ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis: Data Science and Marketing Analytics

Adaptive Kano Choice-Based Conjoint Analysis (AK-CBC): Integrating the Kano Model in the design process of the Choice-Based Conjoint Analysis. A study on the market of voice-assisted smart speakers

---

Author: Oscar Flores Sanchez (Student ID: 501967)

Supervisor: Prof. dr. A. (Andreas) Alfons

Co-reader: Prof. dr. N.M. (Nuno) Almeida Camacho

December 7, 2019

Abstract

In this data-led business environment in which we life today, identifying what the customer wants before your competitor does can be the thin line that separates failure from success. Conjoint Analysis and especially its most popular type, Choice-based Conjoint (CBC), is a widely popular tool in the world of Market Research that allows to identify, through statistical modelling, which are the most attractive products and services on the basis of respondents' trade-offs among their atomic features. In the last years, there have been numerous studies seeking to maximize both the statistical performance and the user engagement of these techniques, mainly through the creation of experimental designs that allow to enhance the information that could be inferred from each respondent by presenting better suited choice alternatives. This would not only enable to obtain more reliable data, but the managerial decisions that could be drawn from it would be far more accurate, valuable and relevant. To this end, various models have been developed with the goal of optimizing the product alternatives presented through the use of preliminary screening questions. This work seeks to go one step further and become an innovative contribution in the field of Machine Learning Model Deployment with the introduction of the Adaptive Kano Choice-Based Conjoint (AK-CBC). This novel model uses the well-known Kano Model to remove irrelevant features and a preliminary CBC to filter out non-appealing levels to fine-tune the design of a CBC exercise on the fly.

Key words: Individual Preferences, Conjoint Analysis, Kano Model, Design Generation, Adaptiveness

JEL: C11, C13, C35, C9, M31, 03

Acknowledgements

First of all, I would like to thank both Andreas Alfons and Nuno Almeida for taking the time to go through this work and provide their experienced assessment. They are, without a shadow of a doubt, a crucial element in the achievement of this personal milestone.

I am profoundly grateful to SKIM and specifically some SKIMmers that have made this thesis possible. I would like to personally thank Nicolas, my supervisor at SKIM, not only he provided the idea for it but his extreme competency and his guidance during the whole process has been vital for making this work a reality. I am sure his immeasurable skills will serve SKIM and himself well to achieve great things in the future, and I am also sure that this has been a great learning experience for both of us. On a related note I cannot stress enough how valuable Jesse's technical input has been to this work. His expertise and warmth made me thankful to connect with him not only on a professional but also on a personal level. Additionally, I would like to thank the quick-witted Jayme, as his technical resources improved drastically this work. Marco, whose extensive experience and knowledge in Conjoint added robustness from the methodological perspective and Marcel, who has been the cornerstone of the whole project and whose trust has made this work possible.

On a more personal note, I would like to thank my family who have been a constant support during the whole journey and to all friends with whom I have shared good moments but also the struggle during these past months; Camilla, Filip, Ricarda and Isabel to name a few but many more that have always been there.

Contents

1 Introduction

1.1 Context and motivation

In 2018, we assisted to a landmark event. For the first time in history, the smartphone market sales came to a halt on a full year basis with a global decline in sales of approximately 4%. According to experts this drop is due to longer replacement time spans (phones show higher durability), saturation in major Asiatic and Western markets, and lack of what is known as wow models. While companies are struggling to come up with the new leaps of innovation that might reignite growth again, such as 5G, foldable displays, punch-hole cameras or eye sensors to lure consumers to upgrade their models, the question arises: *What is the next game-changer in the sector of smart devices?* On this behalf, I attended a presentation of an omni-channel company specialized in technology and communication processes which was currently focusing its efforts on analysing the potential future scenarios for world interconnectivity. On one side, they reaffirmed the idea that the industry of smart phones was in a mature stage and starting to decline. On the other hand, they strongly believed in the potential of success of voice-controlled devices such as smart speakers or home-assistants.

As a second motivating factor, I am currently working in a company (SKIM) that operates in the sector of market research studies through Conjoint Analysis and its different varieties. By focusing on capturing customer preferences and modelling choice-behaviour, we specialize in projects ranging from new product development, pricing strategies or portfolio optimization among others. Another standard tool that we use for those purposes is the Kano Model, which is used to classify the different features or attributes (will be used interchangeably in this work) that shape a new product concept into different categories by its level of integration and its capacity to improve customer satisfaction. A more elaborated explanation of these two methods (Conjoint Analysis and Kano Model) is provided in the next section “Theoretical framework”, but in a nutshell, the target of this dissertation is to integrate them into a hybrid model that would be ultimately used for designing a new product concept. This idea is underpinned by the fact that both techniques have components that could be combined and which might complement each other.

1.2 Research Questions

Henceforth, this work introduces a novel methodology in which the results obtained from an initial Kano exercise can be used for tailoring and fine-tuning the product concepts shown in the choice tasks on the fly, which means that for each individual respondent we will be able to adapt the alternatives presented in the computer-administered Conjoint exercise on the basis of the responses that he/she provided earlier in the survey. Besides, the hybrid model will be used and tested through the development of a new smart speaker concept. The goal of the thesis is threefold then; on one hand I would like to assess its performance for capturing customers' needs and preferences against a random-generated Choice-based Conjoint exercise. For that, the following Research Question was formulated:

RQ1: What would be the predictive power, in terms of identifying customer preferences, for the AK-CBC as compared to a random-generated Choice-Based Conjoint study?

Secondly, but no less importantly, one of the main goals of the newly-designed model is to adapt the Conjoint experiment to the particular and individual needs and preferences of each respondent. With this, the intention is to increase respondent's engagement in the choice task to ultimately obtain more deliberated and judicious responses. Subsequently, the following Research Question has been formulated as well:

RQ2: How would the new model perform in regards to user engagement as compared to a random-generated Choice-Based Conjoint exercise?

Finally, and as we will see in the next section, the Kano Model classifies the various atomic features that shape a product or service into different descriptive categories. By initially categorizing these attributes and using the non-irrelevant ones for designing a CBC, we would be able to assess the relationship between each category and their estimated utility part-worth. Considering this, the final Research Question was defined:

RQ3: What is the relationship between each Kano category with their corresponding part-worth utility?

1.3 Structure

The structure of the ensuing study will be organised as follows. Firstly, a thorough presentation of the theoretical framework, founded on relevant literature review and technical reports, and with the principal goal of uncovering and exposing potential gaps or points that could be improved in traditional Conjoint Analysis. In this section, the main methods that will shape the conceptual groundwork of the study (including the Kano Model), their main strengths and weaknesses along with the different hypotheses will be introduced. Furthermore, a section to explain the methods used for estimating utilities as well as the key notion of study design will be comprehensively discussed.

Next, the core of the dissertation will be presented, as the new model scheme (AK-CBC), its main components and a brief discussion of how the gaps described before could be filled will be dissected in detail.

Following it, the data and the methodology that would be used to analyze it will be described, including the data collection process, the data cleaning procedures and a thorough analysis of the main sociodemographic patterns. Additionally, the list of attributes and levels that were included in the study will be introduced.

In the subsequent sections, the results of the data-analysis will be laid out. A response to the different Research Questions and hypotheses will be delivered in the main findings and discussion part as well as potential actionable managerial implications.

Finally the most important conclusions will be drawn and summarized, followed by limitations of the study, and potential considerations regarding future research lines on the topic.

2 Theoretical Framework

Technology is playing a key role in how businesses operate nowadays. The steady switch from brick and mortar stores to a globalized and interconnected online shopping environment with giant multi-channel players like Amazon or Aliexpress has radically transformed the way in which shopping is done. On these days, from the palm of a hand and just by having access to a working internet connection, any customer can access a myriad of products and services to choose from. This factor has stressed the importance of understanding how the mind of the customer works. Without a shadow of a doubt, those companies that could better comprehend what it is that the customer truly and deeply desires, and manages to deliver a product or service that most closely resembles those needs, is going to have a clear head start for succeeding.

From a neuromarketing point of view, customers show their positive emotional response, and ultimately their preference towards a particular good or service, through choices. On a daily basis we make thousands of these choices, most of them have a trivial nature and are not instrumental to the day to day activities, such as choosing what to eat or wear. Others are far more intricate, as the relevant and multifaceted nature of the alternatives to be chosen, make the decision-making process involved more complex. In this kind of decisions, like for example choosing a Master Program to enrol in, numerous aspects and factors need to be taken into account for making a final choice so as to maximize the gratification level. When we talk about products or services such as smart phones or data plans, we assume these aspects to be the various features that configurate the offering integrally. As a case in point, when we want to purchase a new smartphone, we measure its attractive by assessing and comparing the different atomic features that define it, such as its hardware like screen size or colour, software specifications like RAM capacity, or intangible aspects like its brand prestige. Considering all these, the customer makes trade-offs among the different attributes (Safizadeh, 1989) and the configuration that most closely suits his needs and demands would have the highest chances of being chosen.

Conjoint Analysis and the Kano Model are two well-known tools, extensively used in the market research industry to uncover customer needs and requirements. Both techniques, despite having substantially different foundations, mutually aim at providing a direct elicitation of customer's preferences towards a set of a product or service's atomic features. In terms of existing literature, the main developments in the field of Conjoint Analysis that could be of direct application for this work, will be thoroughly described in the coming subsections. On that behalf, I will introduce the most relevant methodologies that are being used at the moment in the sector, with a special focus on their most important advantages and specially their limitations, which is the main reason that has triggered the design of this new model.

On the other hand, in regards to the existing literature on the Kano Model, there has been extensive academic research when it comes to using it as the main pillar for a new product development scheme (von Dran et al, 1999), for enhancing customer satisfaction (Chen & Chuang, 2008) or as an integrative component to an already existing method, such as combinedly integrating it with the Quality function deployment method (QFD) for optimizing product design processes (Tan & Shen, 2000; Tan & Patrawa, 2001; Tontini, 2007; Lee et al, 2008) or with the failure mode and effect analysis (FMEA) for anticipating potential failures and mitigating risk in the early stages of developing a new product concept (Shahin, 2004).

In terms of combining the Kano Model with Conjoint exercises, and before delving into the detailed theoretical notions of each tool, it is pertinent to mention that both in the academia and industry sphere it is fairly common to observe that the results of the Kano questionnaire can be used ad-hoc to identify which are the most relevant attributes in order to leverage them in the Conjoint Analysis on a complementary basis. On that behalf, a study by Min et al. (2011) used the Kano exercise to reveal which were the most significant e-book reader features in the Korean Market in order to use those in a Conjoint experiment at a later stage. Similarly, Suzianti et al. (2015) employed an identical approach to identify which were the most important service attributes in Indonesian fashion online shops for ultimately calculating preferences statistically using Conjoint. On a more related note, Wang & Wu (2014) proposed a sophisticated theoretical framework based on the use of Conjoint Analysis to determine which core smart phone features were most relevant in order to split respondents into homogeneous segments, and the Kano to identify which optional attributes were preferred for each of these clusters. By using a multi-criteria decision making method called VIKOR, the researcher could assess and rank which smart phone configurations (combining core and optional features) would maximize overall customer satisfaction for each of the segments. Continuing this line of research this works introduces a contribution to the existing academic literature with the presentation of the AK-CBC, which analogously uses the Kano questionnaire and the utilities from an initial Conjoint exercise to build on the fly and on a respondent level a tailor-made Conjoint design that adapts to each respondents' needs and preferences. Overcoming previous technical limitations, the model takes a step forward and deploys intro production a generalizable methodology that will be tested through the development of a new smart speaker concept.

2.1 Conjoint Analysis

Conjoint Analysis (CA), introduced by Green & Rao (1971), is a quantitative market research technique which intends to provide a direct elicitation of preference by measuring the customers´ trade-offs for a multi-attribute item. McFadden (1973), laid the cornerstone to numerically compute discrete preferences based on probabilities, by developing the method Multinomial Logistic Regression as a way of inferring individual part-worth utilities in a choice task. These part-worth utilities can be understood as the importance weight that each individual allocates to each of the feature levels that form a product.

For this section, apart from suitable papers from the most relevant journals of statistics and marketing research, I would use the broad repository of technical papers from the world's leading software provider in the area of choice and Conjoint Analysis (Sawtooth Software) as a core information source. From these technical papers, we can infer that this survey-based statistical technique assumes that an individual´s affinity for a given alternative can be computed by the sum of the part-worths of each of the integrating attribute levels. Henceforth, by numerically assessing these individual part-worths we can ultimately appraise the preference for a given alternative configuration.

2.1.1 Choice-Based Conjoint Analysis (CBC)

Within the realm of Conjoint Analysis (CA) there are different mechanisms to assess or approximate these preferences, such as ranking or rating the different product concepts (Baier, Pełka, Rybicka & Schreiber, 2015). However, Choice-Based Conjoint Analysis (CBC) is far more popular than the other subcategories given that the process is designed to realistically mimic a real market environment in which consumers can choose among various competing products (Ben-Akiva, Mcfadden & Train, 2019). On that same line, Asioli et al. (2016) argued that the data-gathering for a CBC study closely matches real life situations as individuals reveal their preferences through choices, but found no clear empirical evidence on its superiority as compared to the Rating-based Conjoint analysis. However in the study it was acknowledged that CBC resulted more suitable for those consumers with middling preferences.

What is clear according to the latest reports from Sawtooth Software (Orme, 2013) and my own experience working in a market research agency, is that CBC has a predominant position in the sector as compared to the other types (around 70-75% of all Conjoint studies are choice-based). Furthermore, the commercial software provides user-friendly solutions that allow the user to set flexible and well-balanced CBC experimental designs allowing to make direct predictions of realistic choice behaviour, as we will see in the upcoming sections.

For illustrative purposes, we can see in Table 1 how a typical CBC task would look like (with multiple tasks comprised in a given study), in which a respondent would need to choose among the following hypothetical laptop alternatives (in this work the terms product alternative/concept/profile/configuration will be used interchangeably). On this behalf, four concepts would be presented and the following 6 attributes and its respective levels would be included in the study; 1. Price (499$, 699$, 899$, 1399$), 2. Brand (HP, Apple, Lenovo, Dell), 3. RAM (4GB, 8GB, 16GB), 4. Touch Screen (Yes/No), 5. CPU (i3, i5, i7) and 6. Storage (1TB, 256 GB SSD, 512 GB SSD).

