UNIVERSIDADE DE SÃO PAULO

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE

PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

ESCOLHAS BASEADAS EM MÚLTIPLOS OBJETIVOS:

a evolução do homo economicus ao homo aptabilis

MULTIPLE GOALS-BASED CHOICE:

the evolution from *homo economicus* to *homo aptabilis*

Luis Eduardo Pilli

Supervisor: Prof. Dr. José Afonso Mazzon

Co-Supervisor: Prof. Dr. Joffre Swait

SÃO PAULO

2017

LUIS EDUARDO PILLI

MULTIPLE GOALS-BASED CHOICE:
the evolution from *homo economicus* to *homo aptabilis*

Tese apresentada ao Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, como parte dos requisitos para obtenção do grau de Doutor em Ciências.

Área de Concentração: Marketing

Orientador: Prof. Dr. José Afonso Mazzon

Co-Orientador: Prof. Dr. Joffre Swait

Versão Corrigida

(versão original disponível na Biblioteca da Faculdade de Economia, Administração e Contabilidade)

SÃO PAULO

2017

Catalogação na publicação

Serviço de Biblioteca e Documentação

Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo

FICHA CATALOGRÁFICA

Elaborada pela Seção de Processamento Técnico do SBD/FEA/USP

Pilli, Luis Eduardo
Multiple goals based choice: the evolution from *homo economicus* to *homo aptabilis* / Luis Eduardo Pilli. -- São Paulo, 2017.
183 p.

Tese (Doutorado) – Universidade de São Paulo, 2017.
Orientador: José Afonso Mazzon.

1. Comportamento do consumidor 2. Psicologia do consumidor 3. Marketing 4. Econometria I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

CDD – 658.8342

DEDICATÓRIA

Aos meus filhos: Lucas, Gabriel e Gustavo,
Que nossos caminhos sejam feitos de boas escolhas.

Aos meus pais: Waldyr e Lucy,
Por acreditarem e apoiarem, sempre e de novo.

À Flavia, meu amor.
Pelo caminho, pelo carinho, pelo amor.

AGRADECIMENTOS

Agradeço à FAPESP?CAPES (através dos projetos de números 2015/03862-9 e 2015/21771-0), a USP e ao CNPq. Tive a oportunidade de passar os últimos anos estudando em uma da melhores universidades do Brasil, contando com bolsas de estudo, inclusive no exterior. Agradeço a todos os que me deram a oportunidade de me beneficiar do ensino superior público e gratuito no Brasil, infelizmente uma contradição concentradora de renda. Espero que que tenha me feito merecedor e espero que o conhecimento que gerei seja usado para o benefício dos que tornaram possível essa experiência. Evidentemente, as opiniões, hipóteses e conclusões ou recomendações expressas ao longo desta tese são de minha responsabilidade e não refletem, necessariamente, a visão da FAPESP ou do CNPq. Da mesma forma, as pessoas cito a seguir contribuíram decisivamente para a conclusão dessa tese. Com todas eu divido os méritos desse projeto, lembrando que eventuais falhas ou omissões são de minha responsabilidade.

Tive a fortuna de contar com dois orientadores fora de série. Prof. José Afonso Mazzon que me acolheu desde o mestrado, abriu as portas da vida acadêmica, me estimulou, compartilhou projetos e me deu liberdade para fazer minhas escolhas. Encontrei o Prof. Joffre Swait na reta final do doutorado, quando minhas escolhas me colocaram diante de problemas maiores que meus conhecimentos. Pesquisador de classe mundial, me acolheu como orientando no período de estágio de pesquisa no exterior, me ofereceu recursos do *Institute for Choice* da *University of South Australia* e me acompanhou até a entrega da tese, após minha volta ao Brasil; os benefícios de sua orientação são decisivos para a conclusão desse projeto e, sobretudo, para o desenvolvimento dos próximos. Não teria o prazer de chegar ao final desse projeto de vida sem a participação dos dois. A fortuna está em encontrar dois grandes mestres que têm em comum a generosidade; me ofereceram o caminho, o conhecimento e a paciência.

Enorme é a gratidão aos meus pais, Waldyr e Lucy, sempre acreditando, sempre incentivando e sempre oferecendo todo o apoio, inclusive logístico, para que este sonho pudesse ser realizado.

I want to say thanks to all my colleagues, and the visiting researchers that I had the opportunity to meet, for supporting me during the eight amazing months that I spent in Sydney at the Institute for Choice. Everyone always willing to help either with my academic or my *personal necessities. Elisabeth Huynh and Subodh Dubey have contributed decisively to the project that resulted in the 4th Chapter of this thesis; Akshay Vij was always available for talking about ideas, obstacles and possible solutions for my research; Frances Nolan, Karen Cong and Elisabeth (again) for the support after my small accident; Maria Lambides for the support since the arrangements for this journey; Ali Ardeshiri, Habtamu Kassahun, Flavio Freire Souza, Aiste Ruseckaite, Greg Caire, Spring Sampson, Jun Zhang for sharing the path.*

As disciplinas de meu doutorado foram cumpridas na FEA, na ECA, no ICB, no IB e na FGV. Tive a oportunidade de compartilhar a sala de aula com professores e pesquisadores com as mais diversas origens e experiência, o que enriqueceu minha visão como pesquisador e ser humano. À todos que estiveram comigo nesses anos, meu muito obrigado. Em especial, o Prof. Rafael X. V. Ferreira pelos conselhos que permitiram a conclusão do Capítulo 4 dessa tese e ao Prof. Carlos Alberto de Braganca Pereira pela sua constante atenção, disponibilidade e bom humor para me transmitir um pouco de seu vasto conhecimento. Para completar, obrigado Prof. Eduardo Andrade por se unir aos quatro professores já mencionados na composição da banca examinadora de minha tese.

Nesses tempos em que redes sociais multiplicam as oportunidades de aprendizagem, os novos amigos virtuais contribuíram para meu trabalho. Diogo Leite meu ajudou com orientações sobre os tempos verbais do latim e Cesar Cesarius me deu as referências e a correta conjugação do *homo aptabilis*, tornando preciso conceito de indivíduo fazendo escolhas adotado nessa tese. Anna Cristina Figueiredo foi a grande amiga analógica me colocou na direção das pessoas certas. Obrigado aos três.

Me parece que pais que embarcam na viagem do doutorado, levam junto os filhos. Nos momentos de ausência, nas conversas sobre o tema da pesquisa, na loucura dos prazos... Obrigado ao Lucas, que também me ajudou com a revisão de grande parte do texto que vem a seguir, ao Gabriel e ao Gustavo, parceiro de viagem para Sydney. Obrigado por serem parte do caminho.

Finalmente, mas não menos importante, obrigado Flavia, meu amor. Por me encontrar no caminho, por compartilhar o caminho, o projeto e a viagem, mesmo que cada um de um lado do mundo. Com você ao meu foi mais fácil e mais divertido.

RESUMO

Escolhas são meios para que indivíduos e consumidores atinjam seus objetivos. São objeto de estudo em diversas disciplinas e eu me apoiei em três delas para desenvolver estas tese. Modelos normativos da economia que definem o *homo economicus*, modelos descritivos das teorias de decisão comportamental que forjam o *homo aptabilis*, capaz de fazer escolhas que permitam o atingimento de seus objetivos. E modelos econométricos de escolha discreta, que permitem o uso e teste de teorias comportamentais usando métodos flexíveis e realistas. O argumento central desta tese é que modelos econométricos devem considerar a heterogeneidade do comportamento individual em todo o processo de escolha, incluindo objetivos do consumidor, estratégias de decisão, formação de conjuntos de escolha subjetivos, além de preferência. O não reconhecimento desta complexidade nos processos de escolha produz modelos falsos, capturando a heterogeneidade no nível das preferências e induzindo organizações a tomarem decisões equivocadas. Para desenvolver este argumento, esta tese se organiza em três seções. Na primeira eu faço uma revisão da literatura com foco nos diversos níveis do processo de escolha onde a heterogeneidade se manifesta e relaciono os modelos de escolha com as teorias comportamentais de decisão. Na segunda seção é desenvolvido um estudo sobre os vieses provocados quando a heterogeneidade na formação de conjuntos subjetivos de escolha não é levada em consideração. Através de experimentos de Monte Carlo fica comprovado que os parâmetros de modelos econométricos de escolha são generalizadamente viesados, levando à estimadas equivocadas de probabilidades de escolhas das marcas e das elasticidades das probabilidades de escolha. Esses resultados são usados para motivar uma abordagem de teoria dos jogos que resulta em equilíbrio distante do ideal, do ponto de vista de resultados das empresas. Finalmente na terceira seção, é desenvolvido um modelo de escolha discreta baseado em múltiplos meta-objetivos e em diferentes processos de escolha individual. Mais um experimento de Monte Carlo comprova que o modelo é capaz de recuperar o parâmetros do processo gerador de dados. O modelo além de reconhecer a existência de diversos meta-objetivos que ativam diferentes regras comportamentais, também permite estudar a adaptação do processo de escolha individual em função de variáveis de contexto, de situação e individuais. O modelo articula modelos econométricos com teorias comportamentais de decisão e oferece suporte para a compreensão do *homo aptabilis*.

Palavras-chave: Comportamento do consumidor; Modelos de escolha discreta; Escolhas baseadas em múltiplos objetivos; Teorias comportamentais de decisão; Heterogeneidade na resposta do consumidor

ABSTRACT

Choices are the means for individuals and consumers to attain their goals. They are the objects of study for several disciplines and I relied on three of them to develop this thesis. Normative models from economics defining the *homo economicus*, descriptive models from behavioral decision theories that forge the *homo aptabilis*, able to adaptively pursue multiple goals through choices. And discrete-choice econometric models that allow the use and testing of behavioral theories using flexible and realistic methods. The central argument of this thesis is that econometric models should consider the heterogeneity of individual behavior throughout the choice process, including consumer goals, decision strategies, choice set formation, and preferences. Failure to recognize this complexity in the choice process produces false models, capturing process heterogeneity at the level of preferences, and inducing organizations to make the wrong decisions. To develop this argument, this thesis is organized into three sections. In the first one, I review the literature focusing on the different levels of the choice process where the heterogeneity manifests itself and I relate the choice models to the behavioral decision theories. In the second section, a study is developed on the biases caused when heterogeneity in the choice set formation is not accounted for. Through Monte Carlo experiments it has been proven that the parameters of econometric choice models are generally biased, leading to misleading estimates of brands choice probabilities and of attribute's choice elasticities. These results are used to motivate a game theoretical approach that results in far-fetched equilibrium from the point of view of business results. Finally, in the third section, a discrete choice model based on multiple meta-goals and on different individual choice processes is developed. One more Monte Carlo experiment proves that the model is capable of retrieving the parameters of the data-generating process. The model, besides recognizing the existence of several meta-objectives that activate different behavioral rules, also allows studying the adaptation of the consumer choice process as a response to context, situation, and individual variables. The model articulates econometric models with behavioral decision theories and supports the understanding of the homo *aptabilis*.

Keywords: Consumer behavior; Discrete choice models; Multiple goals choice bases; Behavioral Decision Theories, Consumer response heterogeneity.

LIST OF FIGURES

LIST OF TABLES

SUMMARY

1 THESIS PRESENTATION

The object of study chosen for this thesis is consumer response heterogeneity, specifically its integration into discrete choice models.

Consumer response heterogeneity is the outcome of individual differences revealed throughout judgment and decision-making and the psychological processes involved in this kind of human activity (Desarbo et al., 1997). It is a fundamental concept to marketing strategy, supporting segmentation, targeting and positioning decisions, as well as to operational marketing given its importance for marketing mix management (Kamakura, Kim, & Lee, 1996). To make my point clearer, before deepening into more precise definitions, heterogeneity may rest on consumers' tastes as it is the dominant practice in choice modelling or on, what I am loosely naming by now, the choice process, i.e., everything related to decision-making including tastes.

In the study of consumers' choices, Adamowicz et al., (2008) identify three relevant schools of thought that emphasize the understanding of consumers' decision-making processes. These three different perspectives, that offer the building blocks for developing this thesis are: the economic theory of consumer behavior; the behavioral decision theories approach that comprehends fields like consumer behavior, mathematical as well as cognitive and consumer psychology and, more recently, behavioral economics; and the choice modeling stream, which is concerned about the development of econometric models of choice used in a variety disciplines like marketing, applied economics, transportation, and sociology to name a few.

