneity.

Table of Contents

Introduction

We are constantly faced with hundreds of brands every day, consciously or unconsciously. Through the food we eat, the electronics we use, and the internet we surf, brands are fighting to gain our attention and make us buy their product or service. Decoding the choice process in consumers' minds and knowing how to influence it is the main challenge of marketing managers in all fields. How do I ensure consumers are aware of my product? How do I ensure consumers consider and show interest in my product? And finally, how do I ensure consumers buy my product? These are all relevant questions that are top of mind for marketing managers.

Every consumer has a unique choice process when purchasing goods or services. Consumers' minds traditionally function based on a funnel, thoughtfully removing alternatives as the funnel narrows down to the purchase decision (Court et al., 2009). Marketing managers are tasked to influence this process by identifying and influencing consumers at certain touch points, fighting for a place in the set of alternatives considered by the consumer in the vast jungle of competing brands. This jungle of competition can be very perilous as consumers are increasingly well-informed, digital channels are exploding and there are more options to choose from than ever.

More specifically, marketing managers navigate the jungle of competition and cater to consumers' different needs and wants by tailoring their marketing mix through the product, place, price, and promotion to appeal to their target consumers (McCarthy et al., 1979). Understanding the influence of strategic decision-making by marketing managers on the choice process in consumers' minds and purchase probability can be very valuable to manoeuvre a brand to a more profitable market position (Isoraite, 2016). This can be done through the consumers' unique process of choice set formation. Choice set formation is defined by Swait and Feinberg (2014) as the process of creating subsets from all the possible options of a particular product, which specifically includes alternatives that have a significant likelihood of being selected by an individual during their purchasing choice.

To gain a better understanding of the competitive landscape and overall market structure, we can map certain brands in strategic groups. Hunt (1972), in his definition, characterizes a strategic group as a collection of companies operating within the same sector that share significant similarities in terms of their cost structure, level of product diversification, formal organization, control systems, and the perceptions and preferences of individuals within the group. Understanding the distinction between strategic groups and the dynamics between them can help to get a better understanding of a brand’s positioning and strategy from a market-level perspective.

In summary, to build on the existing literature on strategic groups and choice set formation, I want to dedicate my thesis to researching the potential effects of strategic managerial decisions concerning positioning and repositioning on the market structure. To make this feasible, I will adopt an econometric specification of a choice model to estimate the probability of consumers adopting a certain choice set and alternative. The focus will lie in uncovering if heterogeneity in choice sets, in addition to taste heterogeneity, determines the purchase probability of a certain brand by an individual. How do consumers’ choice sets and tastes respond to changes made by marketing managers, and what implication does this have for market structures? These are questions that lead to the main research question:

“Can we predict what happens with market structures, identified through the probability distribution of consumers’ latent choice sets, as a consequence of marketing mix changes?”

Answering the main research question can give marketing managers a better understanding of how to influence the target consumer choice process and what this does to the competitive landscape. Navigating the marketing strategy of firms in the highly competitive, informed, and connected landscape we are currently in is more relevant than ever. This thesis will attempt to provide a map for predicting and influencing choice probabilities. This is done through an empirical approach by studying the choice probabilities of consumers on different brands with certain attributes and comparing them between two different strategic scenarios. Thus estimating choice probabilities more accurately and adding the concept of choice process heterogeneity to the already highly acknowledged existence of taste heterogeneity amongst individuals. Furthermore, the modelling of both the differences in choice sets and tastes of consumers can give valuable insight for future choice modelling and estimation of purchase probabilities.

Firstly, the existing literature on relevant topics such as consumer response heterogeneity and marketing strategy will be discussed in the literature review. This can provide us with a solid foundation to get a better understanding of what the implications of this research could be. The empirical experiment and method of analysis will be elaborated on in the data and methodology section. Next, the experiment's results will be discussed, setting up the conclusions and implications that will relate back to the main research question.

Literature review

Marketing Strategy and Strategic Groups

When we look at market structures and marketing strategies of competing firms within a certain industry we can identify various ‘strategic groups’. These groups can be segmented based on different strategic variables and implications for market structures. Firstly, we group the papers by Ferguson et al., Dranove et al., and Caves and Porter (table 1) on the strategic group implications they describe for mobility barriers and firm performance. Furthermore, we can also group the papers by Mayor et al., Meilich and Desarbo and Grewal (table 1) based on describing strategic group implications for firm mobility and the competitive landscape. Other papers such as the one by Hunt (table 1) describe other, distinctive implications.

Table 1: Overview of studies on strategic groups

Source:	Variables used to identify strategic groups:	Implications for:
-Ferguson, Deephouse & Ferguson, 2000	Reputation	
-Dranove, Peteraf & Shanley, 1998	Group characteristics (e.g. group size)	Mobility barriers, firm performance
-Caves and Porter, 1977	Key strategic dimensions	
Peteraf and Shanley, 1997	Cognitive and behavioral	Intermediate effects (e.g. efficiency, flexibility)
Hunt, 1972	Industry structures and competition	Profitability
Frazier and Howell, 1983	Marketing strategies	Organization and operation
Reger and Huff, 1993	Cognitive data	Performance and acquisition patterns
-Mayor et al., 2016	Size and economies of scale	
-Meilich, 2019	Type of distribution and price	Firm mobility and competitive landscape
-Desarbo and Grewal, 2008	Market value, efficiency, liquidity and leverage, product, size	

Among others, firms can be grouped based on their marketing strategies (Frazier and Howell, 1983) such as the marketing mix (Borden, 1964). We can group these strategic variables into continuous (price) and categorical (distribution channel) differences for instance (Meilich, 2019). As table 1 shows, grouping by different strategic variables goes hand in hand with having different implications for company aspects such as performance or mobility barriers. The fact that there is an infinite number of characteristics on which firms can be grouped, makes for a hard choice to choose those variables that are measurable and best suited to distinguish between firms and strategic groups (Cattani et al., 2017). According to Boynton and Zmud (1984), it is crucial to pick variables that can be classified as “critical success factors”. In other words, factors that differ amongst firms or groups and can be accredited for contributing to performance outcomes.

The result of these strategic groups can be both positive or negative for firms and are closely linked with how strong managers identify themselves with a certain strategic group, according to Peteraf and Shanley (1997). They also argue that strategic groups are not solely based on mutual strategies, but rather on mutual understandings and the existence of cognitive groups. This implies that firms within a strategic group understand the logic of a competitor’s decision-making and can predict each other's prediction function (Edwards, 1991; Tirole, 1989). This mutual understanding of the decision-making process is linked with the central characteristics of the strategic group. These groups can for instance be based upon similar characteristics such as firm size, pricing and product quality, or can be based more on similar relationships such as overlapping social networks or common corporate history. Especially the latter can be open to one’s interpretation.

This means that the distinction between strategic groups is often unclear. Reger and Huff (1993) describe this as the result of core strategic firms with distinctive characteristics, secondary firms with characteristics that match various strategic groups and solitary firms with unique characteristics that stand alone. More distinctive characteristics will go hand in hand with increased mobility barriers which makes it harder for outsiders to enter the specific strategic group (Caves and Porter, 1977). However, over time, the shifting of groups and firms is inevitable. Due to trends and certain managerial decisions, firms may change strategic groups and strategic groups may merge, split, change positioning or even disappear altogether. Consequently, due to the interdependence of firms and groups in the same industry, the competitive landscape will change as well. Mayor et al. (2016) found that firm mobility is higher between strategic groups of closer resemblance, therefore having lower mobility barriers.

To illustrate the dynamic effects of firms and strategic groups, Meilich (2019) introduces a method of constructing a strategic groups map that visualises the competitive landscape in an industry. He also describes an example of how the strategic group of booksellers with a high number of physical stores and low online presence shifted to a combination of both and created a new, so-called “bricks and clicks’’, strategic group in the late 90s. This shift had a serious effect on the industry’s competition, namely on the online-only pioneers in bookselling. This shows how certain trends and company decisions can drastically change the competitive landscape in an industry over time.

We can also look at competition from a market-level standpoint. These are considered competitive market structures and include various strategic groups (Desarbo and Grewal, 2008). Past research discusses the fluctuation of these structures and the existence of pure and hybrid strategic groups in addition to core, secondary, and solitary firms (Reger and Huff, 1993). Pure strategic groups consist of firms with a clear mutual strategy whereas hybrid strategic groups consist of firms that borrow strategies from multiple pure strategic groups. This implies that market structures are not fixed and that firms can belong to multiple strategic groups and therefore not only compete within one strategic group.