Table 1: Example of a CBC task

	Concept 1	**Concept 2**	**Concept 3**	**Concept 4**	
Brand	lenovo	hp	DELL		**None: I would not choose any of these alternatives**
RAM	4GB	8GB	8GB	16GB	
Touch screen	No	No	Yes	No	
CPU	i3	i5	i5	i7	
Storage	1TB	256 GB SSD	256 GB SSD	512 GB SSD	
Price	499$	699$	899$	1399$	

Study Design

At this point, it would be indispensable to introduce the key concept of design of a choice study, as it plays a pivotal role in this dissertation. In the field of Conjoint Analysis, a study design or experiment basically represents the structure of the different experimentally controlled tasks that are shown across all respondents, namely how each hypothetical product profile is defined and the number of tasks, concepts per task, and items/attributes per concept that comprise the exercise. Since it would be unfeasible and cost-prohibitive to enquire the respondent with all possible attribute level combinations or product profiles (known as full factorial experiment), it is inevitable to construct a subset of those that faithfully mimics the reality of the market while at the same time is efficient on the basis of some well-accepted criteria.

There are various methods of creating Choice-based designs, but basically they can be split among manual and randomized designs. Manual ones are normally used for complex studies which requires a great deal of industry expertise as it normally requires to generate unusual configurations originating from the full-factorial design. On the other hand, Randomized designs, used as default by Lighthouse (Sawtooth Software's online survey tool), allow to generate efficient designs that, depending on the settings and the number of attributes and levels, comply with the principles of Level balance (each attribute level appears a similar number of times), Minimal overlap (the number of times a level is shown in a task is minimized) and Orthogonality (the frequency for each pair of levels is proportional across all pair of attributes). On that behalf, randomized designs also allow a certain level of fine-tuning and customization from the researchers' side, as we will see in the following paragraphs.

As a preliminary stage, it is paramount that the designer makes sure that the attributes included are independent to each other because incurring in certain level of overlap may results in an over-inferred influence on product choice for those particular features. Also, it is important to note that the number of levels comprised in each attribute have a significant bearing on the results due to the well-studied **Number-of-levels effect** (Huber, Zandan, Johnson & Wittink, 1992), which basically refers to the well-studied phenomenon in which from both a psychological and probabilistic prism, the attributes with a higher number of levels achieve larger average importance scores. The general rule-of-thumb is to minimize the difference between number of levels across the various attributes included, but at the same time trying to mimic reality as closely as possible.

Another aspect that needs to be taken into consideration for defining the study design specifications and for ensuring that the combinations displayed during the interview are as realistic as possible and in harmony to the current market situation, is to set prohibitions or prohibiting pairs to guarantee that conflicting or incompatible levels from different attributes do not appear together. One of the most typical prohibitions is to disengage explicitly or naturally superior attribute levels with low-tiered price ranges and vice versa. On that line, alternative-specific combinations might be taken into consideration to render further flexible designs, as some or all product alternatives might have their own unique sets of attributes i.e. conditional pricing where each brand might have its own set of prices. Last but not least, another procedure to further improve and balance the design is to create weights and thus determine upfront the probabilities of appearance for each level to mimic market reality to an additional extent.

Preference estimation

As aforementioned, a preference for a given concept might be retrieved by the sum of the utility values or part-worths of each of the attribute levels that shape it. In Choice-based Conjoints, the approximation of this part-worths could be analogous to the counts, namely how many times a product alternative including an attribute level was chosen relative to the number of times it was available for choice. Numerically, these part-worth utilities can be estimated using different variations of the Multinomial Logit Model. The Bayesian approach, for example, allows to compute utilities at an individual level, but it is considerably more computationally-intensive, and yields similar results in terms of reliability as compared to the Classical or Maximum Likelihood approach, and due to those differences both methods were used in this study in different situations or contexts as we will document at a later stage. A more detailed explanation of the models and how the utilities for each level are estimated can be found in section 2.1.4. On this behalf, one important distinction is to differentiate between main effects and interaction effects computations. The former could be understood as the standalone relative importance of each of the levels, whilst the latter indicates the variation in utilities of multiple levels when they are combined as opposed to the sum of the individual parts alone. To shed light on this last notion, Meißner and Steiner (2018) came up with a clever example. The part-worth of a car being red could be for instance 0.2 (Attribute: colour, level: red), and for being a Ferrari (Attribute: brand, level: Ferrari) might be 0.4, but the joint utility of a red Ferrari (which is the color that clearly differentiates the company) could be, lets say, 0.9.

Drawbacks

In spite of the multiple advantages described before, there are several drawbacks in the traditional CBC that may flaw the results of the choice experiment up to a certain degree. On one hand, the concepts that are presented in the choice task are normally not very close to the respondent´s ideal, given that all attribute levels are forced to be included in the different alternative combinations (if no prohibitions are set), when only a few of them are seriously taken into consideration and ultimately deemed as relevant factors to base the decision upon. This, added to the fact that each task round looks the same and are presented in succession, leads to the respondent spending little effort and time at taking the decision (Orme, B. K., 2009), and hence providing rather thoughtless and injudicious data, making it necessary to recruit a bigger pool of respondents. On that behalf, and in direct connection with Research Question 2, one of the challenges of the new model would be to overcome this conflicting situation. By incorporating the Kano Model and an initial CBC with filtering purposes, we would be able to include in the final choice exercise only features that are relevant and desired for each individual (normally known in the Market Research world as evoked or consideration set) and hence weeding out unappealing ones. This would, in turn, encourage respondents to take a more deliberative stance at the experiment since the product alternatives presented would be more in line with their interests and preferences and thus will challenge them with real trade-off exercise. Although the details of how the AK-CBC model functions and aims to tackle this issue will be presented more in-depth in the Conceptual Model and Methodology sections, the following hypothesis was formulated:

H1: The new model would encourage respondents to provide more deliberate and thoughtful responses to each of the tasks presented

On that same line, due to the monotony of the questionnaire, the respondent may develop a non-compensatory behaviour based on the use of shortcut heuristics like the conjunctive model (Dobson & Kalish, 1993), in which respondents would rule out concept profiles if it simply does not include a certain attribute or instead contains an inadmissible one. As a way of illustration, a study from Hauser et al. (2009), showed that two-thirds of Californian car seekers rejected considering General Motors alternatives only due to the fact that GM had a poor brand name in the region, *"Investments in reliability, quality, safety, ride and handling, comfort, navigation, interiors, and Onstar become irrelevant if consumers never get beyond the consideration stage"* (Hauser et al., 2009). A more visual example would be an Apple unconditional supporter who would only consider the choice alternatives that include the brand Apple and disregard the rest of features shown. Contrarily, the CBC assumes a compensatory behaviour of deliberative and effortful weighting of each component attribute. This assumption, in turn, could create some degree of bias, specially for studies in which the product at hand has numerous attributes and levels.

This issue was originally addressed with the release of the Adaptive Conjoint Analysis (ACA) in 1985, which aimed at using past respondents' information for designing the conjoint tasks. The way it worked was really simple, by enquiring the respondent with an initial round of screening questions that focused on groups of two or three attributes at a time, it was possible to detect potentially relevant features and avoid including unimportant ones. The use of only a subset of the features in each product alternative, was later known as **Partial-profile CBC** and experienced a great acceptance among researchers. It allows to include in the study a large number of attributes, as only a subset of them are used in each task, and with a sufficient number of tasks, all features could be assessed. The logic behind it, is that by reducing the complexity of the choice exercises, in terms of limiting the number of attributes that shape each concept, respondents are encouraged to take a more thorough and careful examination of each product alternative which, in turn, might lead to longer experiments (Cunningham, Deal & Chen, 2010). The model has some intrinsic weaknesses as well, such as the inability to capture interaction effects in contrast to a full-profile CBC (as each pair of attributes only appear together a limited number of times) or incurring in bias when the respondent cannot efficiently compare concepts that are constantly showing a different composition of attributes. (Orme & Chrzan, 2017)

The goal of using past respondents' data to enquire them with choice experiments that better suit the idiosyncrasies of each individual was tackled with the commercial release of the Adaptive Choice-Based Conjoint (ACBC) in 2009, as a refurbishment of the declining ACA. By showing more fitting concepts, not only the use of simplifying heuristics could be avoided, but further information could be retrieved from each response. It is explained in the following section.

2.1.2 Adaptive Choice-Based Conjoint (ACBC)

After years of investigation, the Adaptive Choice-based Conjoint Analysis (ACBC), which sought to overcome some of the challenges exposed before, was introduced. As we will see in the next paragraphs, its main components seek to handle the use of simplifying rules of thumb and non-compensatory decision-making. Furthermore, and as I will describe in the following paragraphs, it tests the respondent with highly challenging exercises, leading to more dynamic and engaging interviews. This, in turn, allows the market researchers to capture more information from each respondent, improve the estimation of utilities, and to better predict real-world preferences, as respondents are encouraged to provide in-depth rather than superficial responses (Cunningham et al., 2010). The different sections that are encompassed in a regular ACBC are the following, although apart from the first section the rest are optional:

1. *Build your own (BYO)*: In this initial stage respondents can define their ideal product by picking their most preferred feature level combination. The added price is shown in this screen depending on the chosen level thus providing realistic price-attribute matches and allowing price to be modelled as a continuous function.

2. *Screener*: Using a nearest-neighbour approach, several product configurations resembling the ideal one are shown and the respondent is asked whether it appeals to him or not. Price is summed and randomly varied, and the option "None" is captured in this step.

3. *Must Have and Unacceptable*: Depending on the answers provided in the previous steps, non-compensatory decision-making behaviour is detected and confirmed, as respondents are asked again whether they find any attribute from a list a must-have or an unacceptable one.

4. *Choice tournament*: Personalized relevant choice exercise that looks like a regular CBC, but which only show product concepts as carried forward from the previous steps.

All these aspects served as a great reference and starting point for designing the model that we will present in the subsequent sections. However, ACBC has some noteworthy limitations which might explain why its use rate remains at a steady low of 13% of all conjoint analysis studies conducted with Sawtooth Software (Customer survey, Sawtooth Software, 2014). First and foremost, ACBC is sometimes regarded as a black box method, as despite its proven success in assessing must-have and unacceptable features, not much information could be inferred from these initial screening rounds as they are only in or out filters (for building an optimal configuration). Secondly, according to advanced designers, CBC tasks has more intuitive and user-friendly solutions for fine-tuning and personalizing the design as compared to an ACBC exercise, in which its increased complexity results in more restrictive and inflexible inputs, specially for inexperienced researchers (Jervis, Ennis & Drake, 2012). Furthermore, despite being more engaging and realistic, an ACBC exercise takes on average twice as time than a regular CBC to be finished by the respondent, potentially risking survey drop-put (Chapman et al. 2009). As a final note, for both CBC and ACBC, the software provider recommends to take forward no more than 12 attributes and 7 levels, as preliminary questions would be required to weed out irrelevant attributes. Some of these limitations or gaps will be addressed with the new model, as we will see in the section and which will underpin the novelty value of this study.

2.1.3 CBC design on-the-fly

For the purpose of scheming the AK-CBC, several assumptions are shared with the work of Adrian Martinez de la Torre for generating CBC designs on the fly (Martinez de la Torre, A. 2017). The method plainly consists on using past respondent information in order to gradually approximate the optimal product configuration. It achieves so by iteratively removing the least preferred attribute level in each task or iteration, namely the one with the lowest utility as estimated by the Multinomial Logit Model (which is further elaborated on the next section). Henceforth, by filtering-out non-informative attribute levels, the concepts shown in each choice task would be progressively approximating the ideal alternative and thus require more thoughtful assessment as each round would be more challenging for the respondent (heuristic behaviour would be avoided) and he or she will be more engaged to actually find the perfect concept.

What makes the CBC on-the-fly design different to CBC or ACBC is that the latter aim to unveil part-worth utilities to assess which are the most important features for the respondents, conversely the former´s main aim is to iteratively reduce the space of possibilities until the most preferred option is reached (Martinez de la Torre, A. 2017). On this behalf, and as we will see in the section , one shared objective of this work and the AK-CBC is to use respondents' data to determine which attribute levels to include in the different concepts shown.

2.1.4 Utility estimation

Since the output of the Conjoint Analysis, namely the part-worth utilities estimated for each of the levels, will play an important role in the conceptual algorithm described in the next section, the statistical methodologies used for the computations will be described hereafter. This ensures that the full theoretical frame that characterizes and impacts the model will be conveniently provided in advance.

When modelling discrete choice data, it is generally assumed that respondent j is a rational decision-maker, who bases his or her decisions in some utility-maximization criteria. When faced with a choice set k (given k = 1), the utility for alternative i could be expressed by $U_{ji} = X_i^{\intercal}\beta_j + \epsilon_i$, where U_{ji} represents the utility for the ith alternative, $X_i^{\intercal}$ is the transposed fixed vector of dummy codes describing alternative i (attributes and levels), β_j is the vector of utility part-worths for the jth individual and ϵ_i is the corresponding error term. The model simply assumes that the respondent bases his/her decisions on a set of utility estimates plus an error term or disturbance ϵ, which is included in the equation to represent the various unobserved influences that affect choice, as respondents' decision-making might fluctuate and not follow a perfect utility-maximization criteria.

The **Multinomial Logistic Regression (MNL)** is a regression analysis technique that heavily relies on the assumptions described above. With data from a Conjoint questionnaire or survey, the MNL aims at uncovering a set of utility weights or β's that when multiplied by a design vector (vector of recoded attribute levels that define a product alternative in a boolean fashion) could retrieve a fit to each respondents' choices. In the classical or frequentist approach this fit is approximated by the Maximum Likelihood estimation. On this behalf, it is important to note that the output of the MNL, namely the β's for each of the levels, are zero-centered scores, meaning that the sum of all the utility part-worths of the levels encompassed in a given attribute will add up to zero. As it will play an instrumental role in the analysis phase, it is also essential to introduce the concept of Preference Share for a given level, which can be interpreted as the probability of the jth individual picking the kth alternative for a given attribute, and its also a reliable approximation of the preference weight allocated by an individual to a particular level. For computing it, let Pr_k be the probability that a given individual would choose the kth alternative and $X_i^{\intercal}$ the transposed vector describing the whole space of ith alternative levels in a given attribute, we can define the generalized expression for MNL by $P_k = \dfrac{exp(X_k^\intercal \beta_j)}{\sum_{j=1}^{I-1} exp(X_i^\intercal \beta_j)}$.