I will rely on these three fonts of scientific knowledge to illustrate the sources of choice heterogeneity, i.e., how the consumers' decision-making processes vary between persons as a function of individual differences, and across occasions (within persons) as a function of the environment. As it will be reasoned, I will stand at the side of researchers proposing that consumers' choice is a mean to achieve multiple goals, which are important not only as representations of desired end states but also because they drive the choice process focusing the superior psychological processes, like attention and memory, toward their achievement (Weber & Johnson, 2009). Oriented by these multiple goals the individual decision-making is a process comprising both the selection of an available alternative in any choice task, and also of a strategy that commands the efforts allocated to achieve the goals, i.e., a process of deciding how to decide (Swait & Feinberg, 2014), which I will refer to as a meta choice. The meta choice includes decisions about how to handle the context and task properties, i.e. goals activation and evaluation strategy, to select an alternative that best satisfies the consumer's goals.

Under the premise that choice is a multiple goal pursuing process, enabling a meta choice, the economic rationality gives place to a procedural rationality defined by H. A. Simon (1978 p. 9) as "the effectiveness, in light of human cognitive powers and limitations, of the procedures used to choose actions". And aligned with this view of rationality the *homo economicus* gives place to a decision-maker that I name as the *homo aptabilis*, given its ability to continuously adapt both the meta choice and the choice to the environment in the search for multiples goals.

Back to consumer heterogeneity response, it emerges in its many dimensions as a consequence of consumers striving to achieve multiple goals , leading to the use of different decision rules or strategies, to the manifestation of different preferences and to the observation of different levels of stochasticity throughout the choice process.

It is quite a challenge to organize the knowledge generated to explain the same phenomena as belonging to one source or to the other, among the three above-mentioned. As time goes by, the different disciplines get confronted and inspired by each other, and the relevant findings end up crossing the borders. Nevertheless, I will try to keep the contribution of each discipline as transparent as possible, and my objective is just to use their core perspectives on human decision-making as the vertices of a triangle lending the references about the territory. The economic theory of consumer behavior occupies the vertex offering the normative viewpoint of the phenomenon, with great focus on consumers' tastes; the behavioral decision theories reside in the vertex concerned about psychological processes involved and descriptive theories that challenge the normative view, with emphasis on the choice processes; and the econometric models of choice stream rests in the vertex that develops statistical tools building stochastic models that can empirically combine and test the propositions from the other vertices, to realistically explain and predict human choices.

Econometric models of choice, specifically discrete choice, receive a particular emphasis in this thesis development since this is the tool that I have chosen to shape my empirical contribution to the state of the art. This choice of mine rests on some good reasons: (i) these models are flexible enough to study several sources of data, like revealed or stated preference, experimental or non-experimental data in different aggregation levels and even to formally fuse data from different sources; (ii) this school of though is strongly concerned about prediction, but it also considers increasingly important to offer realistic behavioral explanations to support its predictions, and (iii) although choice process heterogeneity has been debated in the literature for quite a while, only in the past few years, efforts to model the phenomenon are slowly spreading across the discipline, meaning there's lot of room for contribution on extending the knowledge's frontier in the area.

Resulting from my motivation, when working through examples I will focus my efforts on identifying the econometric models of choice that illustrates the behavioral theories proposed in the other vertices of the triangle, i.e., I will try to identify the choice models developed to incorporate behavioral decision theories. This is also a challenge, and to do this I count on a small group of researchers that has been pioneering the exploration of these possibilities and that have already achieved promising results.

As this thesis unfolds I will: (i) present a literature review bridging the disciplines; (ii) followed by an empirical evidence intended to work as a compelling argument of the inevitability of considering choice process heterogeneity in choice models; and (iii) finish proposing and testing, with synthetic data, an econometric choice model that reflects some of the ideas presented in this document.

1.1 Justification

> Consumer choices concerning the selection. consumption. and disposal of products and services can often be difficult and are important to the consumer, to marketers, and to policy makers. As a result, the study of consumer decision processes has been a focal interest in consumer behavior for over 30 years. (Bettman, Luce, & Payne, 1998, p. 186)

And so the statement above is true 20 years after it was published. A reason to the scientific curiosity about choice processes is that making choices is a natural state of activity in any society, in any domain of human action (Louviere, Hensher, & Swait, 2000 p. 1). From a positive psychology perspective, to exercise choice among valued options fosters individual autonomy and competence, which facilitates intrinsic motivation development, supporting a natural human tendency toward activity and integration (Moller, Deci, & Ryan, 2006; Ryan & Deci, 2000).

Choice is also intimately related to the notion of freedom and Berlin (as cited by Bavetta, 2004) proposes positive freedom as the individual preservation of a personal sphere within which she exercises liberty through possibility to act, which requires alternative courses of action, and the existence of conditions under which one is capable of following her own desires. The same author defines negative freedom as the exercise of freedom with no constrains, nor imposed by other individuals neither by the State.

This notion of positive freedom describes our everyday choices, including those which are made much beyond the consumption domain, shaping the natural and social environment surrounding us. Usually, our possibilities to act, or choose among alternatives, are bounded by factors that may relate to our possible psychological states or by exogenous constraints. Remember that we can decide to vote or not to vote (when living under democratic systems) once we are registered and have our political rights intact. Then we choose among political parties or candidates that must be able themselves to run in electoral disputes. We can choose a professional career and when to change it or not, but our skills and knowledge restrict our choices. We can also decide how to share our time between work and leisure, but more leisure implies less money in and, probably, more money out, and this equation must be balanced. We may decide to have kids or not, to have safe sex or not, but these choices imply agreement with our partners. We may choose to recycle our garbage or not, but we need to find a proper place to dispose of it. Thus, in any choice domain, the freedom of choice is constrained by personal characteristics and contextual or situational variables, as I will detail in this dissertation. However, if, on one hand, the constraints are part of the choice process and integrated to a view of positive freedom, on the other hand, if individuals feel that free behaviors are eliminated or threatened with elimination, a psychological state of reactance will arouse toward the restoration of those behaviors (Miron & Brehm, 2006). These ideas are also consistent with the conception of free will, which is supported by self-control, conscious reasoning, and internal decision and it is “understood as one of these abilities that humans developed to be able to create, to function in, and to benefit from culture” (Baumeister, Sparks, Stillman, & Vohs, 2008).

Focusing on this thesis's domain, as consumers we also make choices every day, and they can be more elaborate like when we need to choose a neighborhood to live in; or our next vacation destination; or whether we are going to buy a new car or to solve our mobility needs using the myriad of available transport modes in the urban areas, which include public (bus or train) versus private (taxi or bike) alternatives. And even choosing a private mode we may opt to be alone in an Uber X or to share the ride in an Uber Pool. The choices may also be more trivial, like which alternatives to eat in one of the self-service restaurants where we have our everyday lunch. And they can even become automatic, as the habitual choice that a smoker uses to make when she, every time, lights a cigarette right after drinking a coffee. All these choices are also constrained by our cognitive ability to process information, by our budgetary restrictions, by our time availability to decide or even by the social desirability of the choices we make. Despite all these restrictions, experimental evidence show that longer choice sequences are preferred to shorter ones even when leading to the same outcomes (Bown, Read, & Summers, 2003), and that consumers are more satisfied when they have the chance of choosing an incentive after completing a task than when the incentive is chosen by someone else (Iyengar & Lepper, 2000).

To conclude, the study of human choice process and specifically of choice process heterogeneity is, firstly, justifiable from a theoretical-methodological perspective, but also to develop the strategic management in private and public organizations and, finally, to develop individuals' decision-making skills in every domain, including the consumption one.

From a theoretical-methodological viewpoint, despite the variety of choice models investigating heterogeneity beyond preferences and the growing interest in choice process heterogeneity, the adoption of these ideas, as an alternative to the pure random utility models based on the economic normative standpoint, is still a challenge to the discipline, as noted in Hensher (2014 p. 1):

> These presumptions have been questioned in the broader literature on heuristics and decision making that has evolved in a number of literatures, notably, psychology, economics and marketing; however the migration of ideas from this literature, which we refer to as process heuristics, has been slow to influence the way that discrete choice modeling has been represented. This is changing now, with a growing number of studies questioning the standard fully compensatory choice paradigm.

Another evidence of the necessity of the discipline to incorporate choice process heterogeneity is given in Dellaert et al. (2017 p. 2):

> The different goals we select and how we prioritize them affect our behavior and the choices that we make. Yet only recently has research begun to address the question of how goals can be directly incorporated in econometric models of individual decision-making to test theories about goals and improve our understanding and prediction of individuals' choices.

These two recent citations from leading scholars support my conviction that there is a lot of room to contribute to the enlargement of the knowledge's frontier in this discipline, attending to the persuasive call from Adamowicz et al., (2008) to close the gaps among the three scientific traditions involved in this thesis.

The variety of applied disciplines using choice models to develop strategies and policies, to the many stakeholders involved, includes marketing as well transportation planning, urban planning, health economics, environmental economics, labor economics, transport and sociology (Dellaert et al., 2017; Swait & Feinberg, 2014). It means that the knowledge originated from this kind of approach can support the efficacy of strategic planning, the policy formulation and the operations of a wide variety of public and private organizations.

Last, but not least, choices are individuals' means of pursuing their goals (Austin & Vancouver, 1996; Bettman et al., 1998; van Osselaer & Janiszewski, 2012). To know how individuals make decisions, in a world that demand conscious consumers, is to be able to empower better choices as a way to pursue individuals' and social's objectives.

1.2 Thesis structure

This thesis is comprised by six chapters articulating the ideas presented in this introduction. The three initial chapters composes a theoretical block, including the introduction and two other reviewing the literature. A second block is formed by the next two chapters portraying two empirical studies. Finally, I lay down my final considerations in the sixth chapter..

The chapter 2 has the objective of setting the normative reference and to present the tools to smooth the comprehension of the remaining content of this thesis. The first section is a brief review of the economic theory of consumer behavior, and it has the objective to make clear to the reader what the economic rationality is and what the *homo economicus* means, in behavioral terms. After a conceptual warm-up, a pause is required to present the initial notation used in discrete choice models and basic concepts from the behavioral decision theories. Thus, the objective of the second chapter is to introduce the initial layer of concepts, which supports the presentation of the last sub-section of the literature review and the empirical sections of this thesis. The reader who is unfamiliar with the specificity of either the behavioral decision theories or the discrete choice models is the one who I expect to benefit the most from this sub-section.

Chapter 3 is the most important of the theoretical block, and it is organized around goal based choice, once this is the concept that brings the proposed decision-making perspective of the *homo aptabilis* and, also, completes the structure to explore the different sources of heterogeneity proposed either in the behavioral decision theories or in the discrete choice models.

After reviewing the literature, the first empirical study in chapter 4, is a Monte Carlo experiment followed by a game theoretical approach. The objective is to study the effects of misattributing consumers' choice process heterogeneity into preferences. The Monte Carlo experiment is used to build demand representations for two marketing contexts, being two representations for each context. One representation is the true model accounting for choice process heterogeneity, and the other is the biased one, allowing taste heterogeneity and imposing choice homogeneity. The results are used as inputs to a game theoretical analysis to understand the effects of not accounting for consumers' choice process heterogeneity in a focal firm's payoff and in the market equilibrium.

In chapter 5, I use the knowledge generated in the previous ones to develop an econometric discrete choice model. This model assumes that choice is driven by the pursuit of multiple goals, it allows for individual adaptation based on context, task and individual heterogeneity, and also accommodates a two-stage decision process allowing for choice set heterogeneity.

Finally, in the final considerations I review the main findings of the four initial parts and also indicate some directions for future research.

2 LITERATURE REVIEW

Before starting the literature review, I define a choice as a sequence of behavioral and cognitive events resulting in a selective response. This selective response occurs over a set of alternatives described by attributes or consequences, and it is bounded by contingencies or conditional probabilities connecting the consequences to the actions or alternatives (Bettman et al., 1998; Jacoby, Chestnut, & Fisher, 1978; Tversky & Kahneman, 1981). Obviously, the scientific interest in the subject is drawn by the uncertainties connecting actions to consequences. Uncertainties arise as a result of the temporal displacement between current actions and future consequences, and because the individual may know her preferences in the present, but not in the future (Simonson, 1989).