In conclusion, the term ‘strategic groups’ was first mentioned by Hunt in 1972 and has since had many definitions. However, this thesis defines them as groups that consist of firms that have similar strategies (Ferguson, Deephouse, and Ferguson, 2000; Peteraf and Shanley, 1997) and compete more intensely with each other (Dranove, Peteraf, and Shanley, 1998). We can identify strategic groups based on various firm characteristics and corresponding outcomes, ranging from tangible ones like firm size and price to intangible ones like strategy and product quality. The positioning of firms within these groups does not have to be fixed, firms shift between strategic groups and new strategic groups are formed while outdated ones disappear over time. For this reason, we distinguish core, secondary, solitary and hybrid firms. Core firms strictly belong to one pure strategic group that experiences strong competition within the group and follows the same strategic recipe, secondary firms adhere to the same strategies to a lesser extent, solitary firms follow a unique strategy and form their own strategic group and hybrid firms blend the strategies of multiple strategic groups.

Consumer Response Heterogeneity

Taste heterogeneity

Product heterogeneity and the drive to differentiate from the competition stems from the consumers' urge to seek variety across consumption occasions and the difference in taste or preference among consumers (Lancaster, 1990). This heterogeneity in taste is defined by Price, Feick and Higie (1989) as the extent to which an individual's taste or preference may vary across the population. Thus, consumers vary on their most preferred mix of attributes or on the amount of each attribute they wish to see. In other words, consumers give different weights of importance to each attribute, resulting in a unique product composition that yields the highest personal utility.

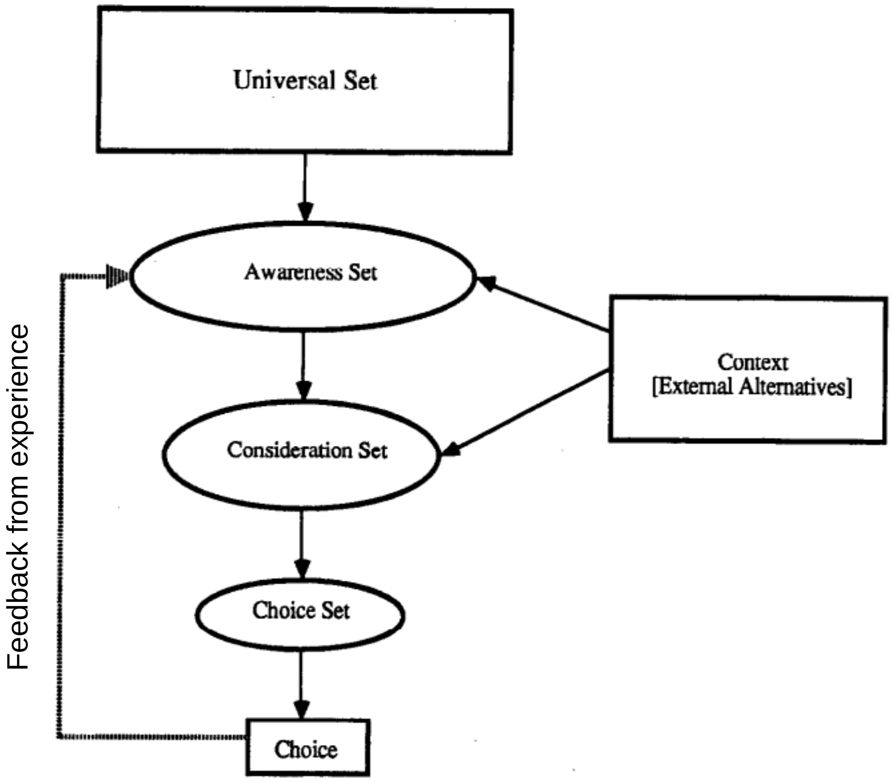
Therefore, goods and services that encounter high taste heterogeneity among consumers have more opportunities to differentiate and offer unique benefits. For example, the clothing industry is subject to strong taste heterogeneity and therefore a lot of clothing styles exist to cater to benefits ranging from functionality to trendiness. Whereas low taste heterogeneity logically means that consumers have similar attribute weightings and seek similar benefits (Feick and Higie, 1992). An example of a product with low taste heterogeneity is a paperclip for instance, because most people will only evaluate a paperclip on its ability to hold together multiple papers. This understanding of consumer taste heterogeneity is the foundation of differences in marketing strategies and the existence of strategic groups and market structures (Kamakura et al., 1996).

Furthermore, various authors have researched the correlation between socio-demographic characteristics and taste. Herve and Mullet (2009) have looked at the correlation between age and preferences toward certain aspects of clothing. They found that young people have a favourable taste towards low price, middle-aged people towards suitability and elderly people towards durability when it comes to buying clothes. All of these correlations were significant and this implies that socio-demographics can play an important role in the determination of taste heterogeneity.

Choice set heterogeneity

However, out of the vast selection of product alternatives, consumers tend to eliminate a significant amount before making their choice of preference. Creating a group that is smaller than the number of brands available (universal set) and smaller than the number of brands the consumer is aware of (awareness or knowledge set). This is the result of a complex psychological choice-making process (Weber and Johnson, 2009) and is described as unobserved heterogeneity in subjective choice sets (Hauser, 1978). This means that individual consumers have a way of forming a set of alternatives in their minds to choose from when purchasing a certain product or service. Successfully identifying the set of alternatives with a non-zero probability of choice explains around 80% of the variance in probabilistic choice models (Hauser, 1978).

Figure 1: Flowchart of individual choice (Shocker et al., 1991)



So in addition to traditional taste heterogeneity and consumer choice-making under utility maximisation of different product attributes (Rieskamp et al., 2006), there is a preceding psychological process that encompasses choice process heterogeneity (Swait and Feinberg, 2014). This is where subsets are made out of all the available alternatives of a specific product and it solely contains alternatives that have a non-zero probability of being chosen during the purchase decision of a certain individual. For instance, an individual with a gluten allergy will eliminate all products containing gluten when forming a specific food choice set. This process of nesting is illustrated in figure 1. It defines a framework to understand the choice-making process in a consumer's mind, from all the available alternatives in the universal set to the eventual choice.

Although there is conflicting literature on the distinction between consideration sets and choice sets, we make a distinction based on the following definitions (Shocker et al. 1991). The consideration set consists of all the alternatives that satisfy the goal of the consumer and are accessible to the consumer on a particular occasion. The consumer purposefully measures out each alternative it is aware of based on the attributes and the specific context, forming his or her consideration set. This means that consideration sets can vary among and within use cases and are dynamic by nature. On the other hand, the subsequent choice set has a more static nature. Most importantly, the choice set is the unobservable set of alternatives evaluated during a specific choice occasion immediately before the purchase decision. In this thesis, the latter (the latent choice set) will be adopted as the focal construct.

Neglecting choice process heterogeneity would result in misattributing differences in demand solely to taste heterogeneity, thus making inaccurate estimation of choice probabilities (Pilli, Swait and Mazzon, 2022). The importance of certain product attributes will be misjudged and brands will struggle to produce the optimal amount, eventually withholding them from profit maximisation. The concept of choice set formation is discussed by Hauser and Wernerfelt (1990). They describe it as a trade-off between the costs and benefits of a certain brand. Does the benefit of including a certain brand in my choice set outweigh the cost of gathering information about this brand in a specific consumption occasion? Moreover, a thorough understanding of choice process heterogeneity will enable marketing managers to influence the choice-set formation process of target consumers through strategic and tactical marketing activities (Shapiro et al., 1997). For instance, using billboards to communicate the sublime quality of your product or service could decrease the cost of gathering information about your brand for the consumer. This way, marketing managers can increase the chance their product is included in the choice set and increase the purchase probability.

In addition to marketing activities, choice sets can also be influenced by the goal a product serves (Shocker et al., 1991), through context (Kardes et al., 1993) and through situational variables (Chakravarti and Janiszewski, 2003). For the first matter, this means that through changing attributes of a product, it can be included in choice sets of consumers seeking a certain goal that this attribute serves. The second matter implies that the competitive environment plays a role in choice set formation. More specifically, pioneering products are more likely to be included in choice sets as they have an increased awareness due to an initial (short-term) monopoly. Additionally, the formation of choice sets is more likely in complex market structures with a lot of product offerings than it is in more simple structures with fewer offerings. With fewer products, assessing all available brands is more feasible and choice sets are less intuitive. Lastly, depending on the comparability of attributes amongst products, products with more aligned attributes have a higher likelihood of being included in the choice set. Meaning that more distinctive products are more likely to be left out as it is harder to compare to the other alternatives.

To summarise, we have discussed the heterogeneity amongst taste and choice sets of consumers during purchase decisions. Based on the amount of taste heterogeneity towards a certain good or service, marketing managers will choose a suitable strategy of benefit segmentation. High taste heterogeneity implies stronger differentiation and low taste heterogeneity implies more uniformity. In addition to taste heterogeneity, we can also identify a preceding step of choice set formation. Choice set formation is also subject to heterogeneity and the result of a complex psychological process in the consumer’s mind. However, a thorough understanding of these processes is crucial for the success of marketing managers and can be influenced through strategic marketing decisions.