To meet the maximum log-likelihood criteria, the variables need to be encoded into dummies as the model only functions for discrete data. Hence, the maximum likelihood estimate β for the parameter vector is obtained by maximizing the log-likelihood function, as denoted by the expression $\ell(\beta) = \sum_{i=1}^{N}(\sum_{j=1}^{J-1} y_{ij}\beta_{kj}x_{ik} - \log(\sum_{j=1}^{J-1} \exp(\beta_{kj}x_{ik})))$, and assuming that N respondents (i = 1,2, . . . , N) evaluate the same set of k alternatives and $J - 1$ parameters (j = 1,2, . . . , J-1) (one of them is set as reference, with utility 0).

It is also indispensable to introduce **Hierarchical Bayes (HB)** as a second method for utility estimation. The Multinomial Logit Model, as we will see in the section Conceptual Model, results highly relevant for the purposes of this work as it provides robust and specially fast utility estimations, but it has the grand disadvantage that it is not able to compute them on an individual or respondent level (which we would need for the analysis phase), only on a population level (Orme & Baker, 2000). The Bayesian approach, henceforth, came as a reliable and well-founded alternative for this. Sawtooth developed a software to estimate part-worth utilities using HB which has become one of the industry landmarks and prevalent solution ever since. As a consequence I will use the technical paper by Orme (2009) to introduce it.

The name Hierarchical comes from the fact that it has two different levels, while the "lower level" is driven by a regular Multinomial Logit Model, the "upper level" is governed by a Multivariate Gaussian Distribution characterized by a vector of means of the distribution of individuals (α) and a covariance matrix for that distribution (D) such that for the jth individual we have the vector of part-worths (β_j) described by $\beta_j \sim Normal(\alpha, D)$. In that sense, the algorithm follows an iterative process in which in each iteration the model refines and re-estimates the different parameters β_j, α and D. The process stops when the level of improvement in a given iteration is below a certain pre-specified threshold or cut-off point (also known as break out criteria) and thus convergence in estimation has been reached.

In each iteration, a new set of parameters is estimated (Lenk, DeSarbo & Green, 1996), starting by drawing a new estimate of α using present estimates of the β_j's and D. For it, and by assuming that D can be represented as the product of $LL^\intercal$ with L being a square, lower triangular matrix (Cheloski Decomposition), X being a normal vector of independent deviates with zero mean and unit variance and Z the result of multiplying X by L, we can establish that $\mathbb{E}(XX^\intercal) = I$ and henceforth, D can be approximated by $D \approx \mathbb{E}(ZZ^\intercal) = \mathbb{E}((LX)(LX)^\intercal) = \mathbb{E}(LXX^\intercal L^\intercal) = L\mathbb{E}(XX^\intercal)L^\intercal = LIL^\intercal = LL^\intercal$. Using this, and for n number of individuals it is possible to draw the vector $\alpha + LX$, with mean α or the average of the current β's and with covariance matrix $\frac{D}{n}$ from a multivariate distribution.

For estimating D in each iteration we follow a similar procedure. Assuming we have p parameters and n individuals, the prior estimate of D would be equal to the identity matrix I of p order. We then compute a matrix H that joints the prior information with the updated estimates of both parameters such that $H = pI + \sum_n (\alpha - \beta_i)(\alpha - \beta_i)^\intercal$. Repeating again the Cholesky Decomposition, we can draw a subset of n random vectors with zero mean and unitary variance such that the new estimate for D would be S^{-1} assuming that $S = \sum_n (LX)(LX)^\intercal$.

The last step of the iteration, and also the most relevant for this work, is the new estimation or draw of the β's which can only be possible with the new estimates of α and D and by employing what is known as the Monte Carlo Markov Chain (MCMC) (Ishwaran & Zarepour, 2000) at an individual level and until converging to an optimal value. To describe it, it would be important to firstly introduce the Bayes' Theorem, which governs the whole HB procedure, and that can be represented by $P(\theta \mid H_i) \propto \dfrac{P(H_i \mid \theta)P(\theta)}{P(H_i)}$ where:

$P(\theta)$: Probability that the hypothesis about parameter i before observing the data is true. Also known as prior probability,

$P(H_i \mid \theta)$: Assuming a set of values θ, probability that those values are part of the data under the condition of hypothesis H_i. It is also regarded as likelihood of the data,

$P(H_i)$: In this it is assumed to be a constant that normalizes the probability into a unitary value,

$P(\theta \mid H_i)$: Known as the posterior probability, represents the conditional probability of the Hypothesis on the assumption that prior H_i is true and also θ is comprised in the data.

Acknowledging this, the MCMC or the specific implementation used by Sawtooth for its HB software, the Metropolis Hastings Algorithm, considers the β estimated in the previous iteration as prior information (from now on referred to as β_1) and aims at generating a new parameter β_2 with the goal of testing whether improvement is reached with a more fitted estimation.

At the lower level, we compute for both β_1 and β_2 the likelihood of the choices made by the jth individual, using MNL and the P_k expression described on its corresponding section, which would be called respectively p_1 and p_2. At this point we also need to compute the relative densities for both β's given current estimates of parameters α and D which we will call d_1 and d_2 and that are scalars computed by the expression $d_k = e^{(\frac{(\beta_1-\alpha)^\intercal}{-2})(\beta_k-\alpha)D^{-1}}$

Finally, we compute the ratio $\frac{p_2 d_2}{p_1 d_1}$, if it is larger than 1 we can accept the null hypotheses that the posterior probability of β_2 is larger than the one for β_1 and henceforth it will become the estimation for the next iteration.

Assessing performance

The validity of the estimations can be measured using what is known as **fixed** or

hold-out choice tasks. These tasks are not used for training the model and estimating utilities but are rather kept aside and used as a validation set with the object of measuring the performance of our prediction. They have the same format and configuration as any other regular task that could be presented in the Conjoint experiment but the difference relies on the fact that the attributes and levels included on the product concepts shown are fixed and the same for all respondents. For this we assume $\hat{E}_j$ to be the predicted choice, on the basis of being the product alternative or level with the largest overall utility, and hence preference share, across all concepts presented in the task ($\max\limits_{0 \leq P \leq 1} P_{\hat{E}}$). Similarly, if we assume E_j to be the actual choice made by respondent j, we can ascertain the (predictive) performance by the proportion of matching predicted and actual choices such that $\hat{E}_j = E_j$, better known as hits. This metric normally takes the name of hit rate. Resulting from these notions, and in direct connection with Research Question 1, the following hypothesis was formulated:

H2: The new model would improve the predictive performance of a regular CBC in terms of hit rate

2.2 Kano Model

Many authors in the field of psychology and behavioral sciences have focused their efforts at understanding what drives human needs and motivations. Arising from this, numerous models and frameworks have tried to shed light on the potential factors that might help explain, interpret and ultimately fulfill those. Abraham Maslow (1943), with its universally popular Hierarchy of needs or Maslow's pyramid, aimed at uncovering these latent motivations by presenting a five-tiered model based on the idea that individuals must first satisfy the lower level needs before moving on to meet higher-tiered ones, henceforth assuming a different hierarchy among needs. On a related note, the motivation-hygiene or two-factors theory by Frederick Herzberg (1959) stated that the relationship between quality attributes and customer satisfaction does not necessarily correlate symmetrically and linearly, and therefore not just by improving a product or service's atomic features, customer satisfaction would be automatically enhanced. Alternatively, he distinguished between hygiene factors, which cause dissatisfaction and hence should be eliminated, and motivator factors which on the other hand boost satisfaction and should be fostered. The model was applied in the context of job satisfaction in the workplace and its results and innovations had a massive impact in other fields of marketing and consumer motivation.

Heavily influenced by this approach and with the intention of extending those notions to the field of product development, Noriaki Kano (1984) presented the Kano Model as a way of mapping and modelling customer satisfaction under the assumption that not all attributes embedded in a product or service are equal in the eyes of the end-user. With the goal of providing manufacturers with meaningful guidelines on which product requirements should be prioritized in the production workflow, Kano and his team introduced consumer expectations and argued that some factors or features could deliver higher added value by enhancing customer loyalty more than others.

As the Kano Model (Kano, 1984) allows to categorize and prioritize a product's differentiating features in terms of its level of integration and the degree of satisfaction fulfillment, it will play a vital role for designing the AK-CBC. In the field of Market Research, the Kano Model is foundationally used to identify customer needs, particularly useful for configuring a new product or service concept on the basis of its atomic features, and for enhancing the identification of exciting requirements, usually associated with innovations (Tontini, 2007). Although these goals might seem to overlap and have many things in common with the description of the traditional Conjoint Analysis, there are several differentiating components among the two methods, which will be duly compiled in Table 4. However, in essence, the Kano has a more explorative, linguistic and qualitative nature (Sauerwein, Bailom, et al. 1996) and its main purpose is to categorise the different features of a product in terms of relevance for the customer, for ultimately revealing explicit needs. By responding a really simple and straightforward questionnaire consisting on a pair of functional and dysfunctional closed question forms, as shown in Table 2, and on the basis of the Kano evaluation table as per Table 3, the researcher would be able to classify each of the features that shape a product in one of the following categories:

- *Must-Be*: As these attributes are taken for granted in the product configuration, if they are implemented would not affect customer satisfaction but if it is not part of the product definition it would bring extreme dissatisfaction. A visual and extreme example could be having a seat belt in a car or not.
- *One-Dimensional* or *Performer*: The presence of the attribute will increase satisfaction level while its absence will proportionally decrease satisfaction level.
- *Attractive* or *Delighter*: Since it is not expected to be implemented, its presence generates high levels of positive satisfaction while customers will not be dissatisfied at all when it is not fulfilled.
- *Reversal*: As its name indicates, their integration might harm customer satisfaction while its absence could result even appealing.
- *Indifferent*: The respondent feels uninterested about it and the incorporation of an indifferent feature would not affect or have an impact whatsoever on customer satisfaction.
- *Questionable*: Either the question was not understood by the respondent or an illogical response was provided.

This classification results more explicit and comprehensible when taking a look at the Kano Diagram in Figure 1. On this graph we can observe that the vertical axis refers to the level of satisfaction provided by the feature going from dissatisfied to satisfied. On the other hand, the horizontal axis represents the level of implementation of the feature, namely whether or not it has been integrated, and ranges from dysfunctional (not implemented at all) to fully functional (best implementation possible).

Table 2: Functional and dysfunctional form of the questions in the Kano questionnaire, as well as the available response alternatives

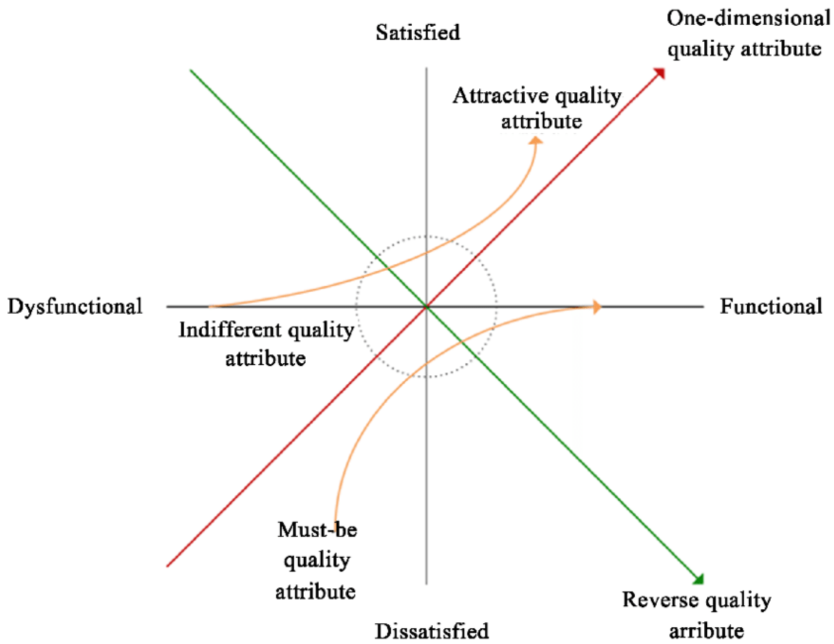
How would you feel if Product X has Feature Y? (Functional form)	I like it that way
	It must be that way
	I am neutral
	I can live with it that way
	I dislike it that way
How would you feel if Product X does not have Feature Y? (Dysfunctional form)	I like it that way
	It must be that way
	I am neutral
	I can live with it that way
	I dislike it that way

Table 3: Kano Evaluation Table

Customer Requirements		**Dysfunctional form of the question**				
		I like it that way	It must be that way	I am neutral	I can live with it that way	I dislike it that way
Functional form of the question	I like it that way	Q	A	A	A	O
	It must be that way	R	I	I	I	M
	I am neutral	R	I	I	I	M
	I can live with it that way	R	I	I	I	M
	I dislike it that way	R	R	R	R	Q

After a careful assessment of both Conjoint Analysis (and more particularly Choice-based Conjoint) and the Kano Model, it is noticeable that both methods share a common goal of identifying which are the most relevant attributes and feature levels of a good or service for prospective users. On the one hand, the Kano might result particularly useful as an orderly way to sort and prioritize which are the features that might have a larger impact on customer preference, and thus yield larger acceptance, when designing and launching a new product concept. However, it does not allow to identify and quantify a specific combination of features, at a specific amount and at a specific price range that would most likely be a hit in the market. Conjoint Analysis, on the

Figure 1: Kano Diagram with the various categories plotted in a bidimensional array. Original picture from Jiawen Huang, School of Management, Jinan University



other hand, allow the researcher to quantitatively estimate market size and potential revenue as well as uncovering expected customer segments by defining and refining what would be the optimal set of product alternatives, namely which features, levels, amounts and price points, to launch into the market. In order to crystallize the main distinguishable elements of each method, a list of key differences between CBC and a Kano Model, from a behavioral and methodological perspective, have been compiled in Table 4.

2.2.1 Analytical Kano (A-Kano Model)

As seen in the previous section, the Kano Model is a widely-used and useful tool to classify and prioritize customer needs, as it captures the non-linear relationship between product performance and customer satisfaction. Despite of this, its qualitative nature and its linguistic-based origins makes the classification of features a good approximation of reality, but sometimes the differences are a bit fuzzy and the margins that separate each category too soft to make strong assumptions (Mikulic & Prebezac, 2011) and (Xu et al., 2009). The work by Xu et al. (2009) also introduced a more quantitative and sturdy extension of the model, called Analytical Kano or A-Kano Model.