The majority of the studies in economics and in cognitive psychology connects alternatives to consequences trough probabilities, for instance studying preferences among gambles in which the individual has information to infer the outcomes' likelihood, meaning that the outcome is objectively verifiable. In this context, the reference is to decision-making under risk and the choice relies on the individuals' inferences about probabilities and preferences for risk. In the marketing literature, the uncertainty from consumers' choices among alternatives is grounded on subjective expectations about the performance of the chosen and the unchosen options. In this case, the choice results from preferences. Given that inferences and preferences are drawn from the same cognitive processes (Weber & Johnson, 2009), it is not my concern to explicitly identify risk or uncertainty in the literature review, although the focal interest sits on decision-making under uncertainty.

2.1 Economic theory of consumer behavior

To get started, the economic theory of consumer behavior offers a normative analytical structure to understand consumers' choices based on a rational perspective that evolves into a utility theory. The theory expresses a view of rationality that is parsimoniously defined through four axioms unfolding decision-making under risk (Von Neumann & Morgenstern, 1947; see Pilli, 2012 for a detailed description of the axioms and the utility theory). For the concern of this thesis, the consequences of the proposed rationality matter as long as it implies that during the decision-making process: (i) the preferences of the focal alternatives are completely ordered; (ii) the preferences among alternatives are transitive; (iii) the preferences are represented by interval scales, meaning that distances between the alternatives' utilities are considered, and; (iv) the inclusion or exclusion of alternatives in the considered set does not affect the preference-indifference relation between any two compared alternatives. In other words, preferences are complete, stable, represented by a subjective latent quantity and independent of irrelevant alternatives, defining what is called, in the behavioral decision theory literature, as the *homo economicus* (Lee, Amir, & Ariely, 2009; Swait & Feinberg, 2014). And the utility theory deals with the individuals' choices, as well as their preferences and judgments, being founded on a set of preference-indifference relationships among a set of objects or alternatives and the respective behavioral predictions that can be derived from this knowledge, given rationality (Fishburn, 1968). Therefore, the expected utility of an object can be expressed through a latent quantity that summarizes its anticipated ability to satisfy an individual's desire, i.e., the expected utility is a quantity that represents a subjective evaluation of a focal object and it is used during the decision-making process such that the alternative having the highest expected utility is chosen.

Still in the field of economics Houthakker (1952) concludes that demand theory faces a limitation by not acknowledging that different products possesses different qualities and proposes a new approach considering products' prices, quantities, and qualities. This new proposition leads to the development of a novel economic consumer behavior theory (Lancaster, 1966) in which the preference is linked to sets of characteristics, and alternatives are indirectly ordered conditional on the characteristics that they possess. Now, consumers choose among products described by sets of characteristics, but do not choose the characteristics that are attributed to each set or product. Moreover, individuals may combine different products to constitute sets of characteristics that are different from those observed in the original alternatives.

The new concepts also evolved into a multi-attribute utility theory that is summarized by Keeney (1972) in the following utility function:

$$\mu(x_1, x_2, \dots, x_n) = \mu_1(x_1) + \mu_2(x_2) + \cdots + \mu_n(x_n) \tag{2-1}$$

This is an additive utility function that expresses that any alternative utility is a function of the marginal probabilities distribution of the vector (*X*) that describes the alternatives' characteristics. This expression can be expanded (see Raiffa[1], 2006) to account for the joint probability distributions of the products' characteristics. One important detail is that this is a deterministic function, meaning that the totality of any alternative utility is explained by its characteristics. See Rieskamp, Busemeyer, & Mellers (2006) for a review and introduction to preference and utility theories, including the random utility theory that will be presented in the discussion about choice modeling and will be used in the empirical sections of this thesis.

In summary, this analytical structure provides, given the observation of consumer choices, an inference of the indirect latent utility of the alternatives in a given choice task, derived from the subjective values of its attributes. There is a decision strategy (or behavioral rule), which is to add the subjective values derived from every product attribute and integrate it to a global latent variable named utility, and then to proceed with this computation for every alternative available and, finally, to choose the one that has the highest utility. Lastly, subjacent to this process there is an objective of making a decision that results in the best possible end state in terms of well-being, given the alternatives available and the individual's preferences. And the compatibility of this process with the rationality assumptions gives predictive power to the theory. Notice that, from the observed choices, a specific perspective of rationality is imposed and the evaluation and the behavioral rule is inferred.

The critics of the utility theory, which will be exposed soon, evolve around its demands over the human brain abilities and about the imposition of a unidimensional criterion to evaluate alternatives. Acknowledging the value of this critiques, which motivate this thesis, it is important to recognize the importance of this normative model to the study of human judgment and decision-making, as Schoemaker (1982, p. 529) briefly did:

> It is no exaggeration to consider expected utility theory the major paradigm in decision making since the Second World War. It has been used prescriptively in management science (especially decision analysis), predictively in finance and economics, descriptively by psychologists, and has played a central role in theories of measurable

[1] Originally published as: RAIFFA, Howard (1969). Preferences for Multi-Attribute Alternatives. Memorandum RM-5868-DOT/RC. Prepared for U.S. Department of Transportation, Federal Railroad Administration, Office of High Speed Ground Transportation. The RAND Corporation, Santa Monica, California.

> utility. The expected utility (EU) model has consequently been the focus of much theoretical and empirical research, including various interpretations and descriptive modifications as to its mathematical form.

To conclude and make sure that the reader has the correct understanding of this paradigm, it is important to dismiss one review that I consider equivocated. To do so, I will rest on the argument used by Luce & Raiffa (1957) to emphasize that utility does not cause choice, but the other way around. Given the observed choices, an unobserved choice process and one analytical structure imposed by the theory, the analyst infers the utility from the choices, i.e. choice "causes" utility.

2.2 Setting the ground: basic notation for discrete choice models

I will start defining discrete choice models' basic notation and concepts using a basic model of the random utility theory class, or simply RUT. These theories are consistent with expected utility theories since they propose choice as the outcome from a deterministic decision rule, i.e. utility maximization, applied to utilities that vary over time and contexts (Rieskamp et al., 2006). This basic model, from which discrete choice modeling has evolved (I will refer to this class as DCM), is the multinomial logit model described by McFadden (1974). I will also use the acronym RUM to the class of random utility models consistent with RUT.

2.2.1 Setting the ground: basic notation of discrete choice models

I will introduce a multinomial logit model (MNL) that is the powertrain of discrete choice models, and it works in agreement with the axiomatic view supporting the *homo economicus* behavior. Here, my objectives are two-fold: firstly provide the basic notation to the following sections; secondly, to use a basic DCM to illustrate where the opportunities to study consumer heterogeneity rests in this class of models, which are RUM. I will start this sub-section deriving the MNL, keeping it very simple, and once I have all the elements of the model I will identify the sources of heterogeneity. Notice that to model consumer heterogeneity implies in relaxing some assumptions of the MNL, and I will not formally work through this since this is not the focus of this thesis. But, the empirical evidence that I will reference in the following sections are the flexible versions of this logit model. The reader who needs a more detailed derivation of this model, and of the more flexible ones, can find good reviews in Train,

(2009) and Louviere et al., (2000). And the derivation of the specific models that I will reference can always be found in the original papers.

To start, a model of individual choice behavior, consistently with the expected utility theory, rests upon "(l) the objects of choice and sets of alternatives available to decision-makers. (2) the observed attributes of decision-makers, and (3) the model of individual choice and behavior and the distribution of behavior patterns in the population" (McFadden, 1974, p. 106). There are two new elements in comparison to the description that I have done for the expected utility theory. The individual characteristics have a role in this model, as well it is concerned with the choice distribution in the population. Now I will present this model following McFadden, omitting repetitive references unless strictly necessary, and completing with other authors when needed. The importance of the concepts that follow is the reason to bring the extensive content available in the referenced author, instead of recommending the reader to learn the details from the original. I am using this author because he was the first, as far as I know, to derive the model.

First, let's denote a universal choice set M as the set of all the possible or existing alternatives, and let's S to denote a set of measured attributes of the decision-maker. Any individual randomly drawn from a population has a vector of characteristics $s \in S$ and will face a set of available alternatives $J \in M$. From now on, I will refer to the set of available alternatives as the choice task or the choice scenario. Each alternative is defined by a vector of attributes $x_j = \{1, \dots, \mathcal{K}\}$ and the subscript j in x means that the attributes may be unique, or not, to any specific option. Moreover, every attribute can be described by its own vector of levels or values such that $\ell_k = \{1, \dots, \mathcal{L}\}$. Now, every alternative $j \in \mathcal{M} = \{x_{j1}, \dots, x_{jk}\}$, i.e. every alternative in the universal choice set is a set of attribute measures, which can be nominal (like brands), ordinal (like the screen size of mobile phones) or continuous (like prices).

One example of a choice task is presented in Figure 1, and here the choice task J, drawn from the universal choice set $\mathcal{M}$, is composed by four alternatives, each one described by a set of generic nominal, ordinal or continuous attributes. Every attribute has different levels or values that can be common or unique to the option.

Figure 1 - Example of choice task

Every alternative's conditional choice probability, from an individual with characteristics s facing a choice task J, is a draw from the multinomial distribution with probabilities given by:

$$P(j|s, J), \forall j \in J \quad \textbf{(2-2)}$$

It is necessary to map the individual characteristics vector s and the vector of alternatives J into a chosen option of this choice scenario. This is done through an individual decision rule $h \in H$, e.g., if the individual is trying to maximize expected utility h is a specific utility function from a set of possible utility functions that would maximize expected utility. To make it clear, a specific decision rule is, for instance, the additive linear function described in equation (2-1); and the one with interactions between any two product attributes, that I have mentioned in the same sub-section, would be a different decision rule h pertaining to the set of utility maximization rules H. Given unmeasured individual characteristics across the population, there are many possible decision rules in H and there is a probability π, as a function of the observed subsets of H, defining the distribution of decision rules in the population, such that:

$$\mathrm{P(j \mid s, J)} = \pi[\{h \in H | h(s, J) = j\}] \quad \textbf{(2-3)}$$

In words, equation (2-3) states that the probability of alternative j to be chosen by a consumer with individual characteristics s, facing a choice scenario J, equals the probability of a decision rule h to be used, conditional in the choice of alternative j. Before moving on, I will use the terms decision rules, behavioral rules or decision strategies to refer to the functions that maps s and J into a chosen alternative.

Given the expected utility maximization decision rule, the multinomial distribution to represent the choice probabilities, and assuming π as a member of a parametric family of probability distributions, an econometric model is ready to be developed.

Once again, let's draw, from the population, a consumer with a vector of measured individual characteristics s, facing a choice task J, with alternatives $\mathrm{j} = 1, \dots, \mathrm{J}$ described by vectors of attributes x_j. Her utility function may be written as:

$$U = V(s, x) + \varepsilon(s, x) \tag{2-4}$$

This function is decomposed into two components, the first one is systematic and it is called the representative utility. The second component is stochastic and represents individual idiosyncratic tastes for the alternative with attributes x. The notation in equation (2-4) expresses tastes as common to the population, but I can index the representative utility as $\mathrm{V}_n(\mathrm{s}, \mathrm{x})$ and allow for taste heterogeneity. Remember the definition of RUT that tastes may vary but the behavioral rule is deterministic, i.e., given the preferences, the alternative with the largest utility is chosen. It means that if the analyst can recover the representative utilities from the measured variables, the stochasticity of the process emerges from unobserved variables, by the analyst, and not from the inability of the consumer to identify her preferred alternative.