Hypotheses

As we have learned from the literature review, consumers are subject to heterogeneity in choice sets and tastes during purchase decisions. In addition, marketing managers can influence consumer response through various strategic decisions. Changing product attributes as part of the marketing mix can be one of these decisions. It is to be expected that changing product attributes will have an impact on the choice sets of consumers and therefore affect the choice probability of certain brands. This is why we identify the following first hypothesis:

H1: Changing product attributes as a result of marketing managers’ strategic decisions affects the choice probability of the brands through shifts in the choice set probabilities.

Furthermore, as stated earlier, we can identify strategic groups based on strategic variables such as the marketing mix. Strategic groups with alternatives that have similar marketing mix attributes are more likely to be part of the same choice set of a consumer. This is because it makes for easier comparison and more varying alternatives will have been eliminated in the preceding consideration set as mentioned in the literature review. As a result of an exogenous shock, it can therefore be expected that when brands within a strategic group or choice set move away from another strategic group in terms of product attributes, the choice probability of the now more distinctive choice set will increase. This leads to the second hypothesis:

H2: The choice probability of choice sets that become more distinctive as a result of the exogenous shock will increase.

We have identified consumer response heterogeneity to consist of heterogeneity in choice sets and tastes. After choice set formation, consumers tend to choose from the remaining alternatives based on their taste. This is unique to each individual. An individual's taste can be related to certain socio-demographic characteristics such as gender, age and educational background. Although the main focus of this research is the latent choice set, taste heterogeneity is acknowledged and this research will also attempt to identify age as a moderator of an individual's taste toward stronger-tasting chocolate brands with 70% cocoa. This is based on earlier research done by Mojet, Christ-Heizelhof and Heidema (2001) that correlates an increase in age to a decreased sensitivity to strong tastes. Therefore the third hypothesis is as follows:

H3: The choice probability for chocolate with 70% cocoa will increase as the age of consumers increases, indicating a change in taste correlated with age.

In other words, this research examines how changing product attributes, through the simulation of an exogenous shock induced by marketing managers' strategic decisions, affect brand choice probabilities through consumer response heterogeneity (H1). In addition, strategic groups, represented as choice set heterogeneity, play a role in how brands are affected by the shock (H2). Lastly, the research identifies the socio-demographic boundary condition age, under which the effect may be positively or negatively correlated (H3).

Data & Methodology

Survey

The survey was held amongst 600 respondents living in the Netherlands or Belgium. Moreover, it is based on a discrete choice experiment and defines the different chocolate brands based on 5 attributes The chocolate brands have been classified as low-, middle- or high-tier as displayed in table 2. These tiers correspond with strategic groups, as we will rely on the product aspect of the marketing mix to identify strategic groups in the empirical approach.

Table 2: Differences between control and experimental condition

	Control condition		Experimental condition	
Brand	**Quality rating**	**Price**	**Quality rating**	**Price**
Lindt	High-tier 3-5 star	€1,75-€2,65	High-tier 3-5 star	€1,75-€2,65
Godiva				
Tony's Chocolonely	Middle-tier 2-4 star	€1,35-€2,05	Middle-tier 1-3 star	€0,95-€1,45
Nestlé L'Atelier				
Chateau	Low-tier 0-2 star	€0,55-€0,85	Low-tier 0-2 star	€0,55-€0,85
Côte d'Or				

Besides quality, the chocolate brands were also defined by price per 100g (ranging from €0,55 to €2,65) and dependent on the corresponding tier, type (70% cocoa and milk) and addition (salted caramel, almonds and pure).

Half of the survey respondents were presented with an experimental condition (table 2) which was the result of an exogenous shock induced by marketing managers' strategic decisions. More specifically, there was a simulation of certain marketing managers changing product attributes for their chocolate brand. This can give insight into how consumers change their choice sets as a consequence of the shock and what implications this has for long-term market structures. Now, middle-tier brands were attributed with lower quality ratings (2-3) and lower prices eliminating the overlap with high-tier brands and positioning themselves as a more distinctive strategic group. One-half of the respondents were therefore faced with the middle-tier brands skewed towards the high-tier and one-half of the respondents with the middle-tier brands skewed more towards the lower end to create a between-subjects design.

Furthermore, all respondents were faced with several socio-demographical questions. These consisted of questions about gender (figure 2), ethnicity (figure 3), employment status (figure 4), student status (figure 5) and age (figure 6).

Figure 2: Gender

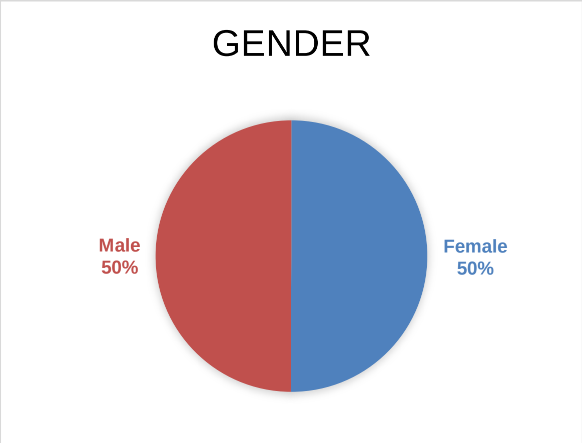


Figure 3: Ethnicity

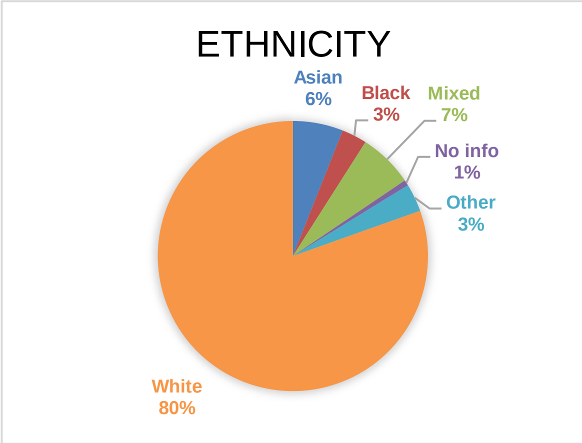


Figure 4: Employment status

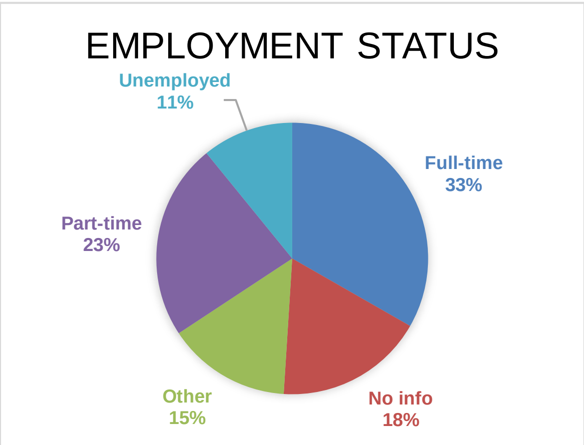


Figure 5: Student status

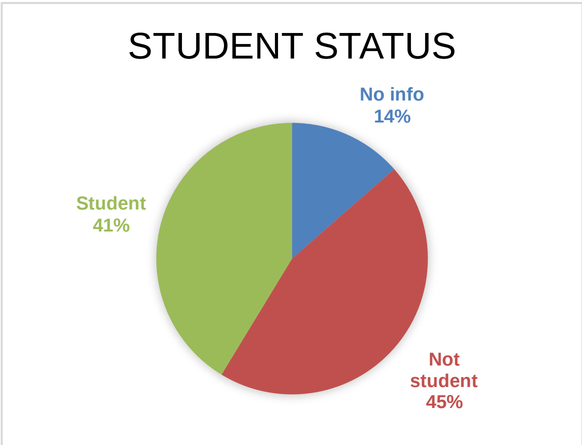
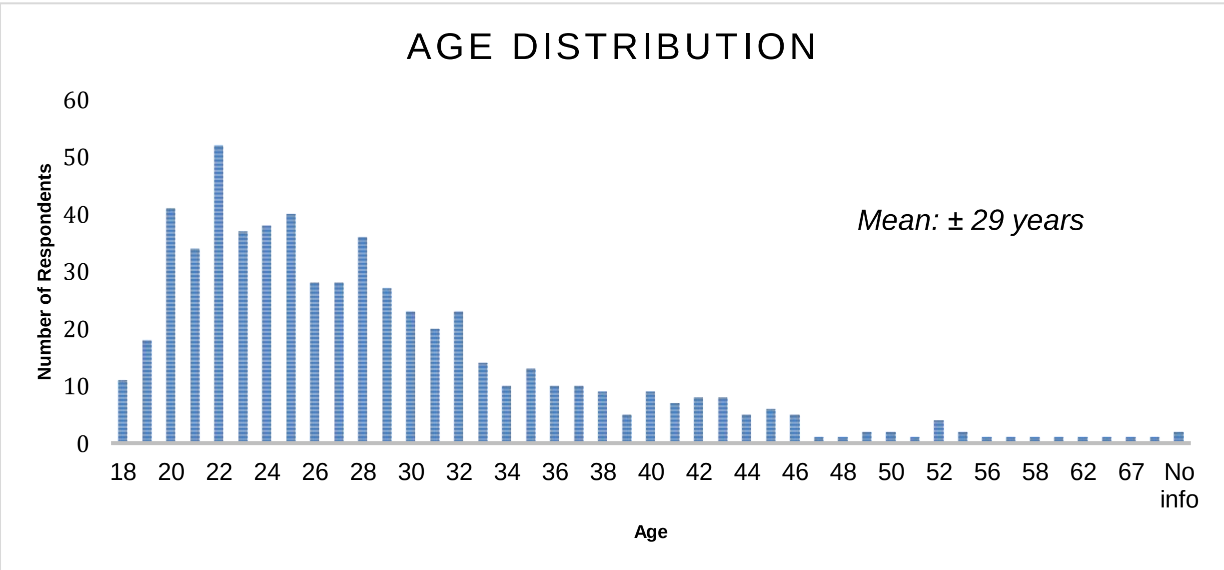