Table 4: Key differences between Choice-based Conjoint Analysis and the Kano Model

	CBC	**Kano**
Interaction between features	The interaction between attributes is the key component in a Conjoint exercise since it is the actual trade-off among different competing levels which makes the utility measurement possible.	In the basic Kano approach, each feature is disseminated and evaluated individually. Despite of this, the features can be plotted in a bidimensional axis on a population level afterwards.
Linearity in preferences	The output of the CBC, namely the list of utilities for each of the levels included in the trade-off exercise, is normally zero-centered. Having normalized utility partworths allows marketers to interpret the results as the magnitude and sign of them are linearly connected to its corresponding preference and importance.	It is noticeable in Figure 1 that not all feature categories are linear, in the sense that the satisfaction is not linearly correlated to its level of integration in the product. Both Attractive and Must-be features have an exponential curve and the other categories have a lineal pattern but with different steepness.
Measurement and simplicity	Methodological expertise is required in order to obtain robust utility estimations. Not only it is indispensable to use advanced statistical methods like Multinomial Logit or Hierarchical Bayes but it is also imperative to build well-balanced and complex Conjoint designs.	Both the set-up (consisting of a pair of functional and dysfunctional questions) and the subsequent analysis has a more simple nature and requires only simple math computations.
Impact of time	Utilities undoubtedly vary in time, and its estimations can be modelled across it using simple regression algorithms. Despite of this, its impact is directly dependable on the set-up used.	Time has a more clearly defined impact on the different categories. As categories heavily rely on customer expectations, these expectations might vary across a product's life cycle. For example, Attractive or Performer attributes might become Must-be as market and its features mature.
Sample size	Although the necessary sample size heavily depends on the complexity of the study design, in terms of number of attributes and levels, a minimum of 100 respondents is normally advised.	According to initial explorative investigations carried out by Griffin and Hauser (1993) it was found out that a sample of around 20 to 30 respondents in an homogeneous segment would be sufficient to determine more than 90% of all possible product requirements.

For the purpose of designing the AK-CBC, I will retrieve some theoretical concepts and notation from this work that may enhance the designer´s understanding of customer needs and define a more robust decision criteria. The basic theoretical framework to understand the model is presented below, which also would serve as a starting point to describe the new model in the early stages. Being *s* the market segment that contains a total of *J* homogeneous respondents $s \equiv \{t_j \mid j = 1, 2, \cdots, J\}$, a set of features is defined by $F \equiv \{f_i \mid i = 1, 2, \cdots, I\}$. In accordance to Table 3, we have that for each respondent $t_j \in s(\forall j = 1, 2, ..., J)$ the evaluation of each attribute $f_i(\forall i = 1, 2, \cdots, I)$ is represented by $e_{ij} = (x_{ij}, y_{ij}, w_{ij})$ in which x_{ij} is the score for a certain feature for the dysfunctional question, y_{ij} is the score for the functional form of the question and w_{ij} represents the self-stated importance for that feature normalized to the range $\{0 - 1\}$.

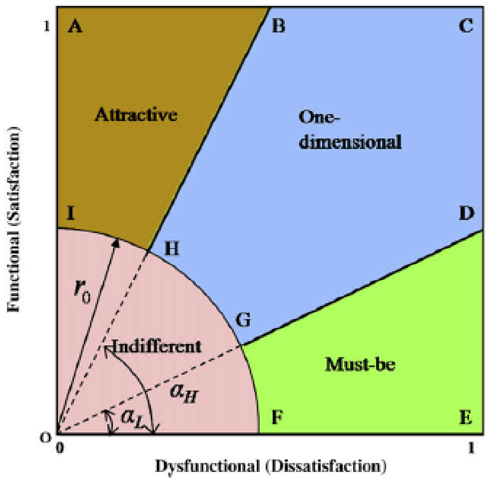
The study further details that for each feature f_i and market segment *s* the average level of satisfaction for the dysfunctional question ($\bar{X}_i$) and functional form ($\bar{Y}_i$) is represented respectively by $\bar{X}_i = \dfrac{1}{J}\sum_{j=1}^{J} w_{ij}x_{ij}$ and $\bar{Y}_i = \dfrac{1}{J}\sum_{j=1}^{J} w_{ij}y_{ij}$.

As seen in Figure 2, the value pair $(\bar{X}_i, \bar{Y}_i)$ can be represented in a bi-dimensional diagram in which the horizontal axis stands for the dissatisfaction score and the vertical axis represents the satisfaction value. From the customer´s side, a certain attribute can be represented as a vector $f_i \sim R_i \equiv (r_i, \sigma_i)$ in which the magnitude of r_i (which is called the importance index) equates $\sqrt{\bar{X}_i^2 + \bar{Y}_i^2}$ and it denotes the overall importance of that feature. On the other hand, σ_i, also known as satisfaction index, is the angle between σ_i and the horizontal index and it can be approximated by $\sigma_i = tan^{-1}(\frac{\bar{Y}_i}{\bar{X}_i})$. Both σ_i and r_i are known as Kano indices and help measure customer satisfaction in a quantitative fashion. The study also introduced the term Kano classifiers which are basically thresholds that help categorize each feature in a more systematic way. By defining $K = (r_0, \sigma_L, \sigma_H)$ we can divide the vectorial space of the Kano Diagram in order to obtain a robust a clear differentiation for each of the features, as seen in Figure 2.

Table 5: Scores for functional and dysfunctional features.

Answers to the Kano Question	**Functional form of the question**	**Dysfunctional form of the question**
I like it that way	1	-0.5
It must be that way	0.5	-0.25
I am neutral	0	0
I can live with it that way	-0.25	0.5
I dislike it that way	-0.5	1

Figure 2: Kano classifier and Kano categories. Original picture from Xu et al., (2009), published in www.sciencedirect.com. For this study, the parameters α_L and α_H as represented in this figure, have been renamed to σ_H and σ_H respectively

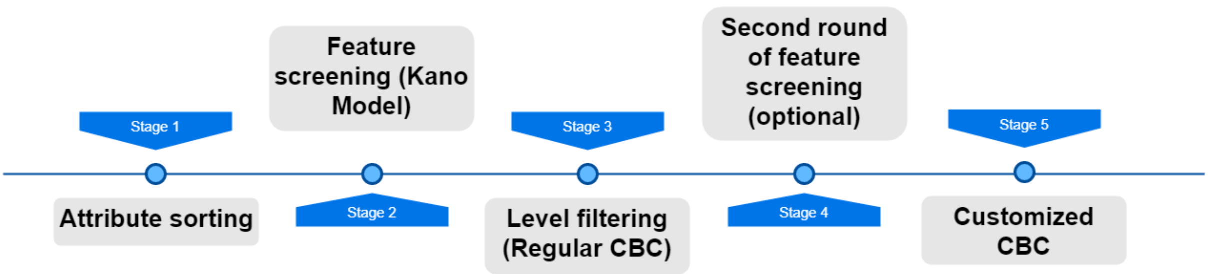


3 Conceptual Model

3.1 Introductory considerations

The new model (AK-CBC) is named after the different methodologies (CBC, Adaptive CBC and the Kano Model) and the extensions of which have inspired its design and structure. As could be noted in the "Theoretical Framework" the multiple methods and approaches described in it had strengths and weaknesses that made them suitable for some studies but relatively unfitting for other cases. On this behalf, the model aims at providing a novel yet standardized and generalizable approach for Conjoint studies that involve complex product concepts with more than 12 asymmetric attributes (different number of levels) such as technological devices that are in the product development stage or any other good or service that is prone to include multiple extras and complementary features besides the base case product configuration. The different steps that shape the AK-CBC algorithm are depicted in Figure 3 and will be further detailed in the subsequent sections, but before delving into the details of the conceptual model, it is important to mention that despite not being inspired by it, this algorithm coincidentally shares some notions and terminology with the work by Wang and Wu (2014) and Wang and Wang (2014). In section 3.3, after the algorithm description, coincidental and differentiating points will be duly described as well as how the limitations of the rest of the methodologies were tackled.

Figure 3: Summary Diagram for the AK-CBC



3.2 Algorithm Description

3.2.1 Attribute Sorting

As mentioned in the introduction, this model would be particularly well-suited for products with many features or extra complements. That being said, the first step would be to sort the different attributes by <u>Core features</u> and by <u>Extra features</u>. The Core attributes, represented as $F_C \equiv \{f_{Ci} \mid i = 1, 2, \cdots, I\}$, are distinguishable over the fact that they will always be part of the base case product configuration and integrated in the product profile regardless of brand or price range. Commonly, multiple levels are comprised in the attributes of this subgroup (i.e. > 2). By contrast, Extra features, represented by $F_E \equiv \{f_{Ei} \mid i = 1, 2, \cdots, I\}$, may be perceived as alternative attributes or add-ons, which might be offered on a brand basis, and normally have a binary or dichotomous nature in the form of either it is included in the alternative or not, namely $f_{Ei} \in \mathbb{B}$. As a practical illustration, we can observe that in Table 7 of section 4.1, attributes 1 to 7 might be handily categorized as Core features with a number of levels denoted by $v_C = \{5, 6, 3, 5, 4, 3, 3\}$, whereas attributes 8 to 18 could be classified as Extra FRs and expressed as $v_E = \{2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2\}$, as said in a dichotomous fashion.

3.2.2 Extra Features Screening (Kano)

At this stage of the process, the goal is to assess how each respondent perceives and categorizes the different Extra features of the study. For that, a first round of Kano is applied at the individual level over a random subset of the full-list of v_E. In order to avoid survey monotony, and thus respondents' fatigue, if the number of Extra features is ≥ 6, then $\frac{F_E}{2}$ would be retrieved for consideration, otherwise the entire selection would be used in this round. For each respondent $t_j \in s_k (\forall j = 1, 2, ..., J)$ the evaluation of $f_{Ei} (\forall i = 1, 2, \cdots, I)$ is represented by $e_{ij} = (x_{ij}, y_{ij})$. Depending on the different combinations of x_{ij} (Dysfunctional form question) and y_{ij} (Functional form question) and as per Table 5 we would be able to classify the different extra features into the various categories contemplated in the Kano. The information captured in terms of how each respondent perceives the different extra features would be duly compiled and recorder for its use at a later stage of the algorithm.

3.2.3 Level Filtering

At this juncture, the goal is to approximate the closest optimal base case product configuration, namely alternatives that only include combinations of Core attributes or F_C. For this, several rounds of regular CBC tasks are presented in order to compute the utility part-worth for each level, with the ultimate objective of removing the most unappealing levels from the final choice exercise, as explained in subsection 3.2.5. The initial idea was to use a CBC design on the fly (as explained in 2.1.3) so the least preferred level in each iteration could be removed. Eventually, this idea was discarded since the implementation this methodology was quite computationally-intensive and could lag the survey considerably. For that reason it was decided to employ a standard randomized CBC, the results of which would be collected to be used at a later stage instead of removing levels on the fly.

3.2.4 Second round of Feature Screening

The model continues by inserting an additional round of Kano questions. Its integration would have an optional nature since it will directly depend on the number of extra features considered for the study (F_E), as mentioned in subsection 3.2.2. Its mechanism would be exactly the same as the previous Kano exercise, and the primary intention of alternating CBC and Kano questions is to maximize respondents' engagement on the survey and to avoid any potential fatigue that may arise from presenting multiple tasks of CBC in succession.

3.2.5 Customized CBC

The model (AK-CBC) ends up with a final set of CBC tasks. One of the key innovations of this dissertation is presented at this point, as the design of this last round of choice tasks will be tailored and adapted on the basis of the information captured in the previous Kano and CBC exercises. The objectives and bearing at this point is twofold. On the one hand, the aim is to bring forward highly-challenging choice tasks by presenting the most appealing product configurations possible. That would stimulate the respondent to put great effort to make a decision instead of using simplifying heuristics and thus allowing the market researcher to collect highly valuable and thoughtful information. On the other hand, by including the different Extra features F_E in the choice exercises, as opposed to the previous round described in 4.2.3 which only included Core Attributes F_C, we would be able to estimate and quantify the utilities, and thus the willingness to pay, for the various extra features as categorized by the Kano Model, which in turn, would deliver a straightforward approximation of how these attributes can perform as drivers of choice and key value generators in the market.

As per each individual respondent, and on the basis of his or her prior responses in the previous sections of the questionnaire, the principles compiled in Table 6 will be followed in a systematic and generalized way. It is important to mention that for this stage, and given that we needed to estimate the utilities for each individual at a time, the regular Multinomial Logit was used for filtering out the least preferred level of the Core attributes instead of Hierarchical Bayes, given that the package for this model resulted significantly faster, easier to implement and equally reliable when computing each individual's part-worth estimates, as introduced in the Theoretical Framework section.

Table 6: Description of the holistic principles of the Model at this stage

	Core Attributes (using CBC and Multinomial Logit)	**Extra Attributes (using Kano Model)**
Filtering	For all F_{Ei} 1) If number of levels ≥ 7 (Remove 2 least preferred) 2) If number of levels ≥ 4 (Remove least preferred) 3) If number of levels <4 (Only remove one level if it is the least preferable in overall terms).	Remove Indifferent (I), Questionable (Q) and Reverse (R) features.
	Goal: Mitigate Number-of-levels effect (Huber, Zandan, Johnson Wittink, 1992) and balance the attribute consideration set.	Goal: Reduce size of attribute set by discarding noisy features which add no value to the respondent or even reduces utility. According to Xu et al. (1996) individuals feel uninterested about these features and the willingness to pay for including them is unaltered.
Partial Profiling	Include in the study all features.	The number of Extra features to include would depend on market requirements and industry guidelines. For this case, 1 Extra will be included in each alternative.
	Goal: As Core features are always part of any product profile, it would realistically mimic a market offer	Goal: Reduce the number of features included in each product profile to avoid overwhelming the respondents

An important note to discuss is that conditional pricing (summed pricing) was also considered to be included in the customized CBC. In the sense that in accordance to the theory that defines the Kano Model, highly appealing features (One-Dimensional and Delighters), which would drive positively the satisfaction of the respondent, would be linked to higher price points (by setting prohibition to the lowest price levels in case of appearance), whereas Must-be attributes, the implementation of which would only avoid dissatisfaction, would be hardwired to the lowest price ranges. With this practice, the goal was to defy further the respondent with more challenging alternatives by linearly connecting their past decisions (in terms of how they perceive these extra features) to a realistic price point. Nevertheless, it was ultimately determined that by keeping the same unaltered price intervals as in the initial CBC task (section 3.2.3) we would be able to assess the value assorted to each of these extra features by comparing the different utilities scored by the Base case product configurations (including only Core attributes) and with the inclusion of the previously categorized Extra features.