Now let h_u denote the expected utility maximization rule and $J = \{x_1, \dots, x_j\}$ be the choice scenario, then, the probability of a consumer drawn from the population, with individual characteristics s and facing a choice task J to choose the alternative x_i is:

$$\begin{aligned} P_i \equiv \mathrm{P}(\mathrm{j}|\mathrm{s}, \mathrm{J}) &= \pi[\{h \in H | h(s, J) = j\}] \\ &= P\left[U(s, x_i) > U(s, x_j), \forall j \neq i\right] \\ &= P\left[V(s, x_i) + \varepsilon(s, x_i) > V(s, x_j) + \varepsilon(s, x_j), \forall j \neq i\right] \\ &= P\left[\varepsilon(s, x_j) - \varepsilon(s, x_i) < V(s, x_i) - V(s, x_j), \forall j \neq i\right] \end{aligned} \tag{2-5}$$

Again in words, equation (2-5) states that the desired result is the probability that each difference in the stochastic portion, $\varepsilon(s, x_j) - \varepsilon(s, x_i)$, of the utility function to be smaller than the difference in the representative utility $V(s, x_i) - V(s, x_j)$. As noted by (Train, 2009 ch. 2) two important observations, which motivated me to present the derivation up to this point, must be made in regard to the empirical identification of this model. The first is that given that only $J-1$ utilities can be identified, we are concerned about the difference in utilities as it can be noticed in the equation (2-5). It leads to the second observation that the scale of utility is arbitrary, and this will be further explored soon. Now, I will follow the demonstration from Train, and equation (2-5) is expressed as the cumulative probability:

$$\int_{\varepsilon} I\big(\varepsilon(s,x_j) - \varepsilon(s,x_i) < V(s,x_i) - V(s,x_j), \forall j \neq i\big)\, f(\varepsilon_n) d(\varepsilon_n) \tag{2-6}$$

Where I is an indicator function that turns to 1 if the expression within parenthesis is true an 0 otherwise, i.e., it points to the chosen alternative that is the one that maximizes expected utility. Different choice models are derived from equation (2-6) and some of them takes a closed form for specific cases of $f(.)$. Remember that I am interested in the simplest case, which is the MNL, and it is derived from (2-6) under the assumption that the stochastic portion of the utility function is distributed independent and identically (i.i.d.) extreme value. The full derivation of the model can be found in detail in (Train, 2009 ch. 3) or in (Louviere et al., 2000 ch. 3). From here, I'll jump straight to the end point, where, in the logit model the probability of a consumer drawn from the population with a vector of individual characteristics s and facing a choice task J, to choose a specific alternative, is given by:

$$P_{ni} = \frac{e^{V_i}}{\sum_{j \in J} e^{V_j}} \tag{2-7}$$

Equation (2-7) results from a deterministic behavioral rule, i.e. utility maximization, consistent with the *homo economicus*. But the model can be always derived as a function of other decision strategies introducing heterogeneity also at this point.

To conclude, I have to add a final detail that results from the fact that the utility's scale is arbitrary, given the identification issue. In the logit model, the scale is normalized by the standard deviation of the stochastic term, σ, of the function, what means that in equation (2-7) the utility and σ are perfectly confounded, as:

(2-8)

$$P_{ni} = \frac{e^{V_i/_\sigma}}{\sum_{j\in J} e^{V_j/_\sigma}}$$

Moreover, Louviere et al. (2000 p. 235) demonstrate that the variance of the unobservable portion of the utility function is:

(2-9)

$$\sigma^2 = \frac{\pi^2}{6\lambda^2}$$

Now, there is a scale parameter λ inversely related to σ^2, i.e., the larger the variance of the stochastic term the smaller is the scale parameter. And under the proper transformation:

(2-10)

$$P_{ni} = \frac{e^{\lambda V_i}}{\sum_{j\in J} e^{\lambda V_j}}$$

In equation (2-10), the scale parameter λ connects the variance of the stochastic component to the representative utility and the larger is the scale parameter (or the lower id the variance of ε) the closer the probability of the chosen alternative is to 1. And the smaller is the scale parameter (or the higher is the variance of ε), the more similar are the probabilities among the alternatives in the choice scenario. As observed by Swait & Adamowicz (2001) λ measures a relation signal (representative utility) to noise (variance of ε), such that as λ increases the signal commands the choice and as λ decreases choices are driven by noise or randomness. Notice, that the preference order of the alternatives does not depend on the scale parameter, i.e., the chosen alternative is always the same. But, given the same preferences, the choice probabilities may vary across consumers or occasions disclosing a series of behavioral effects. The problem is that the normalization of the taste parameters introduces a perfect confoundment between utilities and the variance of the stochastic term. However, properly modeling the scale parameter may uncover an additional source of consumer heterogeneity in the choice process, from the stochasticity. Louviere (2001) proposes that the scale parameter can be decomposed to identify within-subjects, between-subjects, between-contexts, between-measurement instruments, between-time periods and other sources of variability in the random term. From this point on, I will use terms signal and noise to refer to the relevance of the representative utility or of the stochastic terms as drivers of the choice process.

Finally, there is one last source of heterogeneity to be exposed. Notice that either the choice process proposed by the economic theory of consumer behavior and the MNL derived above assume the consumer evaluating every alternative in the choice task. This is a not a reasonable assumption, given the human brain cognitive limitations (H. A. Simon, 1955) and the resulting propensity to simplify the choice process considering only a subset of the information available (Weber & Johnson, 2009). Lot about these issues will follow as this thesis unfolds, but by know it is important to assert that selectivity results in the consumer's consideration of only a subset of the alternatives available in the choice task J. The subset of alternatives considered by any consumer can be any of the $2^J - 1$ nonempty subsets of J; moreover the consumer subset is subjective and unobserved, and I will refer to this source of consumer heterogeneity as choice set and the processes leading to it as choice set formation (Manski, 1977; Swait & Ben-Akiva, 1987). This issue will be largely discussed in the next section, so I will put it to rest by now.

To summarize, I have identified four possible sources of consumer heterogeneity in the RUM: (i) preferences vary among consumer and across occasions, this is a well-established phenomenon and giving equation (2-6) the proper stochastic structure lead to the mixed logit model and, as special case, to the latent class model; (ii) the decision rules are a source of heterogeneity, once we remove the imposition of the utility maximization to allow for the occurrence of a variety of behaviors, as I will describe in this thesis; (iii) the stochasticity of the choice process is a source of heterogeneity that can be modeled though the scale parameter; and (iv) the unobserved choice sets are a final source of consumer response heterogeneity.

Preference heterogeneity is well accepted across disciplines and it is current practice to account for it in DCMs, so I am not especially concerned about this topic in the remaining chapters, except if the explanatory factors that lead to variation are relevant. The same idea applies to scale parameter, that will only be used to illustrate how DCM can explain situational effects trough the stochasticity of the choice process.

Choice set heterogeneity and decision rule heterogeneity will be explored not only in the literature, but mainly in the empirical section of this thesis. In the first empirical application, in section II, I will model the effects of not accounting for choice set formation in the resulting demand function and in the downstream consequences on the firm decision-making. In the second empirical application, I will introduce decision strategy heterogeneity, allowing any population to be described by a mixture of behavioral rules.

2.2.2 Setting the ground: basic concepts from behavioral decision theories

Since its early days, the utility theory developed from the presented axiomatic view of rationality has been challenged by behavioral decision theorists (Pilli, 2012 presented a review of the literature describing the experimental evidence of violation in every of the four rationality axioms mentioned above). A major contribution comes from psychology, either as one of the disciplines involved in the studies or offering the framework used by other researchers. The approach is based on descriptive models that should explain systematic deviations between normative prediction and observed choices (Baron, 2004 p. 19) with a special concern about the validity of the axioms supporting normative theories and about the processes and psychological states involved in the decision-making (Schoemaker, 1982).

Researchers from this tradition proposed an alternative paradigm to individual decision-making, based on the bounded rationality premise (H. A. Simon, 1955, 1990). Accordingly to this author, due to the limits in the human brain working memory and in its power and speed of processing, most activities are executed through approximation methods preventing the optimality of behaviors, as predicted by the normative models. Moreover, given the adaptability of any organism, the behavior is flexible and defined by the interaction between the individual and the environment. As a result of this idea, Bettman et al. (1998) propose that consumers' choice is a constructive process that results from the impossibility of the individual to form , store and retrieve well-defined preferences.

Notice that the concept of preference, defined as the subjective value reflecting liking or disliking of the object under evaluation (Bettman et al., 1998; Simonson, 2008; Weber & Johnson, 2009) is equivalent to the utility concept, except that the constructive perspective denies the viability of an axiomatic view and considers that they tend to be formed on the spot, being contingent to the individuals' and to the environment characteristics. Before proceeding, I just make it clear that preferences and utilities, as well as tastes, are going to be used as the descriptors of this subjective evaluation of the alternatives or its characteristics.

The next step is to examine the interaction between the individual and the environment as a driver of choice. To support this investigation, I introduce the framework in Figure 2, identifying the different kinds of variables, all of them sources of heterogeneity, involved in the choice process with their respective definitions. In the consumer choice domain, the object is the choice scenario describing the alternatives to be considered and a few examples of its characteristics are the number of alternatives in the choice task, the number of attributes describing each alternative and the inter-attribute correlation. The situational variables are environmental elements that do not describe properties of the object but interact with it (Belk, 1975; Thaler & Sunstein, 2008). Some examples of situational variables are the existence of time pressure to make a choice or the necessity of justifying the choice to other people. Lastly, the person's characteristics may condition the behavioral response and the effects caused by these variables will be termed as individual effects. The characteristic can be demographics, like gender or age, or psychosocial, like socioeconomic status, attitudes, personality and other traits. Finally, the behaviors are the choice probabilities distributions made by the consumers and the effects caused by the previous variable on such distributions.

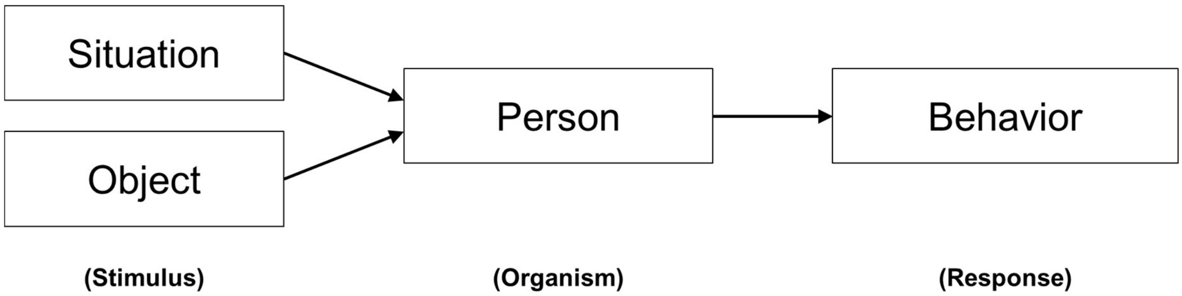

Figure 2 - Revised S-O-R Paradigm
Source: : (Belk, 1975, p. 158)

Next, I will describe, mostly, the evidence presented in the behavioral decision theories to support this structure proposed in Figure 2 and to illustrate these effects in practice. In this sub-section I will reference only DCMs that derive environmental effects from the choice probabilities distribution, since they are good illustrations of the assimilation of this paradigm into econometric choice models. It is important to notice that I am not maintaining that Figure 2 is a territory pertaining to BDT, since the discrete choice models to be reviewed in other chapters also incorporate variables from this framework and, most important, my own empirical investigation will rest upon it.

2.2.2.1 Context effects

The context effects are those resulting from the characteristics of the focal object or, in this case, the choice task. I have mentioned that one of the concerns from psychologists is to test the validity of the axioms proposed in the normative theories. The fourth axiom proposed in the economic theory of consumer behavior is the independence of irrelevant alternatives, also known as IIA. This axiom implies that the preference relation between any two alternatives, in a given choice task, is independent of any alternative other than A and B, i.e., if A is strictly preferred to B in a choice task with only these two alternatives, the inclusion of an alternative C cannot modify the original preference relation. Moreover, given the set of axioms, if an alternative C is included in the choice task, it should draw its choice probability proportionally from A and B, such that, the relation between the choice probabilities of the existing alternatives remain constant independently of the presence of C. If all these implications hold, the choice probability of any alternative cannot increase after the inclusion of a new one.

Huber, Payne, & Puto (1982) and Huber & Puto (1983) have demonstrated a classical violation of the IIA, known as asymmetric dominance or the attraction effect. Figure 3 illustrates the concept considering a choice with two alternatives described by the level or the amount of both attributes that each one possesses. The figure describes that alternative A is superior to B in attribute 1 and that alternative B is superior to A in attribute 2. The dotted line is an indifference curve informing how much of one attribute a consumer is willing to trade-off to get more of the other attribute, and given that both products lie in the indifference curve they are perfect substitutes to each other. The shaded area bellow the indifference curve is a special region of the figure, since any alternative in this region is inferior to alternative B in both attributes, i.e., it is dominated by B, thus it should always be disfavored to B. When compared to A, any alternative in this region is inferior in attribute 1 and superior in attribute 2, meaning that the consumer's choice between A and the alternative in the shaded area will depend on how much of one attribute the consumer is willing to trade-off by one unity of the other attribute, i.e. the marginal rate of technical substitution between them. Thus, asymmetric dominance means that an alternative in the shaded region is dominated by B, but not by A.