Figure 6: Age distribution



Conceptual model

The structure that will determine the choice model estimation is illustrated in the path diagrams below (figure 7 and figure 8). As mentioned before in the literature review, distinctive strategic group characteristics make for stronger mobility barriers. However, as marketing managers make strategic decisions, the competitive landscape and strategic group membership can change. We can explain this through the effect on choice set probabilities and the market structure. This concept is at the root of the conceptual model, where we will simulate two scenarios as the consequence of an exogenous shock induced by marketing managers. The 'before' scenario has the middle-tier brands skewed towards the high-tier brands which results in the middle-high-tier. Whereas the 'after' scenario has middle-tier brands skewed more towards the lower segment which results in more distinctive groups. You can see this difference in figure 7 and 8 through the market structures and the distance between choice sets, highlighting the dynamic nature of competition.

In both cases, there is a need to estimate the choice set probabilities in the upper part of the figure and the brand choice probabilities in the lower part of the figure, given the corresponding choice set probability. This is where the GenL model comes in. The upper model of the GenL probability function will calculate the choice probabilities of the upper paths that lead to the latent choice set and incorporate choice set heterogeneity. Hereafter, the lower model of the GenL probability function will calculate the choice probabilities of the lower paths that lead to the brand, given the latent choice set. This will incorporate taste heterogeneity. The multiplication of these two choice probabilities gives us the joint probability and forms the GenL choice probability function (Swait, 2001).

As the experiment is probabilistic by nature, all possible choice sets will be taken into account when calculating their choice probabilities in the control and experimental conditions. This means that there will be a calculation for the 'low', 'low-middle', 'middle', 'middle-high', 'low-high' and 'all' (containing all alternatives) choice set. In reality, including every possible combination would be impossible as there are a lot more alternatives than just the six this experiment looks at. Some form of constraining would therefore be necessary. The codes that will run the GenL model have been provided by my supervisor.

Figure 7: Path diagram of strategic groups and brands before exogenous shock

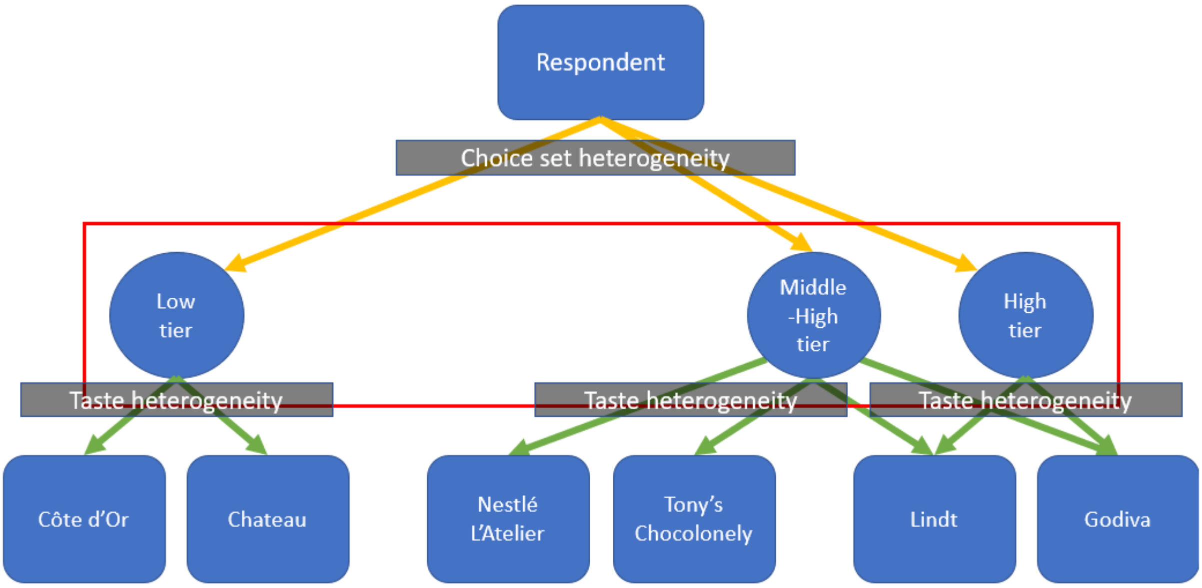


Figure 8: Path diagram of strategic groups and brands after exogenous shock

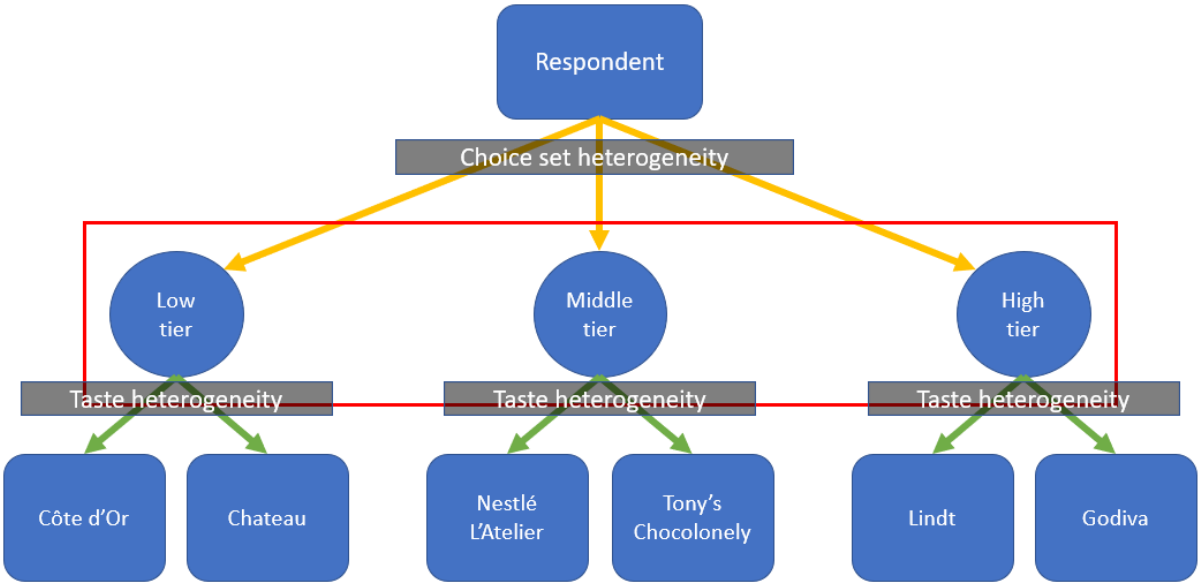


Table 3: Symbol interpretations

Symbol	Interpret
	Choice set probability
	Brand probability
	Market structure

Choice model

To analyse the results, choice modelling will be used. Discrete choice modelling in particular, is a statistical technique used to understand and predict individuals' tastes and choices among a set of alternatives, in this case, chocolate brands (Train, 2009). It provides insights into how different factors influence choice-making and allows researchers to quantify the importance of various attributes in the choice-making process.

Multinomial logit model

One model that will be used within choice modelling is the multinomial logit (MNL) model. The MNL model assumes that individuals make choices by comparing the utilities associated with each alternative and selecting the alternative with the highest utility (Train, 2009). It assumes that the utility of each alternative is a linear function of the attributes and their corresponding coefficients. The MNL model is useful for understanding overall preferences and estimating the relative importance of different attributes on choice probabilities.

The MNL model consists of the following utility function and logit probability function (McFadden, 1974):

The full utility function is given by:

$$U = V_{nj} + \varepsilon$$

In this expression:

- V_{nj} is the systematic component of the full utility of person ‘n’ for alternative ‘j’.
- ε is the random component of the full utility of person ‘n’ for alternative ‘j’.

The systematic component of the utility function can be calculated through the following expression:

$$V_{nj} = \beta X_{nj}$$

In this expression:

- V_{nj} represents the utility of person ‘n’ for alternative ‘j’.