3.3 Closing points

One of the main issues that we have discussed along the study so far, is the respondents' adoption of non-compensatory decision-making behaviour when the product concepts presented in CBC tasks do not accurately suit their needs and desires. In that sense, this translates in rushing through the survey with the use of simplifying heuristics and thus not providing deliberative responses. To tackle this, and in line with the rationale of the Adaptive Choice-based Conjoint also presented in the section "Theoretical Framework", we aimed at using past respondents data to better tailor the product profiles to each individuals' preferences on the fly. In the case of the ACBC this is done by initially asking the respondent to design their own optimal product configuration beforehand. With the AK-CBC we manage to do that by presenting an initial CBC and Kano exercises to weed out the least preferred levels and unimportant features. On a related note, one potential weakness of the ACBC was the "black-box" nature of its screening criteria, in the sense that little information could be inferred from the filtered-out attributes and levels apart from the fact that they were out of the consideration set of an optimal product definition for that particular individual. By designing the AK-CBC, we propose the novel idea of adapting the choice tasks by disregarding uninformative and undesirable features (Kano) and levels (CBC) while at the same time retrieving information on why these features were removed in the first place, as they have been previously categorized in the Kano exercise.

As we discussed in the introduction of the section, the model is particularly suitable for products that involve many extra features or add-ons. The AK-CBC would allow us to estimate the importance or utility part-worths for each of this extra features, that have been previously categorized as Must-be, One-Dimensional or Delighters. With that, we would be able to analyze how each of these categories behave as key decision drivers in an hypothetical market situation. For instance, we would be capable of assessing which categories drive or influence choices and decisions the most, would it be Must-be features or Delighters? On that sense the following hypothesis was formulated in order to be tested and measured on the terms described in the Methodology section:

H3: Must-be features would have larger part-worth utilities than One-Dimensional or Delighter attributes respectively

Finally, we mentioned in the introductory remarks of the section that this study has points in common with the work by Wang and Wu (2014) in the sense that both algorithms use an initial CBC for the core attributes and the Kano for the Extra or optional features. Besides that, the coincidences are purely anecdotal given that both the goals and the methodology employed are substantially different. Most relevantly, the dissimilarities rely on the fact that the AK-CBC uses a script to capture respondent's preferences and design a CBC for each individual on the fly whereas the work by Wang and Wu aim at segmenting them on the basis of their responses to both exercises for ultimately providing and ranking product configurations that maximize overall satisfaction for each of the segments.

4 Data and Methodology

4.1 Attribute formulation

By taking into account all the guidelines related to the design of a Conjoint study, exhaustively researching the market of smart speakers, and brainstorming potential innovative features, the following attributes and levels were determined as the input for this study:

Table 7: List of attributes and levels as input for the study

Attributes	**Levels**
1. Brand	Amazon, Google, Apple JBL, Sonos
2. Price	49.99, 79.99, 119.99, 179, 229, 349
3. Size	Pocket size, Hand size, Desktop size
4. Colour	Charcoal, Dark gray, Light gray, Blue, Red
5. Shape	, , , 
6. Integrated Languages	English, English + elective, 5 or more
7. Material	Fabric cloth, Hard plastic, Aluminum
8. Hi-fi (High fidelity audio)	Yes/No
9. Interchangeable skin/shell (To personalize the speakers' design with detachable shells with different colors and styles)	Yes/No
10. Integrated camera (To take group pictures, check the pet when you are not at home or receive notification if unusual movement is detected)	Yes/No
11. Touch screen	Yes/No
12. In-built projector (Enjoy movies, TV or gaming through the combination of visuals and sound for a Home-Cinema experience)	Yes/No
13. Smart Home Domotic Technology (Control compatible smart devices at home like thermostats, blinds or lights through easy and straightforward voice commands)	Yes/No
14. Compatibility with 5G technology	Yes/No
15. Personalized voice (Instead of the traditional robotic voice, choose from a list of human-like voices or even celebrities)	Yes/No
16. Premium sound recognition (Send a notification to the users' synced smartphone when a particular sound is recognized in another room i.e. a baby crying, or an oven beeping)	Yes/No
17. Personalized Interactions (Adapt the recommendations and conversations depending on the person using the device)	Yes/No
18. Guarantee of confidentiality (Contract for ensuring personal privacy would be attached to the purchase)	Yes/No

4.2 Methodology and Data collection

The practical application of this model was only possible with the use of the latest computing service offered by Amazon, AWS Lambda, which allows to run back-end code in their cloud servers in parallel to other processes. The R script used to make this algorithm possible runs automatically as each respondent answers the survey questions through Lighthouse Studio (Sawtooth Software's tool for Conjoint-related surveys) and their responses to the Kano and CBC exercises are compiled and processed on-the-fly to provide the adapted and tailor-made Conjoint design (AK-CBC) as an end result.

On this behalf, this work not only aims at introducing this new conceptual model and testing its performance, but it intends to put it into practice and serve as a disruptive addition in the field of Machine Learning Model Deployment. This discipline focuses on seamlessly developing and integrating Machine Learning algorithms into existing production environments such as the case of traditional Conjoint Analysis. It is by bringing together and unifying the development of this new ML-related model through the existing AWS software with the current operational Conjoint practice that we could take a major leap in the DevOps domain. Furthermore, since this project was done in cooperation with my current company SKIM, the study aims at laying the foundations to use this technology not only for the current study but for future projects that require adapted designs on-the-fly through scripts from the most used open-source programming languages like Phyton or R.

From a functional and technical perspective, the adapted CBC (AK-CBC) exercise worked better than expected without lags or technical issues (according to respondents' feedback), but in order to prove its validity and to provide empirical evidence to test the Research Questions and its associated Hypotheses, the model was evaluated using primary quantitative data from computer-administered surveys. On this behalf, I wanted to address a pool of respondents that particularly fit this study, henceforth the following initial sample specifications were arranged for the data collection process:

- Target group: Consumers who are familiar with smart speakers, specially focusing on those that already own one or are interested in buying a smart speaker in the future.

- Countries: United States (according to previous studies this country show high rate of acceptance and usage with these devices).

- Number of complete surveys: N = 600 for the Custom Conjoint exercise and and N = 600 for the regular random-generated Conjoint.

- Study type: The estimated study duration (LOI) would be around 12 minutes for each online survey, although since the study design would be completely novel there is some degree of uncertainty regarding the approximate duration.

In order to test the two first RQs, as introduced in the section "Introduction", and their corresponding Hypotheses (both can be found in Table 8), half of the total sample carried out the adapted or hybrid CBC (AK-CBC) while the other half did a regular random-generated CBC which did not incorporate any type of adjustment or adaptiveness and all attributes and levels were shown an equal number of times in a random fashion. In that sense, and in order to ensure unbiasedness, all respondents completed the same survey, all undertaking the same demographic questions, Kano exercise and initial CBC (with only core features) that are required for the AK-CBC. Then, at that point, half of the sample were randomly assigned to undertake the custom AK-CBC and the other half the random-generated one. For maximizing comparability of results, the structure and format of these last choice exercises were exactly the same for both groups, with 6 tasks per respondent, 4 alternatives per task and 8 attributes per concept (core attributes plus the extra feature). The only difference therefore relies on whether or not the final CBC exercise was adapted to the individual preferences of that particular respondent (through undergoing the AK-CBC algorithm) or simply showed all attributes and levels without any adaptiveness of any kind (random-generated CBC).

The number of respondents was set at around 600 for each exercise to ensure robustness in prediction and results. RQ1 would be analyzed in terms of hit rate on a hold-out task in which respondents are asked to choose their favourite level for each of the smart speaker attributes included in the previous CBC (as explained the end of section 2.1.4 Utility estimation). By comparing the utilities computed for the different levels and contrasting it with the alternative chosen in the fixed task we will have a solid approximation of the predictive power of this new model as opposed to the random one. For the second RQ, engagement will be measured on the metric elapsed time, namely how long did each respondent spent making a decision in each of the six tasks of the AK-CBC as compared to the same ones from the random generated experiment.

Table 8: Research Questions 1 and 2 and their respective Hypotheses

Research Question	Hypothesis
RQ1: What would be the predictive power, in terms of identifying customer preferences, for the hybrid model as compared to a regular random-generated Choice-based Conjoint study?	The new model would improve the predictive power of a regular CBC in terms of hit rate.
RQ2: How would the new model perform in regards to user engagement as compared to a regular random-generated Choice-based Conjoint exercise?	The new model would encourage respondents to provide more deliberate and thoughtful responses to each of the tasks presented.
RQ3: What is the relationship of each product feature categorization, with its corresponding utility estimate?	Must-be features would have larger part-worth utilities than One-Dimensional or Delighter attributes respectively.

As for RQ3, the hypothesis will be tested on a straightforward assessment of the utility scores for each of the categories that moved forward from the Kano exercise (Delighters/Attractive, One-Dimensional and Must-be) on both an individual and population level in order to assess which one of them have a more impactful effect on preferences.

4.3 Pre-processing and Data cleaning

Once the final data-file is obtained, it is important to apply several basic pre-processing and cleaning steps in order to ensure that the information compiled is non-fraudulent, robust, meaningful and interpretable. In line with industry practices, a scheme based on points would be used as described below. On that behalf, those respondents that receive a minimum score of 2 points will be removed from the final consideration set.

1. Repeated IPs: The fact that there is more than one line from the same IP, may be a strong indication that either some technical incident happened or the fieldwork agency purposely or accidentally allowed the respondent to take the survey more than once. Either way, those lines should be removed to avoid any potential bias on the results. Those respondents with a duplicated IP would receive 2 points.

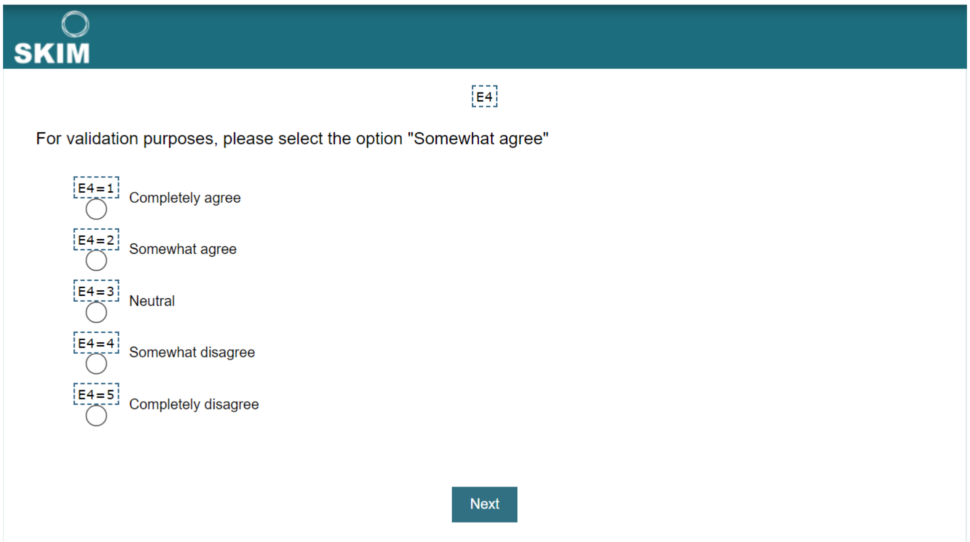
2. Flat-lining: 'Flat-lining' refers to respondents that answer the same option for various questions in succession. It is a sign that a respondent is giving illegitimate answers and/or rushing through the survey. If respondents provided the exact same answer to all 22 questions of the Kano exercise, we assumed that he or she was not allocating the required level of scrutiny and would be therefore useless to the purpose of investigating preference, those respondents would receive 1 point.

3. Elapsed time: The data-file that comes back from the survey software indicates how much time each respondent spent in each of the questions. Respondents that spent an unreasonably short amount of time, below a pre-specified threshold in which they would not have enough time for carefully reading and assessing the exercise (in this case set as below 0,4*median of the whole sample) would be also cleaned out, as no meaningful information can be inferred from those responses, those respondents would receive 2 points.

4. Question for quality control: Below these lines, in Figure 4, we can see the quality control question that was included to analyze whether the respondent

was paying careful attention to the survey and reading all questions thoroughly. If this attention check is not passed, respondents would receive 1 point.

Figure 4: Control question as shown to respondents

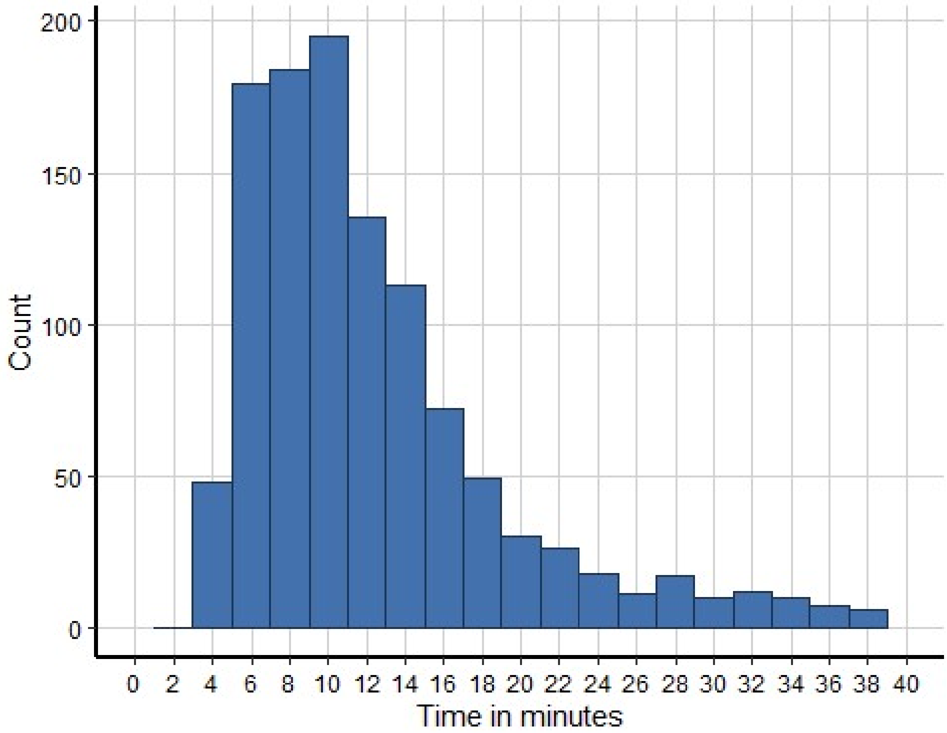


5. Illogical responses and outliers: Extreme numbers that are highly distant from the rest of the data in open-end questions like the one for specifying the age should be thoroughly investigated, as they not only impact descriptive statistics like mean and variance but also could be a strong sign that the survey is not taken seriously. These outliers receive 1 point.