Remember that the IIA states that the inclusion of a new alternative cannot change the preference relation between A and B, and that the ratio between the choice probabilities of the existing alternatives should remain constant when a new one is included. In the two papers referenced above, the authors conducted a series of experiments varying the product category, the placement strategy of the new alternative in the shaded area, and using two or three attributes to describe the alternatives. Systematic increases in the choice probability of B were reported, demonstrating the violation of the IIA.

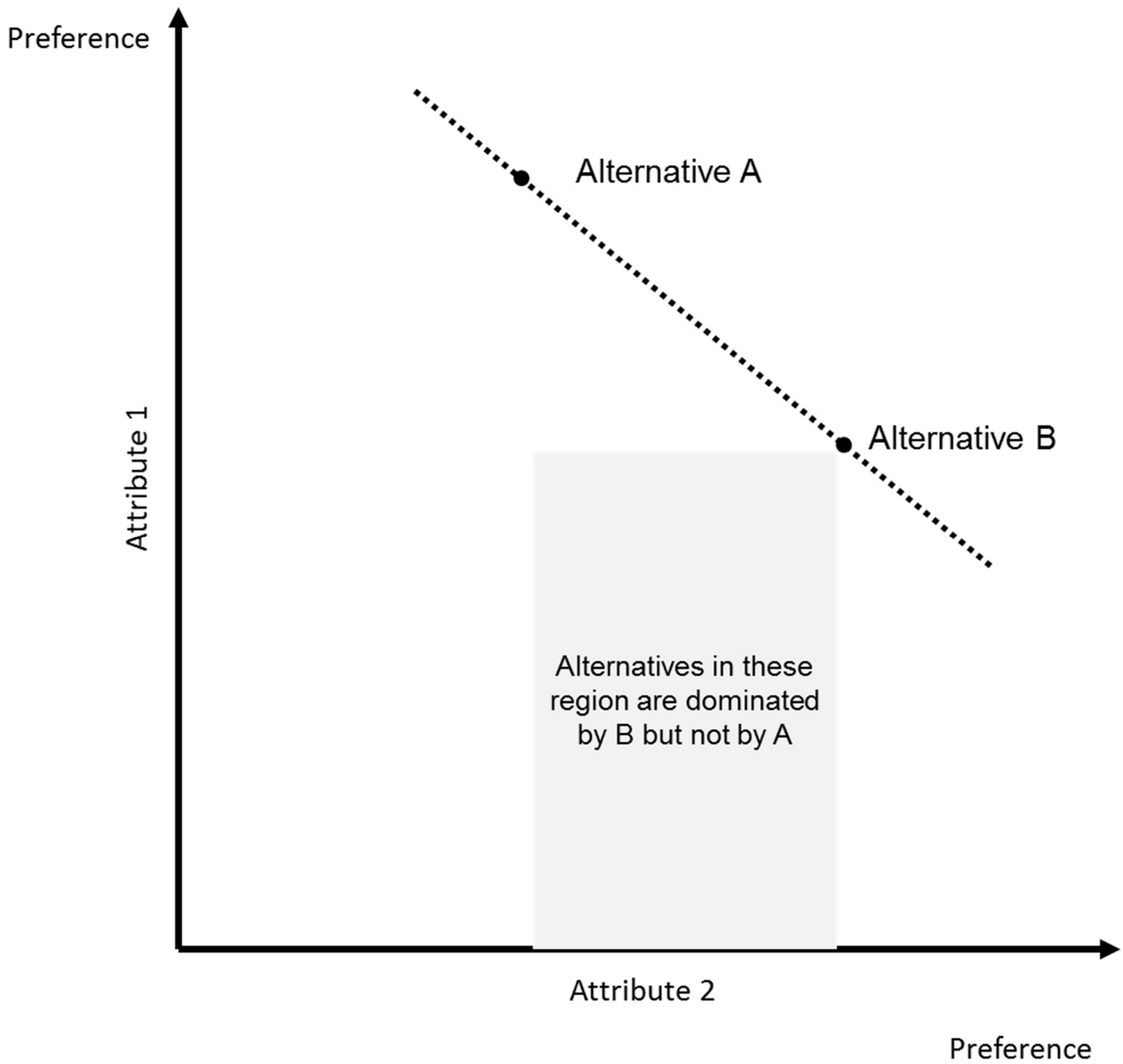

Figure 3 - Asymmetric dominance
Source: (Huber et al., 1982 p.92)

Another classical context effect is portrayed in Figure 4 and it is known as compromise effect. All the possible alternatives lie in the indifference curve, meaning that they are perfect substitutes for each other. In the base scenario the consumer can choose only between alternatives B and C. Now, the region of interest is any section of the indifference curve, except the one between B and C. What makes the focal region special is that any new alternative included in the base scenario turns one of the existing alternatives to be the compromise one. For instance, if the alternative A is included, the new choice scenario is [A, B, C] and A is superior to B in attribute 1, C is superior to B in attribute 2, but B is superior to A in attribute 2 and superior to C in attribute 1. In other words, B is neither the best nor the worst in any attribute. Likewise, if instead of including A, alternative D is included in the choice scenario [B, C, D], the compromise alternative turns to be C.

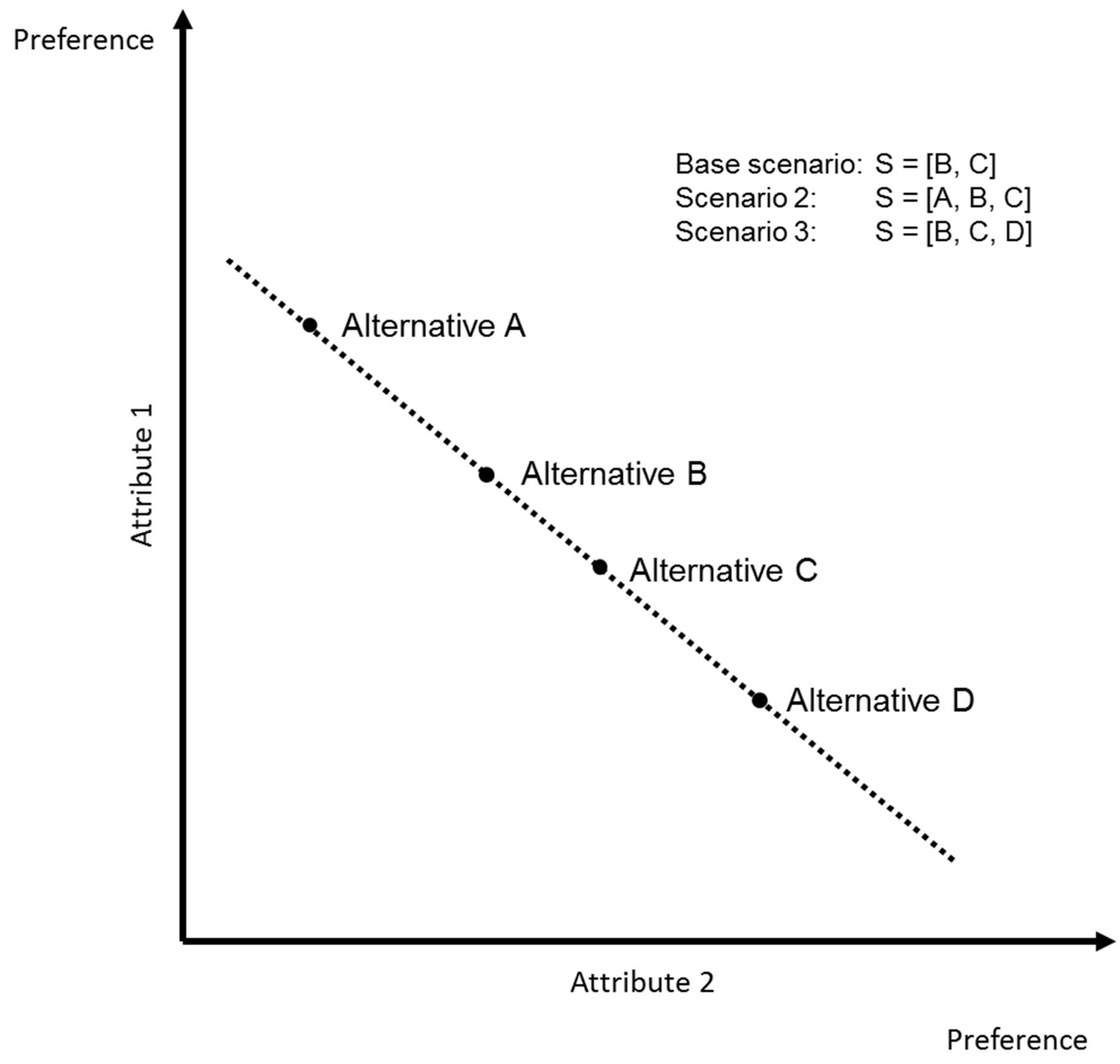

Figure 4 - Compromise effect
Source: (Simonson, 1989 p. 161)

In a series of experiments across different product categories, Simonson, (1989) provided empirical evidence of the increase in choice probabilities of the compromise alternative, i.e., the choice probability of B increases in the choice scenario [A, B, C] compared to the choice scenario [B, C]. On the hand, in the choice scenario [B, C, D] the alternative with increased choice probability, compared to the base scenario, is C.

Remember that I mentioned that the logit model is IIA and as consequence, it is subject to the context effects demonstrated, which results from the substitutability pattern among alternatives. This phenomenon are due to the relation between alternatives, that causes a violation in the IIA axiom. Some choice modelers have relaxed the imposition of the IIA over the RUM, modeling the scale parameter to allow for flexible substitutability patterns.

Dellaert, Brazell, & Louviere, (1999) modeled the scale parameter as a function of the absolute prices and the alternatives' price difference in choice task in a study with tourist choosing among possible bus excursions. The results indicate that as either absolute prices or differences in prices increase, so does the noise in the utility function. One possible explanation is that higher prices or higher differences increase choice difficulty turning harder the trade-offs between attributes and costs. Another account is that, given the possibility of choice avoidance, higher prices or higher differences turn the average utilities in the choice task more similar to the utility of no-choice.

In another study in the tourism industry, DeShazo & Fermo (2002) used context variables to model the scale parameter of the DCM as a function of cognitive burden and choice complexity. The results for measures of cognitive burden are: an increase in number of attributes, ranging from four to seven, increases the noise in the utility function; an increase in the number of alternatives first strengths the signal and then increases the noise in the utility function; the results for measures of choice complexity are: increasing the variability of attribute levels across alternatives, increasing the average variability of each attribute of the alternative across occasions, and increasing the variability of this last measure across alternatives increases the noise in the utility function

In another DCM application, Swait & Adamowicz, (2001) related choice complexity to the entropy in choice probabilities in a given scenario, i.e. the more similar the alternatives' choice probabilities the higher the entropy, and they proposed the entropy results from choice complexity and its variability should be related to the variance of the error term. Across a series of stated and revealed preferences datasets they concluded that, despite specificities in studies' results, there is a convex relationship between the scale parameter and preference similarity (entropy), such that, at lower levels of complexity the signal in the utility function was stronger, reflecting easiness of choice; at intermediate levels of complexity the noise increased reflecting the cognitive effort to make a decision; and at higher levels of entropy, similar choice probabilities reflected in low random term variance and strengthened the signal of the utility function.

To emphasize, the context variables describe the characteristics of the choice scenario and the relation among the alternatives. Either in the attraction effect or in the compromise effect, one change in the context is the inclusion of a new alternative, i.e., the number of alternatives is manipulated. In the case of the attraction effect another context variation is that the new alternative is asymmetrically dominated by one of the existing, i.e., there's a relation between the new and the standing alternatives. In the case of the compromise effect, the inclusion of a new alternative turns one of the existing to be a compromise solution, i.e., the one that is neither the best nor the worst in any attribute.

The substitutability pattern imposed to the *homo economicus* is violated by the context variable, and it requires that the structure of the stochastic part of the utility function of the RUM to be modeled to account for these effects.

2.2.3 Task effects

The task effects are caused by situational variables, present in the choice occasion, that influence the outcome of the process. As delineated by Belk, (1975), these variables must be precisely sited in time and space and may relate to the physical or social environment, to the temporal perspective of decision-making, to the task characteristics, or to any antecedent condition. I will present empirical examples that demonstrate some of the task effects, following up and clarifying Belk's definitions.