- X_{nj} is a vector of observed attributes relating to alternative 'j' for person 'n'.

- β is the beta that gives weight to the observed attribute X_{nj}, notice that it is the same for all individuals.

Logit probability function:

$$P_{ni} = \frac{exp(\mu_k V_i)}{\sum_{j \epsilon C_k} exp(\mu_k V_j)}$$

In this expression:

- P_{ni} represents the probability of person 'n' choosing alternative 'i'.

- V_i is the utility function of alternative 'i' for person 'n'.

- V_j is the utility function of all alternatives 'j' for person 'n'.

- μ_k is the choice set specific parameter that encompasses the unobserved variance in the error-term of the utility function. In the case of the MNL model, the μ_k is unidentifiable and therefore set to 1 with no effect.

- The sum Σ is taken over all alternatives 'j' that are partitioned into 'C_k' choice sets.

The MNL model will lie at the foundation of this research. However, the MNL model assumes homogeneity in both choice sets and taste (Train, 2009), which urges the need for a more elaborate model to account for our main theme of heterogeneity.

The generalized extreme value family

This takes us to the generalized extreme value (GEV) family of discrete choice models. It is a class of models used in choice modeling to analyze decision-making processes. These models are designed to capture a variety of substitution patterns among alternatives and allow for correlations over alternatives, providing a more flexible framework for modeling choice behavior (Train, 2009). The GEV models are an extension of the standard logit model, which assumes independence from irrelevant alternatives (IIA) and proportional substitution across alternatives.

GEV models also have the advantage of usually having closed-form choice probabilities, which allows for estimation without resorting to simulation (Train, 2009). This characteristic makes GEV models a valuable tool for researchers, as they provide new and powerful specifications to meet their needs. This simplifies the estimation process and provides several benefits:

1. More efficient and faster estimation compared to simulation-based methods.
2. Better facilitation of result interpretation.
3. Allows for comparison between different scenarios and evaluates the potential effects on choice outcomes.

Overall, the closed-form nature of choice probabilities in GEV models enhances their usability and applicability in research.

Choice set generation logit

One specific member of the GEV family is the GenL (choice set generation logit) model. The GenL model incorporates choice set generation directly into the GEV framework. It accounts for preferences in the choice set generation process and, unlike traditional models of choice set formation, the GenL model specifies choice set generation endogenously, meaning it also reflects individual preferences. In other words, the GenL model will allow consumers' tastes to be part of the choice set generation component. The properties and characteristics of the GenL model have been examined in detail by Swait (2001).

The GenL model consists of two stages: choice set generation and choice (Swait, 2001). In the choice set generation stage, the model generates the set of alternatives from which the individual will make a choice. The choice set is determined based on the individual's preferences and characteristics, which will be derived from the socio-demographic survey questions. In the choice stage, the individual selects one alternative from the generated choice set.

In essence, the probability function in the GenL model is a multiplication of two MNL probability functions. The GenL probability function can therefore be expressed as follows:

$$P_{ni} = \frac{exp\left(\frac{Logsum(C_k)}{\mu_k}\right)}{\sum_{k=1}^{K} exp\left(\frac{Logsum(C_k)}{\mu_k}\right)} \times \frac{exp(\mu_k V_{i,k})}{\sum_{j \in C_k} exp(\mu_k V_{j,k})}$$

In this expression:

- P_{ni} represents the probability of person ‘n’ choosing alternative ‘i’.
- $Logsum(C_k)$ represents the log of the sum of the exponentiated utilities of all alternatives within the choice set C_k.
- μ_k is the choice set specific parameter that encompasses the unobserved variance in the error-term of the utility function. In contrast to the MNL model, the GenL model allows for a value other than 1.
- $V_{i,k}$ is the utility function of alternative ‘i’ in the choice set ‘k’ for individual ‘n’.
- $V_{j,k}$ is the utility function of all alternatives ‘j’ in the choice set ‘k’ for individual ‘n’.
- The sum Σ is taken over:
    - The exponentiated logsum utilities of all the choice sets C_k divided by the choice set specific μ_k for individual ‘n’.
    - The exponentiated utilities of all the alternatives ‘j’ in the choice set ‘k’ multiplied by the choice set specific μ_k for individual ‘n’.

The GenL probability function can be referred to as being made up of an upper model and a lower model. The upper model refers to the first part of the function where the probability of adopting a certain choice set is calculated. The upper model will account for the main focus of this research, which is heterogeneity in choice sets. This is done by identifying choice set specific utilities and a choice set specific µ which is related to the variance in the utility error-term. The experimental condition will be added in the upper model by adding an interaction effect with the logsum term using the experimental condition binary variable.

The second part of the GenL probability function is referred to as the lower model and gives the probability of choosing a specific alternative, given the choice set. The lower model will account for heterogeneity in taste through interactions between the systematic utility function and socio-demographic characteristics, in this case, the correlation between age and 70% cocoa. The effect of the experimental condition will also be accounted for through an interaction in the utility function. Although choice set heterogeneity is the main theme of this research, taste heterogeneity is included to acknowledge its importance.

In summary, taste heterogeneity has been widely discussed in academic literature and can be accounted for through an interaction effect in the MNL utility function (Train, 2009). To reflect market structures and introduce choice set heterogeneity, we look at the family of GEV models, specifically the GenL model (Swait, 2001). It offers a versatile approach for modeling discrete choice situations. These models incorporate choice set generation modeling directly into the specification and allow for a range of substitution patterns and correlations among alternatives, providing a more comprehensive understanding of individual decision-making processes in various fields.

It's important to note that the GenL model combines choice set generation and choice decisions within a unified framework, allowing for a more comprehensive understanding of the decision-making process. The model estimation and interpretation can be illustrated using empirical data, as demonstrated in the case study on intercity mode choice by non-business travellers presented in the paper by Swait (2001).

Results

As described in the methodology part, the results have been obtained through an MNL and GenL choice model. The results section is structured accordingly and will relate to the formulated hypotheses, starting by taking a look at several statistical measures for comparative evaluation.

Goodness of fit

Firstly, we will look at the descriptive statistics of the MNL model and compare these to the GenL model. Thereby ensuring a good model fit. More specifically we will look at the Akaike information criteria (AIC, AIC3) and the Bayesian information criterion (BIC) which can be found in table 4.

The output gives us a BIC value of 23241.576 (table 4) for the MNL model. If we compare this to the BIC value of 23230.68 (table 4) of the GenL model, it becomes clear that the use of the more complex GenL model to incorporate heterogeneity makes sense as the value is slightly lower indicating a lower likelihood of information loss. It is important to note that the AIC and BIC incorporate a penalty to discourage overfitting and ending up with a model that is too complex. This is because otherwise, the act of adding extra parameters to the model would always result in better goodness of fit. Lastly, the BIC is favoured over the AIC as it practices a more strict version of this penalty.

Table 4: Goodness of fit of the MNL and GenL models

Criteria	MNL Value	GenL Value
AIC	23166.885	23112.053
AIC3	23183.885	23139.053
BIC	23241.576	23230.68

The MNL model parameters

The MNL model parameters were estimated and are presented in table 5. The parameters are defined beside the values to indicate their role in the corresponding probability function. Furthermore, the standard error and p-values are also added to indicate the parameter significance. A p-value threshold of 0.05 has been adopted.

Every tier was given a specific binary variable and its value was estimated to show the role of the tier a brand is in on the choice probability. Looking at the p-values (table 5) of the ASC parameters, it becomes clear that only the high-tier chocolate brands see a significant effect on utility and thus the choice probability, namely an increase of 0.424 (table 5). In addition, an ASC parameter for the mid-tier in the experimental condition was added. This shows us the effect on the choice probability of mid-tier brands in the experimental condition. What we see, is that mid-tier brands see a significant decrease of -0.301 in utility in the experimental condition (table 5). Whereas in the control condition, mid-tier brands have an insignificant and smaller negative effect on utility (table 5).

Within tiers, we see that Côte d'Or is favoured over Chateau, Tony's is favoured over Nestlé and Lindt is favoured over Godiva. All these brands have a positive effect on utility compared to the other alternatives in the same tier and these effects are significant when looking at the p-values (table 5).

When we look at the flavours, it is important to note that the milk flavour was omitted. This means that, compared to milk, 70% cocoa has an overall negative significant effect on utility (table 5). We can also relate to the third hypothesis which predicted a correlation between age and the effect of chocolate with 70% cocoa on one's utility. If we look at the interaction parameter for 70% cocoa and age, we can indeed see that an increase in age is correlated with a marginal increase of 0.009 in utility per year of observed age (table 5). This effect is highly significant.

Next to flavours we also looked at the effect of several additions to chocolate bars on choice probability. Here, the pure variation was omitted from the model. Thus, we can say that the almond variation has a negative effect on utility compared to pure, whereas the caramel sea salt variation sees an increase in utility compared to pure. Both effects are significant (table 5).