4.4 Descriptive Statistics

Fieldwork ended up with N = 1234 completed surveys (N = 600 for the custom design and N = 634 for the random-generated), taking a total of 25 live hours to achieve this quota (September 18th at 11 am - September 19th at 12 am). Incidence Rate or IR (proportion of all population that qualified for this survey) was 74% and Conversion Rate (proportion of respondents that finished the survey to those that started it) was 52.6%. The average lenght of the interview (LOI), namely how much time respondents spent on the survey, was 10 minutes, and on Figure 5 we can see the frequency distribution once the outliers were removed. I considered outliers those respondents that spent more than 40 minutes on it as it was assumed that the individual dropped the interview momentarily and resumed it moments after (4.18 % of the total).

Figure 5: Histogram for the Length of Interview (LOI)



A total of 62 respondents did not qualify after the data clean-up process (5.02 %), as described in section 4.3, and were dropped out from the final consideration set, so data collection consisted of N = 1172 respondents, N = 571 taking the custom design exercise and N = 601 for the random-generated tasks. Since this was the sample that finally moved forward to the analysis phase it is important to take a moment to assess its main characteristics from different angles. In terms of Demographics, we can observe in Table 9 the distribution for gender, age and education.

On a related note it was also meaningful to understand the psychographic idiosyncrasies of this set of individuals as well as their attitude towards smart speakers. In order to amplify the accessible pool of respondents to be reached for the survey, I changed the screening criteria from "individuals that either own or were interested in

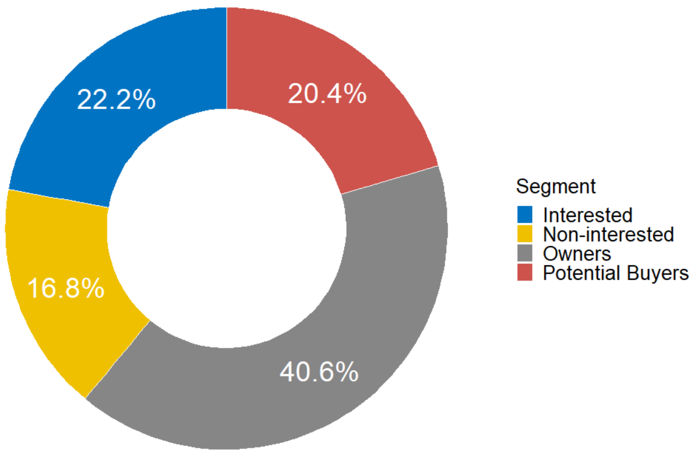
Table 9: Frequency table for the main demographic variables

Group	Frequency	Percentage
GENDER		
Male	421	35.92 %
Female	748	63.82 %
Other	3	0.26 %
Total	1172	100 %
AGE		
18-21	38	3,24 %
22-25	60	5,12 %
26-30	97	8,28 %
31-40	241	20,56 %
41-50	224	19,11 %
51-65	386	32,94 %
Over 65	126	10,75 %
Total	1172	100 %
EDUCATION		
Unfinished High school	21	1.79 %
High school graduate	309	26,37 %
Occupational, technical or vocational program	246	20,99 %
Bachelor's degree	363	30,97 %
Master's degree	181	15,44 %
Doctoral-PhD	41	3,50 %
Other	11	0,94 %
Total	1172	100 %

purchasing a smart speaker" to those that "know what a smart speaker is". This change meant that the quality of the data would slightly drop as owners or purchase intenders would provide a more thorough and careful assessment of the different smart speaker features, but on the other hand fieldwork would result drastically shortened and cheapened. Knowing this, 74 % of respondents passed the screening criteria and thus knew or were familiar with smart speakers. This figure was surprising in itself, as although I was aware that the American market was a pioneer in the use of this kind of technology and it is already widespread in many homes, the figure exceeded any possible expectation. Within that group of people that knew what a smart speaker is and thus passed the screening criteria, we could classify them into these four categories as seen in Figure 6, for the question *"What is your current situation regarding Smart Speakers"*.

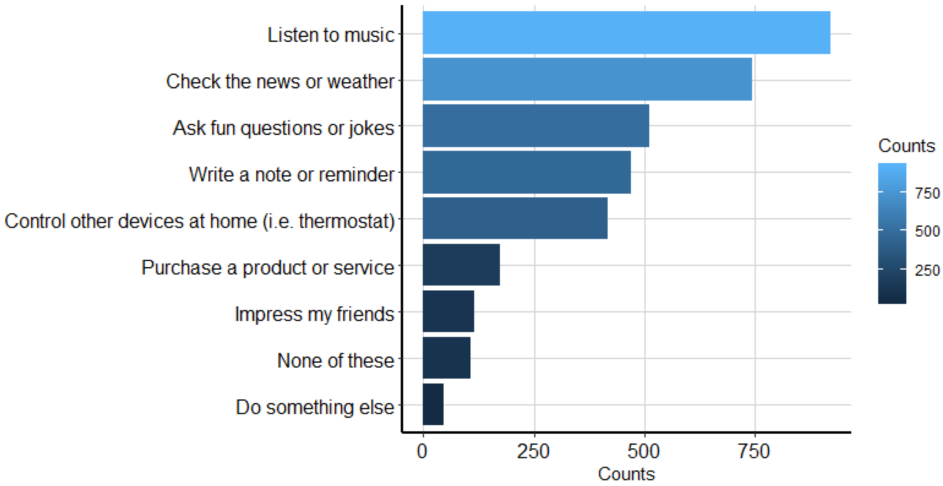
- Owners: I Already have one smart speaker at home (or more than one)
- Potential Buyers: I don't have one but I am highly interested in purchasing one
- Interested: I am somewhat interested in the product
- Non-interested: I am somewhat interested in the product

Figure 6: Half pie chart indicating sample proportion, in percent points, for each segment based on the interest on smart speakers



Before digging into the results of the study, it would be also meaningful to analyze the current consumer behavior when it comes to how they would a smart speaker on their day to day routines. For that, a multiple select question was presented, in which different common and basic uses were shown and respondents were asked to indicate which one they would use these devices for (for non-owners) or they already use them for (for owners of these gadgets). It is intriguing to observe, in Figure 7, that only two of these functionalities were selected by more than half of the sample, possibly indicating a potential lack of knowledge from the general public on the vast array of functionalities that smart speakers provides. This idea is reinforces when we observe that an almost negligent proportion of the respondents selected the option "do something else".

Figure 7: What would or do consumers use smart speakers for? Frequency bar chart for the multi-select option on different current uses for smart speakers

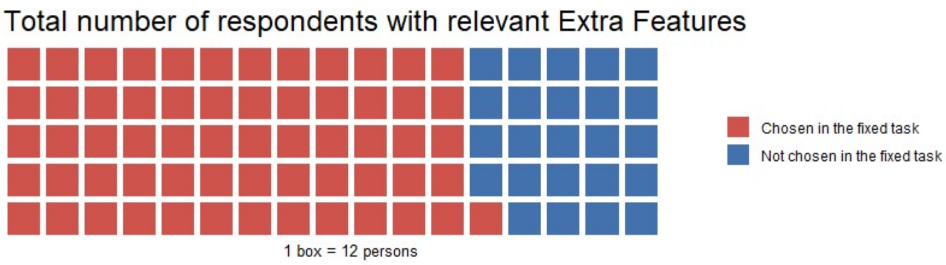


5 Analysis and Results

As reported throughout the study, there are 3 Research Questions and 3 hypotheses that needs to be brought to resolution. In this section, the results that would shed light on them will be presented at length without including interpretations, as they will be provided in the Discussion section. For the first Research Question, and before diving into and contrasting the predictive performance of both choice exercises, it would be insightful to analyze beforehand how well can the Kano Model capture or synthesize an individual's inclination towards a given set of features. As the AK-CBC heavily relies on the capacity of this model to truthfully reveal the preference of respondents in terms of identifying indifferent features and categorizing relevant ones, one important presumption of this model would be seriously compromised if that is not the case. For that, an appraisal was carried out on what proportion of the Extra Features that were categorized as Must-be, Performer or Attractive, and thus move forward in the AK-CBC to the final choice exercise, were freely chosen in the fixed task as their most preferred one. More than three quarters (75.68%) of respondents that deemed one or more of the Extra Features as relevant in the custom exercise (Must-Be, Performer or Attractive) indeed chose one of those as their most preferred ones in the fixed task in which they could only choose one. Nevertheless, the same exercise was conducted on the group that took the regular random-generated CBC in order to have the perspective of the full sample size (N = 1017), and the result dropped significantly to 67.22%. In Figure 8 we can see the average for the two samples, in which out of 1017 respondents that considered any of the Extra features as relevant, 71.19% chose one of these levels in the hold-out task.

As a last preliminary assessment in order to further test the conceptual validity of the Kano, and given that it has a qualitative nature and its categories carry intrinsic and distinguishable meaning, we compared on a respondent level and for each of the categories the proportion of each that was chosen in the fixed task. So, for each respondent that select one or more extra features as Must-be for example, how many chose one of those in the fixed task. It is specially relevant to see the results for this particular category as remember that Must-be features would bring extreme dissatisfaction if it is not included in the product configuration. Accordingly, the ratio for each category was Must-Be (50.66% with 116 out of 229), Performers (55.34% with 145 out of 262) and Delighters (29.85% with 100 out of 335).

Figure 8: Waffle Plot with the proportion of respondents who chose one in the fixed task of the Extra Features that they deemed relevant with the Kano exercise



As explained in the Data and Methodology section the performance of each exercise would be measured on the respective hit rate, namely whether the attribute level with the highest estimated part-worth utility was chosen in its respective hold-out task, in which the whole array of alternatives was available to choose one from. In Figure 9, we can see a visual representation that summarizes the results on the terms of this metric for the two different exercises and in Table 10 we can observe the relative differences respectively. It is noticeable that the AK-CBC improves the random-generated CBC in all 6 tasks presented, but most relevantly, the largest difference can be observed when a minimum preference share threshold is set (as observed in the last column of Table 10). This is a key factor to consider, since the AK-CBC removes uninformative and irrelevant feature levels from the final choice exercise, which is the one used for utility estimation, the expected part-worths for these levels would be greatly reduced, as the estimation would emanate exclusively from the Hierarchical Bayes upper-level. As a consequence, utility estimations and thus preference shares would theoretically polarize around the non-irrelevant feature levels that were included and would not be as widespread as with a the random-generated CBC that includes in the computation all levels. For that, it was set a minimum preference share threshold of 40% for both models in which given the case that it was not reached, that individual would not qualify for the hit rate calculation, as by setting a minimum bar would involve that only respondents with a stronger preference would be taken into account and the levels of uncertainty would be greatly reduced. As a note and reminder for the reader, the highlighted cells in light grey (Material and Size) represent those fixed task attributes that are comprised by less than 4 levels and from which none of those levels were removed.

Figure 9: Grouped Bar Chart plot representing th percentage of Hit Rates for the AK-CBC and a Randomly-generated CBC

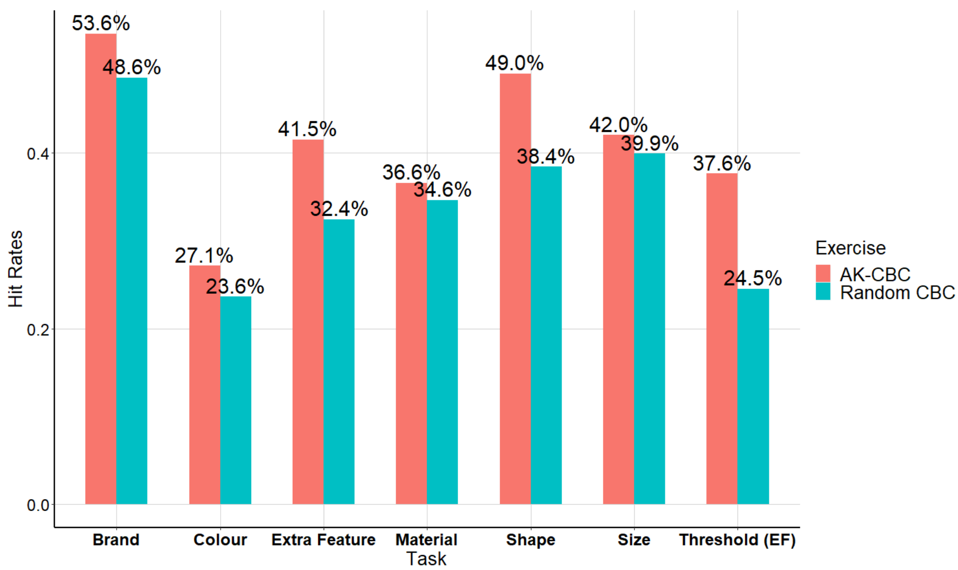


Table 10: Relative differences between the AK-CBC and a Randomly-generated CBC in terms of hit rate for each of the hold-out tasks

	Brand	Colour	Extra Feature	Material	Shape	Size	Threshold
Relative Difference	10.30%	14.89%	27.92%	5.76%	27.58%	5.25%	53.57%

These results will be discussed in detail in the Finding and discussion section, but before moving on to RQ2 it is important to point out that the evidence collected in Figure 9 and Table 10 strongly suggests that the hypothesis for this Research Question is strongly supported.

For elucidating Research Question 2 and its coupled hypothesis, the average time elapsed for both the Custom CBC and the Random CBC in each of the tasks is displayed in Figure 10, with the relative differences among both presented in Table 11. In order to collect the data, outliers were removed to make sure that they did not affect the results. In that sense, it was established that those that spent more than 300 seconds on a given task would qualify as outliers and hence would be removed, as I assumed that taking more than 5 minutes in a simple choice exercise task would mean that the respondent left the survey unattended momentarily. On that note, the ratio of outliers was very similar for both exercises (0,35% for the Custom and 0,36% for the Random).