To start, task characteristics include specific requirement about the choice occasion, like the possibility, or not, to avoid choice allowing the consumer to extend information search or even to just preserve the status quo; or expected differences, by the consumer, in the buyer-consumer role, as choosing for the self or to others (Belk, 1975). See Pilli & Mazzon (2016) for a review on choice avoidance, nonetheless some of the important effects caused by the option to avoid choice are: it weakens the compromise effect and strengths the attraction effect (Dhar

& Simonson, 2003), it also leads to more attribute-based information processing, storage and retrieval, it evokes more evaluative judgments and it rises the importance of attributes performing close to consumer thresholds (Parker & Schrift, 2011).

In another category of situational variables, the variables related to the social environment identify the presence of other persons, their apparent roles and eventual interpersonal interaction during the judgment and decision-making process. Simonson (1989) provides experimental evidence that both the attraction and the compromise effects are magnified when consumer expect to have to justify their choice to other people.

Next, in accordance with Belk, (1975) the temporal perspective is shaped by variables articulating the choice process to past or future events that may impose time constraints to the judgment and decision-making process. Time pressure is a variable largely studied in consumer behavior, provokingly suggesting that it leads to action, and choice, and inducing to less information usage, although it causes psychological discomfort. Empirical evidence reveals that it reduces satisfaction with choice and if, simultaneously, the number of alternatives in the choice scenario is increased, it enlarges the perception of choice difficulty and frustration with the decision-making process (Haynes, 2009). On the other hand, the less time to complete a task the more likely the individual will get it done (Inman & McAlister, 1994; Tversky & Shafir, 1992). Also, time pressure increases the preference for superior quality brands versus inferior quality brands, and this effect holds even if the inferior brand is presented with novel and distinctive features, as well it increases the preference for alternatives which are more expensive and have more features (Nowlis, 1995). Moreover, if choices involve conflict, i.e., there is no dominant alternative in the choice task and the consumer has to trade-off among attributes, time pressure increases the incidence of a choice, in opposition to a condition where time pressure is absent and the consumer prefers to defer the selection of an alternative (Dhar & Nowlis, 1999; Tversky & Shafir, 1992).

Finally, antecedent conditions refer to transitory emotions or conditions that describe some state proximately antecedent to the choice process, in opposition to a resulting state (Belk, 1975). For instance, consumers that have the opportunity to articulate an ideal product through some elicitation method reveal stronger preferences in the choice process compared to those who do not have the same opportunity (Chernev, 2003); assortment mere categorization improves variety perception and satisfaction with the choice process only to unfamiliar consumers, i.e., those who does not have defined preferences (Mogilner, Rudnick, & Iyengar, 2008).

2.2.4 Individual variables

Individual variables are enduring characteristics, describing the decision-maker, holding across choice occasions and time (Clarke & Belk, 1979). Beyond socio-demographic variables, personality is a well-established kind of trait related to decision-making.

Schwartz et al. (2002) developed a psychometric scale relating individual style and decision-making objectives, identifying the propensity of the individual to be a maximizer or a satisficer. Maximizers tend to prefer choice scenarios with more alternatives but to feel less satisfaction and more regret with the choice (Dar-Nimrod, Rawn, Lehman, & Schwartz, 2009).

Need for cognition is defined as the individual necessity of structuring, understanding and assigning meaning to the reality experienced by the individual, and it is considered as the need to orient behavior toward goals and to cause tension and frustration when the goal is not achieved (Cohen, Stotland, & Wolfe, 1955). Cacioppo & Petty (1982) and Cacioppo, Petty, & Feng Kao (1984) present a psychometric scale to measure the intensity of the trait and empirical evidence relating need for cognition to choice process is profuse. For instance, individuals with lower scores in NFC are more susceptible to framing effects (Smith & Levin, 1996); newly formed attitudes toward unfamiliar objects are more enduring and more resistant to counter-arguments among individuals who score high in the scale (Haugtvedt & Petty, 1992); and high-NFC individuals deploy more cognitive effort and search for more information during the choice process (Verplanken, Hazenberg, & Palenewen, 1992).

Nevertheless, the most important variable at the individual level is the goal followed through the choice process, defined in my introduction as the structure that lends rationality to *homo aptabilis* decision-maker. Goals are “internal representations of desired states, where states are broadly construed as outcomes, events, or processes” (Austin & Vancouver, 1996 p. 338). Reaching this point, having defined the main concepts that will be needed ahead, I close this section and start a new one reviewing the literature that represents the goal based choice process, detailing its elements and some examples, and using the three theoretical frameworks that support this thesis.

3 MULTIPLE GOALS BASED CHOICE PROCESS

From a psychological viewpoint, a goal is defined as a subjective representation of a desired end state, connected to positive affect, which motivates the lessening of the gap between the desired and the current state, also subjectively represented. Goals can be attained through multiple means and irrespective of initial state, although it implies in resources' deprivation (Austin & Vancouver, 1996; van Osselaer & Janiszewski, 2012). Then they influence the decision modes, i.e. the means, since affective, analytic or rule-based processes contrast in their potential of satisfying the goals (Weber & Johnson, 2009). Moreover, these subjective representations are knowledge structures, i.e. "goals are connected to other concepts in memory (e.g., means, other goals, contexts) through excitatory and inhibitory associations" (van Osselaer et al., 2005). In summary, individuals are driven to goal attainment and any specific goals' influence in behavior is conditioned by possible strategies and context characteristics available in memory and associated to those goals. In other words, goals, means, and context interact, through its associations in memory, in the production of behavior.

Following these concepts, choices are a mean to achieve multiple goals that are hierarchically organized, subsets of these goals being relevant as a function of the environment (Bettman et al., 1998), i.e., of the context and task demands over perceptual and cognitive resources accessible to the decision-making process. Figure 5, adapted from (Dellaert et al., 2017) illustrates a multiple goals choice based process, consistent with the *homo aptabilis modus operandi*, that I start to describe before extending its concepts and implications. The authors propose that a goal based choice process operates in two separate, but connected, spaces, as follows.

I will start with an overview of the process described in Figure 5 before extending its concepts and implications. The first important notion is that a meta choice firstly results in the selection, by the consumer, of a goal choice strategy that informs which subset of goals, i.e. which $\Gamma_A \subseteq \Gamma$, will be active and what criteria will be used to determine goal attainment. During any choice event, any combination of goals may be activated, i.e., the consumer may use only one goal or, more commonly, a combination of two or three or any number of goals. These active goals may operate through thresholds, meaning that any alternative must have a minimum performance to be accepted as attaining the specific goal, or through the maximization of one goal or of the linear combination of the active goals. In other words, goals operate as constraints and/or as objectives.

This goal determination process includes compromises reflecting the relation among goals and the implications of favoring one goal in detriment of another, and it is also dynamic in the sense that goals are formed and altered given environmental information (Huffman, Ratneshwar, & Mick, 2000). Moreover, the goal choice strategy does not entail any information from the attribute space, since “if one decides to have a healthy dessert, one doesn’t necessarily first ask about which healthy desserts are available” (Dellaert et al., 2017 p. 4).

From the dynamic nature of goals and its associations with other concepts in memory, it results in task and context contingency, leading goals to rest upon a process of learning and forgetting (van Osselaer et al., 2005). Thus, these authors point to accessibility, which depends on frequency and recency of activation, as an important mediator between goals and behavior. They enumerate the following sources of consumer goals activation: (a) direct priming of the goal or of the consumption benefit; (b) priming of other goals associated with the focal one, which will also inhibit the activation of substitute goals; (c) spreading activation from means, i.e., activation of goals or behaviors that are a mean for attaining the focal goal; (d) task and contextual cues; (e) previous goal achievement, through the increasing of relative importance of goals not attained.

The multi-attribute space, on the left side of Figure 5, is where it rests the choice task available to the consumer, with the alternatives described by its attributes. Notice that this is the multi-attribute space as described by the economic theory of consumer behavior and in the traditional econometric choice models. Thus, the *homo economicus* may found support for making choices under expected utility maximization in this space, although, even here, other behaviors are also possible.

Given the goal choice strategy, the goal evaluation strategy informs the mechanics of mapping the alternatives into goals, involving attainment functions that use all or part of the information in the attribute space. Notice that the dynamic nature of the goals choice strategy is represented by the bi-directional arrows between the goal and the attribute spaces, denoting that the goal choice strategy influences how the attribute space will be used, but it also can be updated because of the properties of the choice task or the evaluation itself. In other words, goal choice strategy may impose that an alternative achieves a minimum level in a specific attribute or combination of attributes or even that a subset of alternatives will not be considered. In the other hand, the mapping of attributes into goals may determine that a goal is relaxed or another one is activated or deactivated, as a result of the multi-attribute space properties, accommodating the manifestation of context effects.

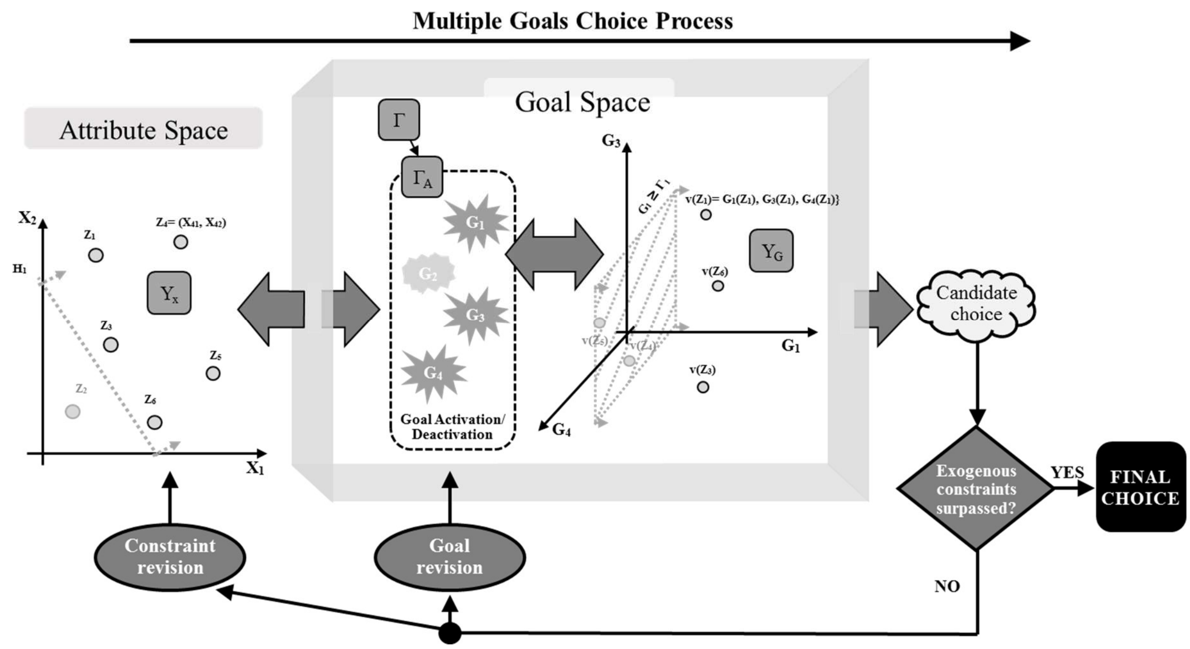

Figure 5 - Multiple Goals Choice Based Process
Adapted from: (Dellaert et al., 2017 p. 4)

This interplay between the attribute space and the goal space is compatible with the framework proposed by Huffman, Ratneshwar, & Mick (2000) and the dynamic view of goal determination is driven by two main psychological levers. Firstly, goal alignment is the process through which goals, at different levels of the hierarchy, influence each other to produce consistency and congruence. It can be a top-down process, called incorporation, with higher level goals shaping and giving meaning to lower level goals and to the attribute space. And it can also be a bottom-up process, named abstraction when the preferences revealed in the attribute space leads to a re-signification of the higher-level goals. Secondly, adaptation is the process by which the context and task configuration update the goals. It is important to notice that this dynamic process, permitting task and context effects to manifest at any point in the choice process, accounts also for nonconscious goals (Chartrand, Huber, Shiv, & Tanner, 2008), including the activation of unaware goals by environmental cues, the nonconscious pursuit of these goals and the effects on consumers' choices and expressed preferences. The authors also report that, even unaware of goal activation and pursuit, consumers are aware of engaging in the behavior and usually choices are the same, regardless the conscious or unconscious goal activation.