Lastly, we see that quality has a positive effect- and price has a negative effect on utility as they increase. Utility will always increase as the quality rating increases from low to high and it will always decrease as the price rating goes up from low to high. More importantly, when considering the quality and price scales (table 2), the increase in utility of higher quality alternatives outweighs the decrease in utility of the higher price. These effects are also significant (table 5).

Table 5: MNL model parameters (rounded to three decimal places)

MNL MODEL

NAME	Meaning	Value	Standard Error	p-value
ASC_LOW	Low-tier specific constant (binary)	-0.304	0.212	0.166
COTEXCHATEAU	0.5 if Cote, -0.5 if Chateau	0.654	0.038	0.000
ASC_MID	Mid-tier specific constant (binary)	-0.120	0.168	0.476
ASC_MID X EXP	Mid-tier specific constant in exp. condition (binary)	-0.301	0.054	0.000
NESTLEXTONY	0.5 if Nestle, -0.5 if Tony's	-0.808	0.035	0.000
ASC_HIGH	High-tier specific constant (binary)	0.424	0.169	0.021
GODIVAXLINDT	0.5 if Godiva, -0.5 if Lindt	-0.176	0.034	0.000
70% COCOA	Flavour	-0.457	0.024	0.000
70% COCOA X AGE	Interaction between 70% Cocoa and age	0.009	0.001	0.000
ALMONDS	Addition	-0.118	0.019	0.000
CARAMEL SEA SALT	Addition	0.066	0.018	0.002
QUALITY LOW	Low star rating	0.435	0.028	0.000
QUALITY MID	Middle star rating	0.602	0.026	0.000
QUALITY HIGH	High star rating	0.475	0.026	0.000
PRICE LOW	Low price	-1.421	0.317	0.000
PRICE MID	Middle price	-1.147	0.087	0.000
PRICE HIGH	High price	-1.139	0.051	0.000

The GenL model parameters

A first look at the GenL model results immediately reveals that this is a more complex model, as there are more parameters and we now introduce the choice set specific μ (table 6). We have already mentioned that the higher complexity of this model pays off as the BIC value of the GenL model is lower than the one of the MNL model.

Starting by looking at the brands and their effect on the utility, we take in mind that the Côte d'Or brand is omitted. This tells us that all brands save Nestlé, which are positioned in a higher tier than Côte d'Or, yield an increasingly higher utility (table 6). The reason for Nestlé being favoured less than Côte d'Or whilst being in a higher tier will be discussed later on.

Within Côte d'Or's low-tier group, it is the more favourable brand over Chateau utility-wise (table 6). Looking at the p-values however, we can disregard the effects of the Tony's, Godiva and Lindt brands as their p-values are higher than 0.05 (table 6).

For the flavours, we see that the GenL model estimates similar values to the MNL model. Again, 70% of cocoa shows a negative effect on utility compared to milk (table 6). The GenL model also shows a positive correlation between 70% cocoa and age with a higher value compared to the MNL model. This is in line with the third hypothesis. Both effects are significant (table 6).

Regarding the additions, the effect of adding almonds or caramel sea salt instead of having a pure chocolate bar is also the same as the MNL model. Almonds will yield a lower utility and adding caramel sea salt to chocolate will result in a higher utility (table 6). Both effects are significant again (table 6).

Finally, the quality and price attributes also show the same effect in both the MNL and GenL models that can be expected when increasing the quality or the price of a product. A higher price will yield an increasingly lower utility and higher quality will yield an increasingly higher utility (table 6).

Now for the parameters that are newly introduced in the GenL model and that have attempted to account for heterogeneity in choice sets and its effect on the choice probability of a certain brand. Firstly, the μ values seem to not drive the model as their effect fades in relation to the corresponding standard errors (table 6). Therefore, they are not different from 1 and the GenL function collapses into an MNL function without identifying choice set heterogeneity (Train, 2009).

Furthermore, the high-tier is the only tier that sees a significant effect on its choice probability as a result of the exogenous shock (table 6). Moreover, this effect is higher than just the choice set effect and can signal that the exogenous shock is driving the model and that it is capable of showing choice set heterogeneity as opposed to the endogenous variables alone.

In addition, we can identify a ceiling effect induced by the high-tier logsum interaction with the exogenous shock. More specifically, the increase in utility of the high-tier caused by the exogenous shock is weakened by the decrease in the high-tier logsum as a result of the exogenous shock. This is displayed by the negative value of -0.550 with a significant p-value, causing a ceiling effect for the high-tier in the experimental condition (table 6).

Table 6: GenL model parameters (rounded to three decimal places)

GENL model

NAME	Meaning	Value	Standard Error	p-value
CHATEAU	Brand	-0.395	0.101	0.001
NESTLE	Brand	-0.550	0.238	0.028
TONY'S	Brand	0.095	0.196	0.618
GODIVA	Brand	0.218	0.204	0.290
LINDT	Brand	0.371	0.207	0.082
70% COCOA	Flavour	-0.385	0.088	0.000
70% COCOA X AGE	Interaction between 70% Cocoa and age	0.076	0.019	0.001
ALMONDS	Addition	-0.091	0.024	0.001
CARAMEL SEA SALT	Addition	0.055	0.018	0.004
QUALITY LOW	Low star rating	0.304	0.074	0.000
QUALITY MID	Middle star rating	0.509	0.106	0.000
QUALITY HIGH	High star rating	0.428	0.092	0.000
PRICE LOW	Low price	-0.888	0.317	0.009
PRICE MID	Middle price	-0.900	0.200	0.000
PRICE HIGH	High price	-1.031	0.218	0.000
LOW (μ)	Choice set specific μ	5.898	3.340	0.087
MID (μ)	Choice set specific μ	1.997	0.815	0.021
HIGH (μ)	Choice set specific μ	1.510	0.379	0.000
LOW MID (μ)	Choice set specific μ	1.390	1.110	0.217
MID HIGH (μ)	Choice set specific μ	1.237	0.868	0.162
LOW HIGH (μ)	Choice set specific μ	1.101	0.469	0.026
LOW_TIER EXP_COND	Effect of exp. condition on low-tier	0.452	0.412	0.277
MID_TIER EXP_COND	Effect of exp. condition on mid-tier	0.264	0.493	0.585
HIGH_TIER EXP_COND	Effect of exp. condition on high-tier	1.510	0.283	0.000
LOW_TIER(LOGSUM) X EXP_COND	Effect of exp. condition on logsum	1.118	0.777	0.158
MID_TIER(LOGSUM) X EXP_COND	Effect of exp. condition on logsum	0.074	0.400	0.838
HIGH_TIER(LOGSUM) X EXP_COND	Effect of exp. condition on logsum	-0.550	0.192	0.008

GenL model choice set logsums

As mentioned in the methodology section, the choice set logsums are closely related to the choice set probabilities and are involved in its calculation. In short, the logsum is a measure of the utility of a choice set and is linked to the utility of the alternatives within that same choice set (Train, 2009). The logsums are presented in table 7. The 'all brands' choice set containing all brands has the highest logsum, followed by the low/high-tier logsum (table 7). This illustrates that a high logsum is linked with a higher choice set probability.

Furthermore, we see that the logsums of all the tiers save the high-tier decrease in the experimental condition compared to the control condition. Thus, within the experimental condition, the ratio between logsums changes in favour of the high-tier (table 7).

Table 7: GenL model means of choice set logsum (rounded to three decimal places)

	LOW-TIER	MID-TIER	HIGH-TIER	LOW/MID-TIER	MID/HIGH-TIER	LOW/HIGH-TIER	ALL BRANDS
CONTROL	0.000	0.613	0.402	1.378	1.536	1.684	1.847
EXPERIMENT	-0.001	0.577	0.402	1.365	1.519	1.673	1.829

GenL model choice set probabilities

As discussed in the methodology part, the GenL model allows for the estimation of choice set probabilities in addition to brand probabilities. This is done by the aforementioned upper model and calculated with the use of the choice set logsums (table 7). The choice set probabilities are shown in table 8. Looking at these results allows us to test the second hypothesis which states that choice sets that become more distinctive as a result of the exogenous shock see their choice probability increase. In this case, that would be the low-, mid- and high-tier.

We can see that as a result of the exogenous shock, the low-, mid- and high-tier choice set probabilities all increase (table 8). This is at the cost of the choice set probabilities for the joint tiers and the tier containing all brands. The high-tier benefits most from the exogenous shock, as it has the highest increase in choice set probability from 0.068 to 0.188 (table 8).

Relating to the second hypothesis, we can say that the increase in choice probabilities for the low-, mid- and high-tier support it. This is because the exogenous shock causes the overlap between the mid- and high-tier to disappear and it makes the low-, mid- and high-tiers to be more distinctive. This was aforementioned and is visualized in figure 7 and 8.