Figure 10: Line plot with the elapsed time or duration in seconds for each of the tasks

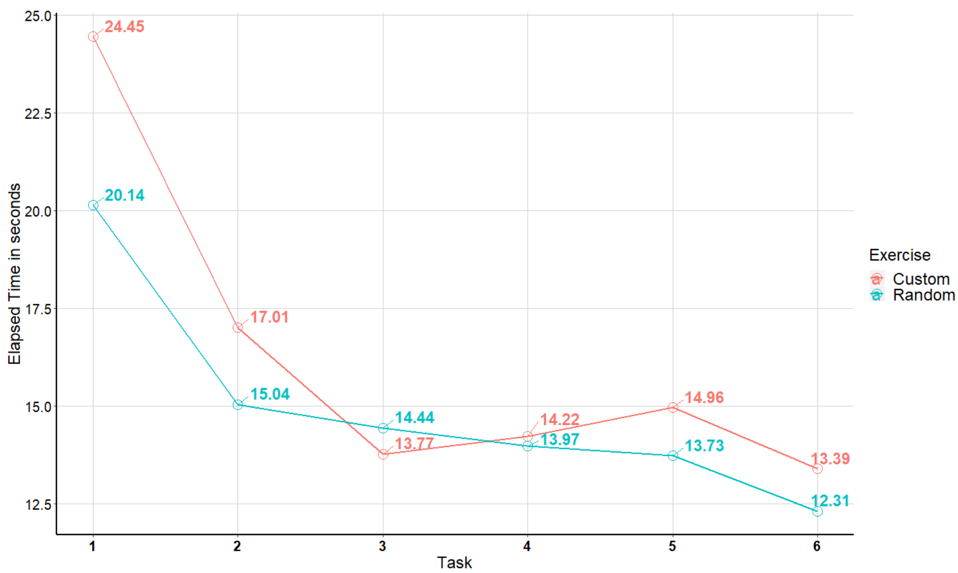


Table 11: Relative differences among the customized CBC design and the random-generated exercise in terms of elapsed time for each of the tasks

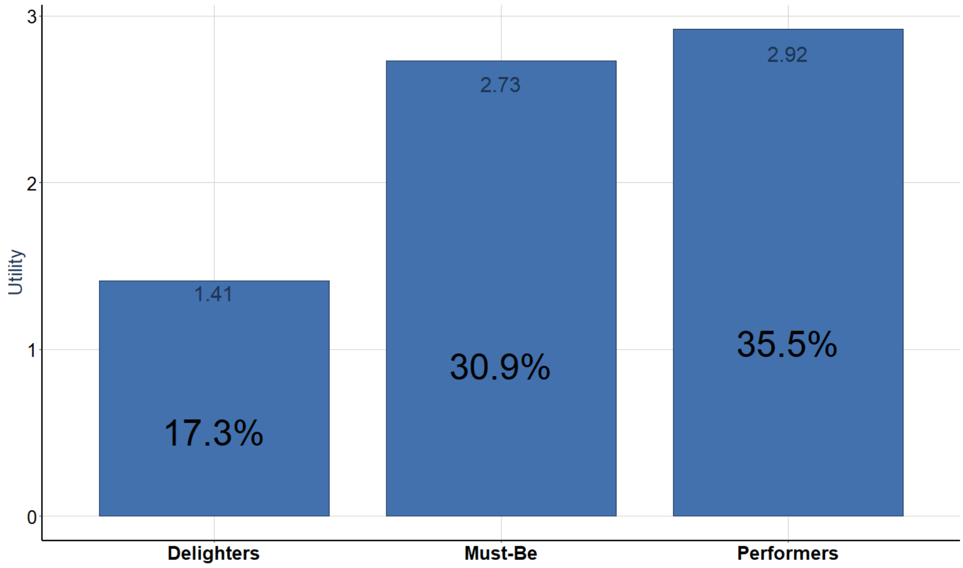
	Task 1	**Task 2**	**Task 4**	**Task 4**	**Task 5**	**Task 6**	**Total**
Relative Difference	21.36%	13.05%	-4.58%	1.80%	9.02%	8.76%	9.12%

The results gathered, underpinning the fact that users consistently spent more time in 5 out of the 6 tasks for the AK-CBC, also supports the hypothesis formulated. An in-depth analysis and its potential interpretations would be discussed in the next section.

Finally, for addressing the third Research Question, it is meaningful to mention that out of the 1172 respondents that completed the survey, 1017 considered as relevant for them at least one of the Extra features presented in the Kano exercise, which corresponds to 86.77% of the whole sample size. Within those, 2860, 3426 and 3426 total features across the whole population were categorized as Must-be, Performer and Delighters respectively with a relative proportion of 29.44%, 35.28% and 35.28% each. We already discussed in the preliminary analysis for RQ1 some metrics and results regarding how the Kano Model and its different categories performed with respect to the Extra feature hold-out task. For this Research Question I wanted to extend those outcomes by comparing how each category from the Kano performed in terms of preference scores and thus their part-worth utilities estimation. The end result would be to scrutinize which one plays a more important role in the decision-making and eventual choice, and for that a two-pronged approach was followed. Firstly, an overall average of utilities and preference shares would approximate at a population-level which of the categories scored higher, in Figure 11 we can observe the results in terms of relative preference share.

This first approach is useful at determining straightforwardly the overall relationship between preference weights and each of the categories respectively, but the fact that the number of features belonging to each category differs, may limit comparability to a certain degree. The second prong aims at assessing which category has a more

Figure 11: Bar Chart with the average utilities for each category (in dark blue) and their corresponding average preference share in bold-faced black



influential impact on preference, but this time at an individual level. For that, by only addressing those respondents that deemed at least one of the extra features as relevant in the AK-CBC (N = 471), I assessed which one was the category in which the feature or the sum of the features that belong to it has more than 45% preference share in sum for that particular individual. That cut-off point was set as there was not a single case with 2 categories over that level, and by setting it at a minimum of 50% many respondents would have not qualified. The results of these analyses are collected in Table 12. The last group "Other" correspond to those individuals that despite having selected one or more extra features as non-indifferent for them, the associated preference share did not reach the aforementioned 45% threshold.

Table 12: Proportion of each Kano category in which their integrating features add up to more than a 45% preference share score

Group	Frequency	Percentage
Must-be	127	26.96 %
Performer	157	33.33 %
Attractive	145	30.79 %
Other	42	8.92 %
Total	471	100 %

The results assembled in Figure 11 and Table 12 contradicts the hypothesis for this Research Question. Further discussion with follow in the next section.

6 Findings and Discussion

The objectives of this dissertation revolves around three main pillars as coinciding with the three different Research Questions. Accordingly, this section would also be structured in the same way, and the key findings will be summarized as per each RQ.

Model Performance

For RQ1, the main aim was to explore whether the new model would improve the performance of a regular random-generated CBC in terms of hit rates. As the AK-CBC uses a preceding CBC and the Kano Model to bring forward to the final choice exercise only those features and levels that are deemed relevant and appealing for each respondent, it was crucial to do a preliminary assessment on the validity of the Kano in the sense of whether it is actually capable of revealing the respondent's preferences truthfully or not. Although the results are susceptible to interpretation, the data reasonably supports this assumption as more than 7 out of 10 respondents included in both samples (N = 1017) chose one of the extra features in the fixed task which they previously categorized as Must-Be, Performer or Attractive in the Kano. In this sense, I considered this figure to be sufficient evidence that the Kano robustly exposes which features each respondent would like to be included in a smart speaker concept in a consistent way, although the minimum cut-off point could be open to discussion. On this behalf, it would also be interesting to analyze and deliberate whether the phrasing of the questions or the way the questionnaire was set up might have influenced the decision in the hold-out task or the Kano itself.

As a second preliminary evaluation before delving into the predictive power of the AK-CBC, and for continuing the line of exploring the Kano's reliability, I assessed on an individual basis what proportion of each Kano category was chosen in the final hold-out task. As a way of example and illustration, for each individual that chose one or more extra features as Must-Be for them (and the other two categories) I calculated the proportion that also chose one of those in the fixed-task against those who did not. Since, as we saw in the description of the Kano exercise in the Theoretical Framework, each of those categories carry a distinct connotation in terms of how it would impact their satisfaction if that feature is integrated in the product concept or not, I assumed that the Must-Be features would have a larger ratio given that their absence would bring extreme dissatisfaction. However, this seemed not to be the case, as only around half of those categories were later chosen in the fixed task, and the Performers or One-Dimensional showed a greater relative proportion with 55.34% of the cases.

These two analyses combined might imply, in my opinion, that the Kano exercise is a solid model to identify which features are relevant and which are not, but its blurry linguistics-based nature and the thin line that differentiates each of the categories, which is only based on two closed-form questions, might make the categorization process relatively fragile and weak in line with the conclusions from Mikulic & Prebezac (2011).

In terms of predictive performance, there are several key findings that are worth bringing up for discussion. In line with the stated hypothesis, the AK-CBC unmistakably outperformed the random-generated CBC in every single fixed task that was presented at the end of the survey, in which respondents could build their own smart speaker by freely choosing their preferred level among the whole array of alternatives. The data also strongly suggests that the biggest and most substantial differences are observed on the extra feature task in which the Kano Model is involved and thus reaffirms the notion that the model proves useful for screening relevant from non-relevant attributes.

Another crucial aspect to note is that the difference with the regular CBC skyrockets when a minimum preference share threshold is introduced in the hit rate estimation.

One of the takeaways or implications that this may exemplify is that the AK-CBC is especially valuable for identifying needs and preferences when some degree of certainty and confidence is required in the estimation and we aim to identify, for example, the desires of those respondents with stronger preferences and convictions.

Finally, it should not go unnoticed that the initial CBC, included in the algorithm to filter out unappealing levels from the core attributes, plays an important role also in terms of enhancing and boosting the predictive performance of the AK-CBC. This can be evidenced by observing in Table 10 the relative improvement for those tasks that had one of their levels removed (remember from Table 6 where it was indicated that only core attributes that had more than 3 levels would be manipulated) and which in this case was brand, colour and shape with 10.3%, 14.89% and 27.58% respectively, against those that were not adapted and were highlighted in light grey in the same table, which are material (5.76%) and size (5.25%). It needs to be taken into consideration that for the latter tasks, the same exact levels were shown in both exercises, and thus the differences are not due to any AK-CBC's fine-tuning. As a consequence, the data might suggest that a plausible explanation to those slight differences (around 5%) could be simply attributable to noisy baseline errors arising from the convergence of β's in the Hierarchical Bayes run. It is hence meaningful to benchmark those results to notice the improvement in the other tasks.

User Engagement

Along the dissertation, it has been commented at length how important it is for researchers to attract respondents' attention and maximize their involvement and engagement in the survey. This is specially relevant for companies, as a careful and thorough examination of all alternatives shown in the choice exercise provide high-quality and valuable data that can be derived into actionable marketing and managerial decisions. It is therefore that many experts in the field of Conjoint have deliberately sought for clever and innovative ways of bringing together methodological soundness and rigor with custom and adapted designs to make the choice alternatives as tailor-made and fitted to each respondents needs and desires as possible. In line with the literature consulted on the matter, and for this Research Question, time-elapsed per choice task was used as a proxy for user engagement.

In this direction, Hauser et al. (2009), empirically proved that when respondents use a task-completion mindset due to the non-relevance of the choice alternatives presented, or when presented with many features and levels, it is highly common that they might resort to the use of simplifying heuristics instead of the desired full-information compensatory decision process in which all included attributes in the product concept are carefully evaluated. On the basis of these notions, the use of cognitive-simple heuristics would unarguably lead to a shorter time elapsed in each of the choice exercises, so with the development of the AK-CBC I aimed at testing if this metric would also be boosted.

The results clearly indicate that the hypothesis is supported, as we can observe at first sight in Table 11 and Figure 10, there is a slight yet stable and consistent improvement in 5 out of 6 tasks. Although the relative differences are not substantial, each respondent spent on average more than 8 seconds in the AK-CBC final choice exercise than in the random-generated one which exhibits a clear pattern once the outliers have been dismissed.

Kano categories preferences

For this last Research Question the goal was to address the correlation between each Kano category and their part-worth estimates. That would allow to investigate what potential connection or relationship could be between the qualitative definitions of each of those categories to their associated preferences, as derived from their utility estimations. In line with the preliminary analysis conducted for RQ1, I hypothesized that Must-Be features should not only have the highest average utility but they should be consistently chosen in the fixed task. Thin conjecture arises from the intrinsic definition of this attribute type, in which its integration in the product configuration is crucial as its absence would produce extreme dissatisfaction. We already observed in the discussion for the first Research Question that this assumption does not hold when we only take into consideration the results for the fixed task. On that same line, and contrary to the hypothesized association, when we average the utilities for each of the Kano categories across all respondents, we can assert that this assumption is not supported either and there is no such connection. On this behalf, and as suggested in the data displayed in Figure 11, One-Dimensional or Performer Attributes outperform again Must-Be features also in terms of average utility.

Because the differing number of features in each category might influence and limit the interpretability of these results I carried out a separate but complementary analysis with the ultimate goal of finding out which one had a stronger impact on overall preference. For that, we assess which category had a larger fraction or magnitude of their included features adding up to a preference share of 45% or more out of the total population of respondents that regarded at least of the extra features as relevant. Again, the results contradicted the main hypothesis in favour of Performer attributes, and even Attractive features also scored higher in this case as can be observed in Table 12. Besides the argument that these attributes might be indeed more preferred than the other two, a plausible reason for this pattern might also be explained on the way the Kano is set up, given that as seen in the Kano Evaluation Table in Table 3, the only combination of responses for the Performer (I like it that way for the functional question and I dislike it that way for the dysfunctional one) might conveniently express the most straightforward and simple responses to be provided in a questionnaire of this kind, although the exact reasons for that are yet to be identified.

6.1 Managerial implications

There are several key ideas that can be extracted from the results and findings of this work and that can be translated beyond the purely academical theory to more actionable business decisions that might be of special relevance for managers and other decision-makers.