Finally, the choice is driven by the goal choice strategy that specifies how the goals' attainment will be combined to produce a preferred alternative. Thus, the influence of the attribute space in choice is only indirect, since it informs about goal attainment through the goal evaluation strategy. Once a preferred alternative is defined from the goal choice strategy it will be evaluated against any eventual exogenous constraints that could prevent the final choice. Suppose that the consumer is considering the purchase of a new car and this process produces a candidate choice. But the total cost, involving the car's price, the insurance and legal and financial expenses are above a consumer's budgetary restriction. Then the process must be restarted, and either the goal or the attribute space will be accessed through goal revision or imposition of new attribute constraints. The substantive interpretation of my adaptation is the same of the original authors, but for sake of clarity, I've made it explicit in Figure 5 through the representation of the preferred alternative as a candidate choice that will become final if there is no exogenous constraint to prevent it.

In this multiple goals choice based process the procedural rationality is given by the effectiveness of the goal evaluation strategy as the support for the attainment of the desired end states specified by the consumer's goal choice strategy. The goal activation processes and the relation between goals and the attribute space, through psychological goal alignment and adaptation, give room to behaviors that are contingent on task and context characteristics, demanding the skills of the *homo aptabilis*. To conclude, if the contingent behaviors prevent normative or economic rationality, the described multiple goals choice based process supports the procedural rationality observed in the *homo aptabilis* behavior.

Now it is time to expand the understanding of the goal space, examining the different types of goals, the processes involved in goal activation and deactivation and some empirical evidence to support the idea of a multiple goals choice based process.

I have already presented, from the relevant literature, goals as desired end states, linked to positive affect, that motivates action and that are represented through knowledge structures.

Additionally, the literature suggests some categories of goals that help to understand the phenomenon; van Osselaer et al. (2005) propose that goals can be consumption, criterion or process goals.

3.1 Consumption goals

Consumption goals are benefits afforded by the consumption of products, derived from a combination of attributes. Dellaert et al. (2017) suggest this kind of goal to be termed functional goals since they are directly related to attributes performance. The tastiness or freshness of a food or drink result from their ingredients combination, as well, the safety or comfort of a car result from specific characteristics as body style, internal space, the presence of specific equipment and others. Notice that the same attribute may contribute to the attainment of more than one goal, e.g., the power of the engine contributes to the sportiness of a car when it allows for speed and it also supports the attainability of safety, since it requires less time and space if an overtake is required. This possibility also reinforces the idea of context dependence, since the meaning of the attribute may vary as a function of the goal that is active.

Li (2013) developed a multiple goal based choice model for digital cameras using functional goals developed from exploratory research and used those goals to model the data from a discrete choice experiment. Some of the attributes were brand, price, resolution, and size of the LCD; and some of the functional goals were “keep up with new technology”, “take good quality pictures” and “have a reliable/durable camera”. The choices were modeled as a function of the previously identified goals and the attributes were used in an attainability function instead of a utility function, i.e., each alternative was valued accordingly to its capacity to attain each of the functional goals, which were used to predict choice probabilities. Moreover, the model accounted for the presence of latent classes to capture desired goals heterogeneity. To validate the latent classes the respondents also self-reported their goals when choosing digital cameras and the author demonstrated the correspondence between latent classes and self-reported goals. Moreover, a latent class model accounting for heterogeneity in the utility function was fit to the data, as described in Kamakura & Russell (1993). It means that this last model did not account for goals but choices were predicted (as in the traditional models) only by the utility function that accounts for taste heterogeneity. The multiple goals choice based model outperformed the latent class model both in model fit and in out of sample prediction. Finally, Li’s model accounts for context adaptation, being a context variation defined as the inclusive value or the expected maximum attainability of each goal over all the options in the choice task (see Ben-Akiva & Lerman, 1985 for the definition of inclusive value). The propensity to adapt goals to the context was modeled as a function of individual variables, specifically the experience with the category. The adaption module improved the model fit, supporting the hypothesis that goals are context dependent and consumers adapt goals depending on the characteristics of the choice task.

Dellaert et al. (2017) propose the existence of non-functional goals, which are more abstract and higher order goals (or superior in a goal hierarchy), that can unfold through lower order sub-goals that can be functional. I understand that this kind of goal comprehends most of what van Osselaer et al., (2005) exemplify as criterion goals like the facility to justify a choice to other people or the desire to cause a good impression. These kind of goals are marks of the person-environment bond (Austin & Vancouver, 1996) and can be understood as related to the self-identity, in the sense they are a consequence of the individual interaction in the social world and, on the other hand, they drive the relations in this social world. (B. Simon, 2004 p.2). Despite its higher construal level, these goals are also derived from the attribute space and are, therefore, consumption related. As examples are the choice of premium brands or luxury products that may facilitate the attainment of a goal like status or the consumer's desire to feel different from the others. Chartrand et al. (2008) report a series of experiments using nonfunctional goals and have demonstrated that the activation of prestige and parsimony goals led to the subsequent choice of premium and value brands, respectively. Also noticeable was the nonconscious activation of these goals, through a game in which participants construed a grammatically correct sentence from a set of scrambled words related to prestige or to thrift, empirically confirming that awareness is not a necessary condition for goal based choice.

These nonfunctional goals may also be attainable as a combination of functional goals as the choice of a car with attributes related to comfort and safety could make the decision-maker to be viewed, by others, as a family oriented individual. In summary, multiple goal based choice models may be grounded in a consumption goal category, which contains functional and non-functional goals, and functional goals tend to be lower level ones while non-functional goals tend to be higher level.

3.2 Process or meta-goals

Following Bettman et al. (1998) I will refer to the second category of goals as meta-goals, although some authors name this group as process goals (Dellaert et al., 2017; van Osselaer et al., 2005). The defining characteristic of this category is that the desired end state relates to the choice process itself, instead of being a consequence of the product being chosen. The main meta-goals in consumer behavior are: (a) to maximize choice accuracy; (b) to minimize cognitive effort deployed to make a choice; (c) to minimize negative emotion experienced during the choice process; and (d) to maximize the ease of choice justification (Bettman et al., 1998).

The economic consumer behavior theory assumption is that the meta-goal is to maximize accuracy since the subjacent psychological process supporting decision-making is to choose the alternative with the maximum expected utility. One important detail is that accuracy is a subjective criterion, not observed by the analyst, which means that even if two consumers chose different alternatives in one choice task they can both be maximizing accuracy. Subjective criteria have also been used with an ideal alternative subjectively formulated through compositional methods, i.e. that consumer states the desirability of every attribute, and the following choice is observed and compared with the expected one (see Hahn, Lawson, & Lee, 1992; Malhotra, 1982). Other researchers have used objective criteria, making information available to the consumer before the choice process, to activate a social norm (Lurie, 2004; Malhotra, Jain, & Lagakos, 1982; Scammon, 1977)., Some of the used information sources were the product itself, specialized media or a socially relevant source (to the consumer)

I will examine with details the effort-minimization meta-goal given its importance in many streams of the choice literature and because it motivates, or at least cause the effects that motivate the empirical section of the thesis. I will also introduce the negative emotion minimization since it motivates the fifth chapter section of these thesis. And in that empirical section I will explore this meta-goal with more detail.

3.2.1 Effort minimization

To minimize the cognitive effort involved in any choice process is a goal derived from information processing theories that acknowledge that the human brain has limited storage and processing capacity (H. A. Simon, 1955, 1990). Accordingly to this perspective, a choice process aims to result in a decision that is good enough to the consumer while not depleting the limited cognitive capacities. Dar-Nimrod, Rawn, Lehman, & Schwartz, (2009) developed and validated a psychometric scale showing that propensity to maximize or to satisfice, i.e. to choose a good enough alternative, is an individual characteristic and presented experimental evidence that the satisfiers tend to be more satisfied with their choice than the maximizers. To minimize cognitive effort individuals deploy heuristics, defined as mechanisms that reduce the complexity of evaluation of probabilities and attribution of values, in a decision under uncertainty, to more simple judgmental operations (Tversky & Kahneman, 1974). From an information processing perspective, Payne (1982) identifies three theoretical frameworks to explain how individual decision-making responds to context and task properties: (a) cost-benefit principles; (b) perceptual processes; and (c) adaptive production systems. Although in many situations the two last frameworks may confound, in terms of predictions and decision rules, I will keep them apart because they start from different assumptions, especially in the motivation for the adoption of heuristics, and they are positioned as competitive frameworks.

a. Cost/benefit principles

The proposition supporting the development of this framework is that the individual's choice process is a compromise between the desired goals - that include higher choice accuracy, reduced decision-making time and choice's justifiability, among others - and the necessity to minimize efforts involved in information acquisition and processing (Payne, 1982). It means that effort minimization as a goal is not separable from other goals during the choice process (Russo & Dosher, 1983) and that strategy selection is a higher level decision-making, that antecedes alternative selection, imposing rationality to the choice process even when the observed choice seems to disrupt such rationality. This is also consistent with a neoclassical optimization reasoning, which concludes that information will be used up to the point in which the expected marginal incentive equals its expected marginal cost of acquisition or processing (Stigler, 1961).

In a classical paper using the cost/benefit framework Shugan (1980) proposes a model in which the cost of thinking is an increasing function in the number of comparisons, performed by attribute, involved in the choice process. This number will depend on the utility difference in the attribute comparison between the focal alternatives, the desired accuracy of the process and the variability in the alternative differences. The larger the utility difference in the sampled attribute, the fewer comparisons will be required to produce a choice and, in the other hand, the higher the choice accuracy desired by the consumer, more attributes need to be compared. Finally, the less variability in attribute comparison, the less comparisons will be needed, i.e. it will be easier to choose an alternative if it tends to be systematically better than another in the attributes comparisons. One issue that arises from this model is that the computational cost is dependent on the alternatives being processed by attributes and, while there are many attribute based decision rules, it is not feasible for alternative based choice strategies.

In a recent choice modeling paper adopting a cost/benefit approach, Swait, Popa, & Wang (2016) developed a model proposing that "individuals are information managers who simultaneously solve the joint problem of deciding what information to use and which alternative to choose" (p. 647). There are two functions, one to model the benefit of including every attribute and other to model the cost of using information, allowing for the assessment of a net benefit of using each possible attribute and every combination of attributes. The benefit of using a piece of information rests on its diagnosticity, which is related to its own range in the choice task. It means that the larger is the range or the variety of attribute levels across the alternatives in the choice task the larger is the benefit of using the information. The cost of using the information is a function of the number of attributes in use, the number of alternatives in the choice task and the interaction between these dimensions. The model accounts for the simultaneous presence of an archetype that always uses all the information, as expected from the economic consumer behavior theory, and an adaptive archetype to whom the information usage is contingent to the context. Across four empirical applications, the authors report selective usage of information even in simple choice environments. When presented with choice tasks with five alternatives and only two attributes, 35% of the consumers have used information selectively, adjusting behavior to the context. This figure grows up to 71% of consumers managing the information used in the choice process when the choice tasks have nine attributes and the number of alternatives varies from two to eight within subjects.

b. Perceptual processes

This stream of knowledge is strongly based on the research from two prominent psychologists who proposed that the lack of invariance ruling individual choice behavior can be explained by elementary principles of human perception. Daniel Kahneman and Amos Tversky are important sources of knowledge in all scientific disciplines involved in the judgment and decision-making arena, and they are so important to the field that the first won the Economics' Nobel Prize in 2003. A detailed discussion of the theories, and the empirical support, proposed by these researchers is presented in Pilli (2012), and here I just outline a brief overview.

During decision-making, individuals evaluate expected changes in well-being in opposition to the expected end states as defined in the utility theories. The choice process unfolds in two stages, the first one being the edition of the available alternatives to simplify the information. This phase offers the final representation of the alternatives and it is dependent on the problem structure, i.e. how the information is presented and the editing sequence, since the well-being change evaluation depends on a reference point. In the second phase, which is the evaluative one, individuals are more sensitive to losses than to gains, i.e. they are loss avert, what causes a kink in the function around the reference point. The problem structure constrains the choice process inducing alternatives' evaluation on the most accessible characteristics, accessibility meaning that some thoughts come easier to mind given priming, stimuli salience, attention, training or associations (Kahneman, 2003; Kahneman & Tversky, 1979). Additionally, the reference point adopted by a decision-maker depends on her norms, habits, and characteristics (Thaler & Sunstein, 2008; Tversky & Kahneman, 1981).