Table 8: GenL model choice set probabilities (rounded to three decimal places)

	LOW-TIER	MID-TIER	HIGH-TIER	LOW/MID-TIER	MID/HIGH-TIER	LOW/HIGH-TIER	ALL BRANDS
CONTROL	0.044	0.082	0.068	0.163	0.190	0.217	0.264
EXPERIMENT	0.064	0.091	0.188	0.134	0.155	0.178	0.217

Brand probabilities in the MNL model

The code has also estimated the choice probabilities of all six chocolate brands in the MNL model (table 9). The table makes a distinction between the control condition and the experimental condition which gives insight into the effect of the exogenous shock on the favourability towards certain brands. This can give us an indication of the first hypothesis, which states that by changing certain product attributes, the choice probability of brands will change through a shift in the choice set probabilities.

By looking at the control and experimental columns of table 9, we can compare the choice probabilities between control and experimental conditions. Table 9 shows that low-tier brands Côte d'Or and Chateau see their choice probabilities rise as a result of the exogenous shock. The same counts for high-tier brands Godiva and Lindt, with a similar ratio (table 9).

On the other hand, middle-tier brands Nestlé and Tony's have seen their choice probability decrease as a result of the exogenous shock induced by marketing managers decreasing the price and quality. These effects are also reflected in the ratios of brand probabilities between the experimental and control conditions. Thus indicating that choice probabilities for brands do indeed change as a result of changing product attributes.

Table 9: Brand probabilities for the MNL model (rounded to three decimal places)

MNL MODEL BRAND PROBABILITIES

NAME	Ratio	Control	Experimental
CÔTE D'OR	1.092	0.152	0.166
CHATEAU	1.093	0.086	0.094
NESTLÉ	0.829	0.146	0.121
TONY'S	0.871	0.287	0.250
GODIVA	1.099	0.152	0.167
LINDT	1.110	0.182	0.202

Brand probabilities in the GenL model

Similarly to the MNL model, the GenL model has also estimated the choice probabilities for all six brands in the control and experimental conditions. These probabilities are displayed in table 10.

These results support the change in brand probabilities and choice set probabilities caused by the exogenous shock as formulated in the first and second hypotheses. Namely because we see an increase in the choice set probability of the high-tier (table 8) and a sequential increase in the high-tier brand probabilities (table 10).

Some effects are different from those of the MNL model, however. This is revealed by the fact that the MNL model imposes fixed ratios for the low- and high-tier brand probabilities, whereas the GenL model softens this constraint by including the effect of choice sets (table 9 and 10). In the GenL model, the increase in choice probabilities for the high-tier brands Godiva and Lindt is significantly higher than the increase we see in the MNL model (table 9 and 10). This is at the cost of the low-tier brands Côte d'Or and Chateau, their choice probabilities decrease in the GenL model in contrast to the MNL model.

Table 10: Brand probabilities for the GenL model (rounded to three decimal places)

GENL MODEL BRAND PROBABILITIES

NAME	Ratio	Control	Experimental
CÔTE D'OR	0.918	0.159	0.146
CHATEAU	0.897	0.097	0.087
NESTLÉ	0.840	0.144	0.121
TONY'S	0.895	0.285	0.255
GODIVA	1.205	0.146	0.176
LINDT	1.237	0.173	0.214

Conclusion

This thesis has attempted to answer the following research question: *"Can we predict what happens with market structures, identified through the probability distribution of consumers' latent choice sets, as a consequence of marketing mix changes?".* This was done by using discrete choice modelling, namely the MNL model and the GenL model. Within these models, an exogenous shock was simulated in an attempt to identify choice set heterogeneity in addition to the acknowledgement of taste heterogeneity. The chocolate industry was adopted as a testing ground and the estimation was based on a survey amongst 600 respondents living in the Netherlands and Belgium.

The first hypothesis was formulated as follows: *"Changing product attributes as a result of marketing managers' strategic decisions affects the choice probability of the brands through shifts in the choice set probabilities".* From the results, we can indeed conclude that the choice probabilities of the middle-tier brands, whose quality and price attributes were lowered, decreased. Furthermore, we can conclude that in the consumers' minds, the loss of quality is weighed more heavily than the decrease in price for chocolate brands. This is revealed by the significant decrease in the logsum of the middle-tier brands as a result of the exogenous shock. However, the biggest effect of the changing attributes was seen in the high-tier choice set probability and the corresponding high-tier brand probabilities.

This sets up the second hypothesis which states the following: *"The choice probability of choice sets that become more distinctive as a result of the exogenous shock will increase".* The GenL model allowed for the estimation of choice probability for choice sets and from these results we can conclude that the second hypothesis holds. Because the middle-tier brands saw their quality and price attributes decrease, they moved away from the high-tier brands and took a more distinctive middle-segment positioning. This resulted in an increase in choice probabilities for the low-, mid- and high-tier choice sets and a decrease in the choice sets that included brands from multiple tiers. The high-tier choice set especially benefitted from its more distinctive upmarket position.

Lastly, the third hypothesis looked at taste heterogeneity and it stated the following: *"The choice probability for chocolate with 70% cocoa will increase as the age of consumers increases, indicating a change in taste correlated with age".* By adding an interaction parameter between the socio-demographic age and the 70% cocoa flavour in the systematic utility functions of both the MNL and the GenL model, the third hypothesis could be tested.

Both models showed a significant correlation. The older the respondents, the more preference they show towards chocolate with 70% cocoa.

Considering all the pertinent results and returning to the main research question-*"Can we predict what happens with market structures, identified through the probability distribution of consumers' latent choice sets, as a consequence of marketing mix changes?"*-it can be concluded that market structures are indeed influenced to some extent by marketing mix changes implemented by marketing managers. In this particular case, the middle-tier brands' repositioning had a notable impact primarily on the high-tier brands, Godiva and Lindt, which benefited from the middle-tier brands' shift towards lower quality and price. However, the identification of choice set heterogeneity in addition to taste heterogeneity was not possible through the endogenous variables alone; the exogenous shock was needed to do this. Moreover, the selection of chocolate brands and their tier allocation can be called into question, particularly concerning Côte d'Or and Nestlé. These aspects will be further explored in the discussion, along with the implications and recommendations of this thesis to the field of consumer behaviour.

Discussion

Implications

Based on the conclusion, there are several insightful takeaways for marketing managers. First, choosing your brand's positioning and adapting it wisely based on the behaviour of your competitors is crucial for success, especially for competitors that are positioned in the same strategic group. The act of changing your company's marketing mix has winners and losers. We observed that high-tier brands benefited from middle-tier brands changing their product attributes and shifting to a more low-end market position. In contrast, in the control condition, middle-tier brands were positioned relatively close to the high-tier brands. However, after changing their positioning, they became further away from high-tier brands with lower quality and price. The price decrease was not able to offset the effect of the decrease in quality, leading to a reduced preference for middle-tier brands.

Preference for low-tier brands also saw a small reduction due to the middle-tier brands positioning themselves closer to them and therefore facing more competition. This is in line with what the GenL model brand probabilities showed us. In contrast, the MNL model estimated higher choice probabilities for the low-tier brands in the experimental condition. Based on intuition, it is more logical that both the low- and middle-tier brands share in the loss of choice probability as a consequence of the middle-tier moving downmarket, due to increased competition.

The high-tier brands can be considered the winners as they are perceived more clearly as having high quality and face less competition. Consumers that were formerly hesitant between middle- and high-tier brands, will now be more in favour of high-tier brands as higher quality outweighs lower price in the consumers' minds. On the other hand, more price-sensitive consumers on the lower end of the market will now also consider middle-tier brands next to low-tier brands. This can explain the decrease in choice probability for low-tier brands in the GenL model. Overall, this implies that marketing managers should be very careful when changing certain attributes of their marketing mix, as it can have an insignificant or even adverse effect on themselves and change the competitive landscape to the benefit of other brands.

If we specifically look at choice sets, we can see that as a result of the middle-tier brands adopting a more mid-segment position in the experimental condition, the probability of adopting a choice set with brands from the same strategic group increases. This means that when brands choose a more distinctive positioning, they have a higher likelihood of competing more intensely with brands in the same strategic group in the consumers' minds. On the other hand, if strategic group boundaries are more unclear, consumers tend to include brands from different strategic groups in their choice sets more frequently. The GenL choice set and brand probabilities support this finding.

Adopting the GenL model and including the latent choice set in the estimation of purchase probability is therefore a good idea. This is to ensure that the effect of an exogenous shock on the latent choice set and therefore market structure is included. This benefit was highlighted in the difference in the brand choice probabilities between the MNL model and the GenL model. The GenL model was able to incorporate the effect of the exogenous shock on choice sets in the estimation process, whereas the MNL model simply looked at the change in utility that was yielded from the new marketing mix. This shows that the GenL model incorporates a significant benefit in the estimation of purchase probability and demand forecasting.

Moreover, marketing managers should be aware of the scope of their competition when choosing a certain marketing strategy. In a market with less clear strategic group boundaries, chances are that they are competing for a place in their target consumers' choice set with more brands than they think. Understanding what benefits consumers are seeking and communicating the value proposition accordingly can increase the chance of being included in choice sets and eventually lead to profit maximization.