First and foremost, the data obtained from this study strongly suggests that creating custom and adapted designs for the Conjoint exercises using the Kano Model not only boosts respondents' engagement in the different tasks but also reflects compelling improvements in terms of identifying customer preferences and needs more accurately. These findings go in line with the theory consulted on the matter in which the search for further adaptiveness, such as the ACBC, yields tangible betterment and underpins the need for Market Research agencies to unceasingly revise the traditional methods used in the industry towards new advanced and innovative mechanisms. This would not only be advantageous for the agency itself as it would allow them to provide more accurate recommendations and insights, but also for the final clients that would be able to translate this data into actionable and profitable business decisions.

Secondly, and directly linked with the first point, productionalizing the AK-CBC would not have been possible without a heavy investment in time and knowledge for new DevOps tools. It has not been detailed at length since it escapes the scope and goals of this dissertation, but for making the AK-CBC a reality a great deal of technical expertise was inputted through Amazon Web Services Lambda to deploy the AK-CBC R script into a production environment such as the server of the survey software provider. In that sense, the future of Machine Learning and Big Data is rapidly changing and is also moving beyond a seemingly abstract idea with unlimited possibilities to a more actionable and automated integration of algorithms and deploying them into actual production environments. The companies that do not swiftly adapt to these new business conditions would definitely have a hard time providing solutions that are on par with the ones that rely on the true power of data.

Finally, and independent of the Research objectives defined for this study, it would be of high relevance to make a reference on the actual results of the Conjoint Exercise. On this note, during the brainstorming stage for deciding which Extra Features could be included in the study, I decided to incorporate "Guarantee of Confidentiality" in light of the latest events in which big technological players in the sector, like Apple or Google, allegedly used their models of smart speakers to collect information from their clients. In this study, that particular feature was the most relevant one in the Kano Exercise with 53.32% of respondents choosing it as either Must-Be, Performer or Attractive but also the preferred one by a great distance in the fixed task in which respondents had to choose their favourite Extra Feature, as it was chosen in 38.57% of the cases. These results only reaffirms the current concerns and sensitivity on the matter and stresses the importance that big companies respect their data security policies given the importance that the average user allocates to confidentiality these days.

7 Conclusions

The AK-CBC was created with the goal of becoming an important extension in the field of Adaptive Conjoint Analysis Designs. With a novel algorithm that uses the Kano Model to filter out non-relevant attributes for the final choice exercise on the fly I aimed at both testing its performance in terms of predictive power and user engagement against a random generated design with no adaptation whatsoever, as well as providing a broader understanding on how the different feature categories of the model affect preferences.

The results were encouraging from a two-pronged perspective. From the technical side, the integration of the conceptual algorithm into a fully-functional production environment like the survey software platform was a complete success. In concordance with the proposed model, the script satisfactorily compiled the data from the initial CBC and Kano exercises in order to deliver an adapted final Conjoint design on the fly. On that sense, I can confidently say that the model deployment was an accomplishment in itself and I am truly optimistic on how it can contribute and trigger further research works on this domain. Secondly, regarding the research objectives, the results obtained in this study were in line with the literature consulted and strongly confirmed and substantiated the two initial hypothesis. A preliminary assessment reasonably supported the idea that the Kano Model is capable of identifying relevant features and thus serve as a pillar for the AK-CBC. In spite of this, it yielded more ambiguous results when it comes to its internal categorization, as features categorized as Must-be were not predominantly chosen in a fixed task where respondents could choose their single most proffered extra feature, hence contradicting the model's internal logic somewhat. Regardless of this conclusions, the data clearly suggests that the AK-CBC outperformed the regular random-generated CBC both in predictive performance, with a clear improvement in hit rates, and the user engagement, as respondents consistently spend more time in the tasks with the new methodology.

On a final note, an appraisal was conducted with the aim of having a broader understanding on which of the Kano categories has a higher preference rate among respondents, in terms of their utility part-worths. As the Kano has a linguistic-origin and qualitative nature, I deemed insightful for both academia and industry to contrast the conceptual categories with actual preference estimates. On this behalf, the results contradicted the hypothesis, as Performer features scored a larger average utility than Must-be and similar conclusions could be extracted when we analyze the proportion of each category that summed more than a minimum preference threshold, in which both Performer and Attractive features scored higher than Must-be ones respectively

8 Limitations and Further Research

Scope

During the proposal stage, the scope of this study was clearly bounded to a specific and limited domain. In the sense that the performance of the new model AK-CBC was only contrasted against a regular random-generated CBC. On this behalf, this line of research can be extended by comparing the AK-CBC with additional models, such as the ACBC (which was discussed in previous sections) or the newly-developed

Preference-Based Conjoint Analysis (PBC) which was out of the boundaries of this project but follows the same research line of maximizing adaptiveness. With this, more light could be shed on which model is more powerful and can perform better in understanding consumers decision-making and preferences.

Literature Review

One key element that the reader needs to take into consideration is that Sawtooth Software is not only the largest provider of Conjoint-related surveys and other complementary tools like Hierarchical Bayes Software, but it is also the provider for the Market Research agency with which I have collaborated for this study (SKIM) and has one of the most important repositories in terms of technical papers and other well-suited specialized articles. On that sense, the theoretical framework for the Conjoint sections was primarily sourced on these technical papers and enriched and complemented with other methodological, psychological and marketing-related academical work. In light of this, plus the fact that other external sources consulted for Conjoint material resulted out-dated, irrelevant for this work or simply incomplete, might lead to contents being one-sided or too suited to this particular software procedures and idiosyncrasies and thus weakening generalizability of this section.

Methodology

In order to provide clarification for RQ2, user engagement was measured on the time elapsed/spent on the different tasks of the choice exercise. This could be a robust approximation but undoubtedly results in such a broad metric that many factors are not taken into account. For this study it resulted unfeasible from both a technical and budgetary perspective, but the development of technologies like wearable eye-tracking googles that allow to create heat maps which points out where the user is focusing on would allow to make a stronger and more reliable assessment on the level of user engagement in the survey if used complementarily to the elapsed time.

Data Collection

It is also pertinent to discuss the sample profile used for the survey. For the data collection process, the original idea was to address respondents that either owned or had an active intend on purchasing a smart speaker. Due to time and budgetary constraints, as accessing this particular pool of individuals would be more complex, it was decided to disregard this idea and instead seek for individuals that knew what a smart speaker was. This approach swift might have lead to a slight decrease in data-quality, as disinterest or unfamiliarity on the different features included in the choice task or the product itself might compel less careful examination of the product alternatives. On this same line, it is also important to stress that the results of this dissertation should only be regarded as a market environment simulation, but it cannot be taken as ground truth, as individuals might react differently in a real-life scenario than in a computer-administered survey.

Technical limitations

One of the most important limitations that I encountered was during the functional application of the script in Amazon Web Services Lambda. As it is still under development, the cloud platform only allows a constrained number of packages to be used, as there is restricted space available. This resulted in an important hurdle to overcome at the time of building efficient and smooth code. On the bright side, once this space restriction is unlocked, a new array of opportunities to fine-tune and refine Conjoint designs will be opened.

Other

Finally, for answering RQ1, an initial assessment of the Kano performance to detect relevant features was conducted. For the sample that took the AK-CBC more than three quarters of the sample chose in the fixed task one of the features that they deemed non-irrelevant in the Kano, but for the group that made the other exercise, this figure drastically falls to. It would be interesting to carry out a follow-up analysis on the reasons that may have led to this situation, as it cannot be explained by mere sampling bias at the groups were assigned on a straight random fashion.

9 Bibliography

- Asioli, D., Næs, T., Granli, B. S., & Lengard Almli, V. (2014). Consumer preferences for iced coffee determined by conjoint analysis: An exploratory study with N orwegian consumers. International journal of food science & technology, 49(6), 1565-1571.

- Baier, D., Pełka, M., Rybicka, A., & Schreiber, S. (2015). Ratings-/Rankings-Based Versus Choice-Based Conjoint Analysis for Predicting Choices (pp. 205–216).

- Ben-Akiva, M., Mcfadden, D., & Train, K. (2019). Foundations of Stated Preference Elicitation: Consumer Behavior and Choice-based Conjoint Analysis, 10(2), 1–144. https://doi.org/10.1561/0800000036

- Chapman, C. N., Alford, J. L., Johnson, C., Lahav, M. & Weidemann, R. CBC vs ACBC: Comparing results with real product selection. Reseach paper series. Sawtooth Software.

- Chen, C. C., & Chuang, M. C. (2008). Integrating the Kano model into a robust design approach to enhance customer satisfaction with product design. International journal of production economics, 114(2), 667-681.

- Cunningham, C. E., Deal, K., & Chen, Y. (2010). Adaptive Choice-Based Conjoint Analysis. The Patient: Patient-Centered Outcomes Research, 3(4), 257–273. https://doi.org/10.2165/11537870-000000000-00000

- Dobson, G., & Kalish, S. (1993). Heuristics for pricing and positioning a product-line using. Retrieved from https://search.proquest.com/docview/213232386

- Green, P. E.,& Rao, V. R. (1971). Conjoint Measurement for Quantifying Judgmental Data. Journal of Marketing Research, 8(3), 355. https://doi.org/10.2307/3149575

- Griffin, A., & Hauser, J. R. (1993). The voice of the customer. Marketing science, 12(1), 1-27.

- Hauser, J., Ding, M., Sawtooth, S. G. Proceedings of the Sawtooth Software Conference (2009), undefined. (n.d.). Non-compensatory (and compensatory) models of consideration-set decisions. Academia.Edu.

- Hauser, J., Toubia, O., Evgeniou, T., Befurt, R. & Daria Silinskaia, D. (2009), Cognitive Simplicity and Consideration Sets, forthcoming, Journal of Marketing Research.

- Herzberg, F. (1959). The motivation to work. New York: Wiley.

- Ishwaran, H., & Zarepour, M. (2000). Markov chain Monte Carlo in approximate Dirichlet and beta two-parameter process hierarchical models. Biometrika, 87(2), 371-390.

- Jervis, S. M., Ennis, J. M., & Drake, M. A. (2012). A Comparison of Adaptive Choice-Based Conjoint and Choice-Based Conjoint to Determine Key Choice Attributes of Sour Cream with Limited Sample Size. Journal of Sensory Studies, 27(6), 451–462. https://doi.org/10.1111/joss.12009

- Lee, Y. C., Sheu, L. C., & Tsou, Y. G. (2008). Quality function deployment implementation based on Fuzzy Kano model: An application in PLM system. Computers & Industrial Engineering, 55(1), 48-63.

- Lenk, P. J., DeSarbo, W. S., Green, P. E., & Young, M. R. (1996). Hierarchical Bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. Marketing Science, 15(2), 173-191.

- Martinez de la Torre, A. (Adrian). (2017, October 16). Conjoint Design Generation On-The-Fly. Econometrie. Retrieved from http://hdl.handle.net/2105/39744

- Maslow, A. H. (1943). A theory of human motivation. Psychological Review, 50(4), 370–396.

- McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. University of California at Berkeley.

- Meißner, M. & Steiner, M (2018). A User's Guide to the Galaxy of Conjoint Analysis and Compositional Preference Measurements. Faculty. Marketing ZFP.

- Mikulić, J., & Prebežac, D. (2011). A critical review of techniques for classifying quality attributes in the Kano model. Managing Service Quality: An International Journal, 21(1), 46-66.

- Min, S. H., Kim, H. Y., Kwon, Y. J., & Sohn, S. Y. (2011). Conjoint analysis for improving the e-book reader in the Korean market. Expert Systems with Applications, 38(10), 12923-12929.

- Orme, B., & Baker, G. (2000). Comparing Hierarchical Bayes Draws and Randomized First Choice for Conjoint Simulations. Sawtooth Software Research Paper Series.

- Orme, B. (2009). The CBC/HB system for hierarchical Bayes estimation version 5.0 Technical Paper. Technical Paper Series, Sawtooth Software, Orem, UT.

- Orme, B. K. (2009). The Adaptive Choice-based Conjoint (ACBC) Technical Paper. Sawtooth Software, Sequim, WA.

- Orme, B., & Software, S. (1996). Article originally published in Sawtooth Solutions. Retrieved from www.sawtoothsoftware.com

- Safizadeh, M. H. (1989). The Internal Validity of the Trade-Off Method of Conjoint Analysis. Decision Sciences, 20(3), 451–461. https://doi.org/10.1111/j.1540-5915.1989.tb01560.x

- Sauerwein, E., & Bailom, F. (1996). The Kano model: How to delight your customers. Faculty.Kfupm.Edu.Sa.

- Shahin, A. (2004). Integration of FMEA and the Kano model: An exploratory examination. International Journal of Quality & Reliability Management, 21(7), 731-746.

- Suzianti, A., Faradilla, N. D. P., & Anjani, S. (2015). Customer preference analysis on fashion online shops using the Kano model and conjoint analysis. International Journal of Technology, 6(5), 881-885.

- Tan, K. C., & Pawitra, T. A. (2001). Integrating SERVQUAL and Kano's model into QFD for service excellence development. Managing Service Quality: An International Journal, 11(6), 418-430.

- Tan, K. C., & Shen, X. X. (2000). Integrating Kano's model in the planning matrix of quality function deployment. Total quality management, 11(8), 1141-1151.

- Tontini, G. (2007). Integrating the Kano Model and QFD for Designing New Products. Total Quality Management & Business Excellence, 18(6), 599–612. https://doi.org/10.1080/14783360701349351

- von Dran, G., Zhang, P., & Small, R. (1999). Quality websites: An application of the Kano model to website design. AMCIS 1999 proceedings, 314.

- Wang, C. H., & Wang, J. (2014). Combining fuzzy AHP and fuzzy Kano to optimize product varieties for smart cameras: A zero-one integer programming perspective. Applied Soft Computing, 22, 410-416.

- Wang, C. H., & Wu, C. W. (2014). Combining conjoint analysis with Kano model to optimize product varieties of smart phones: A VIKOR perspective. Journal of Industrial and Production Engineering, 31(4), 177-186.

- Wittink, D., Huber, J., Zandan, P., Johnson, R. M., & Wittink, D. R. (1992). The Number of Levels Effect in Conjoint: Where Does It Come From and Can It Be Eliminated? Retrieved from www.sawtoothsoftware.com

- Xu, Q., Jiao, R., Yang, X., & Helander, M., (2009). An analytical Kano model for customer need analysis. Elsevier. Retrieved from https://www.sciencedirect.com

- Zwerina, K., Huber, J., & Kuhfeld, W. F. (1996). A general method for constructing efficient choice designs. Durham, NC: Fuqua School of Business, Duke University.