Hence, marketing managers need to acknowledge heterogeneity in choice sets and tastes when predicting demand for their products. Understanding the effect of marketing mix changes on demand through consumer choice sets and tastes is crucial for navigating toward a more profitable market position. Carefully studying the competition and reacting in the best way depends on understanding the actions of the competition and their impact on the market structure and the choice sets of consumers. In this case, we observed that high-tier brands benefited from middle-tier brands competing less intensely with them. Not following the same strategic decision by also lowering quality and price proved to be beneficial for the high-tier brands. This could be explained by the fact that they are perceived to have a stronger high-end positioning by consumers and are more likely to be considered in choice sets of consumers seeking luxury chocolate.

Limitations and recommendations

The foundation of this thesis lies in choice modelling, where survey results from 600 respondents living in the Netherlands and Belgium were used to estimate the results. Choosing the right models within choice modelling has proven to be quite a challenging task. Contemplating the complexity of the model and adding extra parameters to increase accuracy requires careful consideration and can be open to one's interpretation. The MNL model and GenL model with a certain number of parameters have ultimately been chosen in consultation with my supervisor and have yielded useful results. Although the MNL and GenL models have been widely applied, the addition of the latent choice set in calculating consumer choice probabilities, alongside consumer taste, is relatively new in the field of consumer behaviour and can contribute to more accurate choice estimation.

On one hand, this thesis has confirmed that taste heterogeneity plays a significant role in an individual's preference toward certain products. In this case, we found that age can influence an individual's taste preference for chocolate with 70% cocoa flavour. On the other hand, the essence of this thesis is to emphasize the inclusion of choice set heterogeneity in predicting consumer purchase behaviour and enabling profit maximization. With the methodology applied in this thesis, the endogenous variables were unable to identify choice set heterogeneity. The addition of an exogenous shock was needed.

However, by increasing the number of respondents and brands considered in the survey, choice set heterogeneity could become more apparent. Moreover, to increase significance and accurately test the correlation between age and taste, a better age distribution among the respondents would be preferable. Currently, the mean age of all respondents is 29 years, so including relatively older people could better resemble real society.

Finally, a thorough study of the respondents' domestic market is crucial as well. The choice of brands and their positioning in the Dutch and Belgian chocolate industry was based on a relatively brief online study of their pricing strategies. This led to Côte d'Or being placed in the low-tier, directly competing with Chateau. Nonetheless, the respondents perceived Côte d'Or to be a more high-end brand than Chateau and even Nestlé, distorting the results. For future research, engaging in discussions with industry marketing leaders and consumers before conducting the survey can result in a better representation of the real market and yield more accurate results.

References

Borden, N. H. (1964). *The concept of the marketing mix*. Journal of advertising research, 4(2), 2-7.

Boynton, A. C., & Zmud, R. W. (1984*). An assessment of critical success factors.* Sloan management review, 25(4), 17-27.

Caves, R. E., & Porter, M. E. (1977). *From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition.* The quarterly journal of economics, 91(2), 241-261.

Cattani, G., Porac, J. F., & Thomas, H. (2017). *Categories and competition.* Strategic Management Journal, 38(1), 64-92.

Chakravarti, A., Janiszewski, C. (2003). *The influence of macro-level motives on consideration set composition in novel purchase situations.* J. Consum. Res. 30, 244–258.

Court, D., Elzinga, D., Mulder, S. & Vetvik O. (2009). *The consumer decision journey.* McKinsey & Company. Retrieved from https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/the-consumer-decision-journey

DeSarbo, W. S., & Grewal, R. (2008). *Hybrid strategic groups.* Strategic Management Journal, 29(3), 293-317.

Dranove, D., Peteraf, M., & Shanley, M. (1998). *Do strategic groups exist? An economic framework for analysis*. Strategic Management Journal, 19(11), 1029-1044.

Edwards, D. (1991). *Categories are for talking: On the cognitive and discursive bases of categorization.* Theory & psychology, 1(4), 515-542.

Ferguson, T. D., Deephouse, D. L., & Ferguson, W. L. (2000). *Do strategic groups differ in reputation?*. Strategic management journal, 21(12), 1195-1214.

Feick, L., & Higie, R. A. (1992). *The effects of preference heterogeneity and source characteristics on ad processing and judgements about endorsers.* Journal of Advertising, 21(2), 9-24.

Frazier, G. L., & Howell, R. D. (1983). *Business definition and performance*. Journal of Marketing, 47(2), 59-67.

Hauser, J.R. (1978). *Testing the accuracy, usefulness, and significance of probabilistic choice models: an information-theoretic approach.* Oper. Res. 26, 406–421.

Hauser, J. R., & Wernerfelt, B. (1990). *An evaluation cost model of consideration sets.* Journal of consumer research, 16(4), 393-408.

Hervé, C., & Mullet, E. (2009). *Age and factors influencing consumer behaviour*. International journal of consumer studies, 33(3), 302-308.

Hess, S., & Palma, D. (2019*). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application.* Journal of choice modelling, 32, 100170.

Hess, S., & Palma, D. (2019). *Apollo version 0.1.0.* user manual. www.ApolloChoideModelling.com

Hunt, M. S. (1972). *Competition in the major home appliance industry, 1960-1970.* Harvard University.

Išoraitė, M. (2016). Marketing mix theoretical aspects. *International Journal of Research-Granthaalayah*, *4*(6), 25-37.

Kamakura, W.A., Kim, B.-D., Lee, J. (1996). *Modeling preference and structural heterogeneity in consumer choice.* Mar. Sci. 15, 152–172.

Kardes, F.R., Kalyanaram, G., Chandrashekaran, M., Dornoff, R.J. (1993). *Brand retrieval, consideration set composition, consumer choice, and the pioneering advantage.* J. Consum. Res. 20, 62.

Lancaster, K. (1990). *The economics of product variety: a survey.* Mar. Sci. 9, 189–206.

Magidson, J., & Vermunt, J. K. (2004). *Latent class models.* The Sage handbook of quantitative methodology for the social sciences, 175-198.

Mayor, M. G. O., Davó, N. B., & Martínez, F. R. (2016). *Dynamic strategic groups' analysis and identification of mobility barriers in the European mobile phone industry.* International Journal of Technology Management, 71(3-4), 253-277.

McCarthy, E. J., Shapiro, S. J., & Perreault, W. D. (1979). *Basic marketing* (pp. 29-33). Georgetown, ON, Canada: Irwin-Dorsey.

McFadden, D. (1974). Analysis of qualitative choice behavior. *Frontiers in econometrics*, *1*, 105-142.

Meilich, O. (2019). *Strategic groups maps: review, synthesis, and guidelines.* Journal of Strategy and Management, 12(4), 447-463.

Mojet, J., Christ-Hazelhof, E., & Heidema, J. (2001). *Taste perception with age: generic or specific losses in threshold sensitivity to the five basic tastes?*. Chemical senses, 26(7), 845-860.

Peteraf, M., & Shanley, M. (1997). *Getting to know you: A theory of strategic group identity.* Strategic Management Journal, 18(S1), 165-186.

Pilli, L., Swait, J., & Mazzon, J. A. (2022*). Jeopardizing brand profitability by misattributing process heterogeneity to preference heterogeneity.* Journal of choice modelling, 43, 100359.

Price, L. L., Feick, L. F., & Higie, R. A. (1989). *Preference heterogeneity and coorientation as determinants of perceived informational influence.* Journal of Business Research, 19(3), 227-242.

Reger, R. K., & Huff, A. S. (1993). *Strategic groups: A cognitive perspective. Strategic management journal*, 14(2), 103-123.

Rieskamp, J., Busemeyer, J.R., Mellers, B.A. (2006). *Extending the bounds of rationality: evidence and theories of preferential choice.* J. Econ. Lit. 631–661.

Shapiro, S., MacInnis, D.J., Heckler, S.E. (1997). *The effects of incidental Ad exposure on the formation of consideration sets.* J. Consum. Res. 24, 94–104.

Shocker, A.D., Ben-Akiva, M., Boccara, B., Nedungadi, P. (1991). *Consideration set influences on consumer decision-making and choice: issues, models, and suggestions.* Market. Lett. 2, 181–197.

Swait, J. (2001). Choice set generation within the generalized extreme value family of discrete choice models. *Transportation Research Part B: Methodological*, *35*(7), 643-666.

Swait, J., Feinberg, F.M. (2014). *Deciding how to decide: an agenda for multi-stage choice modelling research in marketing.* In: Hess, S., Daly, A. (Eds.), Handbook of Choice Modelling. Edward Elgar Publishing, pp. 649–660.

Tirole, J. (1988). *The theory of industrial organization.* MIT press.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge university press.

Weber, E.U., Johnson, E.J. (2009). *Mindful judgement and decision making.* Annu. Rev. Psychol. 60, 53–85.