These theories are grounded in a dual system psychological model defined in Woodworth (1938), as cited by H. A. Simon, (1990). System 1 is perceptual and engaged in problem-solving, and it has the qualities of being fast, emotional, effortless, parallel, associative and slow-learner. System 2 is reasoning and it describes the thought processes through propositions and logical manipulations, it is rule-based, controlled, slow, serial, flexible, effort based and neutral. Kahneman, (2003) describes a process of intuitive judgment, resting between perception and reasoning, that is a perceptual function processing stimuli that would better fit the reasoning function. This intuitive judgment is grounded on the bounded rationality working mainly through heuristics, defined as the processes that reduce the complexity of decision-making under uncertainty to simpler judgmental operations that are "quite useful, but sometimes they lead to severe and systematic errors" (Tversky & Kahneman, 1974 p. 1124).

In the choice modeling literature, the Categorization Generalized Extreme Value Model (CatGEV) presented in Swait, Brigden, & Johnson, (2014) is an example of an empirical application, in marketing, of the concept of intuitive judgment and, importantly, of the goal based choice as it can be noticed by the statement that "categorization is an essential precursor to evaluation and choice because understanding what type of object we are dealing with determines what the object is useful for and, therefore, which attributes are important" (p. 4). Categorization is a first phase that edits all the alternatives, by coding, creating a simpler representation of the choice task, before evaluation. This model allows the attribute preferences to be homogeneous and conditioned only by the categories, which are unobserved by the analyst. In other words, there are as many preference vectors as there are mental categories, but they are homogeneous across consumers or occasions. The choice process involves, firstly, the attribution of the alternative to one available category and then, the evaluation through the proper attribute preference vector. In one of the empirical applications, the authors used candy bars to deploy the CatGEV model. The authors selected, in a pilot study, five chocolate bars, five protein bars and one candy bar ambiguously positioned between the two categories. A stated preference experiment was conducted in three cells, and in the control cell the initial priming was the presentation of all the bars together and randomized before starting the exercise of the choice tasks. In the two manipulated cells, chocolate and protein bars were present in different columns of the priming stimuli, such that in one cell the ambiguous bar was primed as the first bar in the chocolate column and in the other cell it was primed as the first bar in protein column. The model could recover two categories with different attribute preference vectors, and the categorization of the ambiguous candy bar was a function of the priming. At the end, the model supported the idea of a two phases choice process, the first phase being perceptual and forming the final representation of the choice task for the evaluative second phase. As predicted by Kahneman and Tversky's theories, priming was enough to change the reference used for the evaluation by the alternatives.

c. Adaptive production systems

This framework is supported by the definitions originated in artificial intelligence system and it departs from perceptual framework by disagreeing that heuristics tend to result in systematic biases. The motivation is that, considering the conditions for normative rationality are nonexistent, there is no sense in evaluating the results produced by heuristics with normative expectations. Furthermore, heuristics are defined as decision strategies that ignore information aiming at speed, parsimony and accuracy and the definition is agnostic about the effectiveness of heuristics usage. Actually, in the real world, they often perform as well or better than eventual normative models and, in this sense, heuristics' usefulness is defined by its ecological validity, which results from its adaptation to the environment (Dana & Davis-Stober, 2016; Gigerenzer & Gaissmaier, 2011).

To understand the origin and the motivation of this framework I count on some ideas presented by Newell & Simon (1976), in the tenth Turing Lecture at the Association of Computer Machinery, starting with the assertion that "we measure the intelligence of a system by its ability to achieve stated ends in the face of variations, difficulties and complexities posed by the task environment" (p.114). Firstly, this is, in other words, the procedural rationality oriented towards goal pursuit proposed to be the behavioral character of the *homo aptabilis*. Additionally, the limitation in cognitive ability relates to the fact that, given a constrained number of steps and time, only a limited number of processes can be executed, demanding heuristic search, i.e. the generation of possible solutions and the test of their goal attainment potential. The key to the success of such a system rests on its ability to be selective, in the sense of generating only promising solutions, and to properly identify the gaps between the current and the desired states following up with behaviors that decrease this difference. At this point we get to the concept of an adaptive production system that means the individuals have the ability to generate and to learn several decision rules and to select the most appropriate accordingly to the environmental features, or in the authors' words:

> If the system had some control over the order in which potential solutions were generated, then it would be desirable to arrange this order of generation so that actual solutions would have a high likelihood of appearing early. A symbol system would exhibit intelligence to the extent that it succeeded in doing this. Intelligence for a system with limited processing resources consists in making wise choices of what to do next. (p.121)

Consistent with these ideas, Gigerenzer (2007 p. 33-35) reports experiments with golf and handball players, where the former had to execute plays while the latter had to watch games and predict the best next movement when the game was frozen. The experienced players perform better, in both games, under time pressure while the novice ones perform better having time to think. The experiments support the idea that experience builds a repertoire of solutions, and that the best solution for every situation is more likely to be activated earlier than the others. The more time the expert has, the more information will be considered and the higher will be the number of solutions coming to mind, reducing the likelihood of the best one being selected. Since novices have not developed a sophisticated production system, the more time for deliberation, the better the solution. Another idea behind these results is that intuitive judgments become more effective with training, and once sophisticated behaviors are internalized they do not require consciousness or deliberation to be performed. In summary, a set of potential solutions is made available by learning, and once it happens the selectivity in the use of information drives to the best outcomes.

A detailed review about heuristics proposed within the adaptive production systems framework is presented in (Gigerenzer & Gaissmaier, 2011). Now, I just describe the main heuristics, among those identified by the authors, to illustrate their impact on consumer behavior. The main heuristic classes are based on recognition, on one-reason-only, and on trade-offs.

The recognition based heuristics consider recognition only cues and the two main exemplars are the recognition heuristic, which predicts that “if one of two alternatives is recognized and the other is not, then infer that the recognized alternative has the higher value with respect to the criterion” (p. 460) and the fluency heuristic predicts that if both alternatives are familiar the first to be recognized has the higher value. In marketing, the recognition heuristics explain phenomena like choice set formation based on familiar brands instead of quality cues evaluation.

The one-reason-only is a category of heuristics in which only one attribute bases the judgment or preference and other cues are ignored. The take the best heuristic predicts that only the most important attribute will be used, consistent with a lexicographic process (Bettman et al., 1998; Einhorn, 1970).

The trade-off is the class of heuristics based on compensatory strategies, but equally weighting attributes or alternatives. Tallying is the heuristic through which consumer compares alternatives and code advantages and disadvantages and the one counting more advantages wins. Attributes are processed until one alternative emerges as the winner. The 1/N is the heuristics that equally allocates resources among alternatives and it is compatible with the diversification of an investment portfolio or with variety seeking in consumer behavior.

3.2.2 Negative emotion minimization

Emotion is an automatic psychological process derived from a primary appraisal that rises the perception and identification of risks in decision-making, which may be related either to the possible hazard to a relevant goal or to gains, which trigger positive emotions, or losses, which trigger negative emotions, resulting from the choice process. Moreover, the emotion comprises also a secondary appraisal of the alternative action available to cope with it. Coping is a response, including cognitive processes, aiming to regulate emotional states. The main identified types of coping are: (a) to plan or to approach the situation eliciting the emotion to find a solution for it; (b) to positively reappraise the situation trying to extract a positive meaning from it, (c) to aggressively confront the situation, and (iv) to take distance from it (Folkman & Lazarus, 1988; Lazarus, 1991). M. F. Luce, Bettman, & Payne (1997) define the type (a) as problem-focused coping, which encourages the individual to solve the situation arising the emotion, and the other types as emotion-focused coping, "or indirect actions intended to minimize experienced emotion through changes in (only) the amount or content of thought about the emotion-eliciting situation" (p. 387).

At the first stage of the emotional process, i.e. the one resulting from the primary appraisal, the choice process causes emotional conflict since it requires the consumer to renounce the unchosen, and this conflict may be magnified in specific circumstances, for instance, when both alternatives elicits negative emotions or when the trade-off involves relevant attributes, consequences or goals (Botti & Iyengar, 2006); when moral consideration are demanded (Bettman et al., 1998); and when negative inter-attribute correlation increases choice difficulty or when choice consequences are more salient due to task effects, like visual stimuli (M. F. Luce, 1998).

In the realm of consumer decision-making, the primary appraisal evolved to regret, that is, a result from the comparative evaluation, in the utility function, of the chosen and the nonchosen alternatives, such that if the nonchosen option happen to be better the consumer feels regret, which is a negative and cognitive based emotion, otherwise she feels rejoice. The definition can be generalized assuming that it emerges if the consumer realizes or imagines that the current state could have been better had the choice been different. Now, regret can be about the past or future, action or inaction and, choice process or choice outcomes (Loomes & Sugden, 1982; Pieters & Zeelenberg, 2007; Zeelenberg, 1999; Zeelenberg & Pieters, 2007).

3.2.3 Behavioral rules

Behavioral rules are quite intimate to the meta-goals and, in this sense, it is quite tricky to address it as a separate topic. Nevertheless, it is important to do so in order to define the goal based choice process and to present some additional contributions from the literature.

Remember that a multiple goals based choice process comprises a goal choice strategy, defining active goals and attainment criteria, and goal evaluation strategy mapping the goal choice strategy into the attribute space. The selection of a goal choice strategy, a cognitive structure sitting in the associative memory, should activate the behavioral rule, enabled by constraints and/or objectives, that produces the highest likelihood of goal attainment. Notice the similarity of this description with the procedural rationality that supports the *homo aptabilis*. Different decision rules will result in different outcomes and Pilli (2012) explores a set of decision strategies and their different predictions. Swait & Marley (2013) outline the model for a series of behavioral rules, including exploitation and exploration of the choice scenarios, expected satisficing, random regret minimization, and adherence to other preferences. Most importantly, they provide the optimization framework to incorporate the multiple goal based objectives into a probabilistic choice model.

A multitude of behavioral rules have been proposed by BDT and many of them have been adapted in DCM. Besides being sources of heterogeneity themselves, most of these rules imply in choice set formation, another source of heterogeneity in consumer choice process. This will be formalized with more details in section II, but I will present some important evidence from the RUM literature supporting the heterogeneity in behavioral rules. Obviously, all the examples presented when developing the concepts of meta-goal included behavioral rules and I will only bring new studies at this point.

Gilbride & Allenby (2004) presented a two-stage choice model, the first one being a selection stage and the second one being an evaluative stage consistent with random utility models. Three different behavioral rules were used for selecting alternatives, the first one being a conjunctive rule implying that an alternative must accomplish each criterion from a set of conditions (Dawes, 1964; Einhorn, 1970); the second one a disjunctive rule that determines that an alternative must accomplish at least one criteria (Dawes, 1964), and the third one a compensatory rule that selects alternatives values at the systematic portion of the utility function above a threshold, which is compatible with a satisfying choice process (H. A. Simon, 1955, 1990). The authors report a stated preference study with photographic cameras in which 92% of the consumers used some of the screening rules, and that screening is usually supported by well-known attributes while others are used at both stages of the choice process. In another study, Gilbride & Allenby (2006) used three different behavioral rules to support the screening of alternatives and the comparison against utility maximization. The conjunctive rule was used again; an elimination by aspects model, implying a sequential process that eliminates alternatives bellow a threshold defined by attribute until a choice occurs (Tversky, 1972a, 1972b); and a two-stage economic screening criteria that uses attribute cutoffs to exclude alternatives from the second stage based on a cost-benefit approach (Bettman et al., 1998; Payne, 1982). The authors ran four independent models to understand consumers' preferences for documentary films and they found that the conjunctive and the economic screening rules performed better in fit and predictive power than the utility maximization, while the elimination by aspects model did not perform well. The model is agnostic about meta-goals, although these behavioral rules fit the cost/benefit principle, except for EBA that is associated with perceptual processes. Notice, that all of these rules simplify the choice task through some kind of selectivity process. However, given that the econometric model is compatible with the behavioral model but does assure the subjacent psychological process, more information is needed to assure that the model is describing the expected behavior. To mention examples that explicitly tried to tie the econometric model to the assumed psychological process, remember that Swait et al., (2016) explicitly modeled a cost-benefit function to identify full information users from information managers, and Li (2013) collected self-reported goals to validate the latent goals identified in her model.

Perceptual processes have also been introduced to represent dual stage choice processes in choice models. Hensher & Greene (2010) developed a latent class model, in which the classes are defined by the use of different heuristics based on perceptual process and attribute non-attendance was identified as a component of the choice process used by more than 50% of respondents. These authors also advocate the need for external data to validate the psychological process. Working around the perceptual processes paradigm, think of Swait et al., (2014) that manipulated experimentally the focal product to establish the causality between the behavioral rule and the perceptual process.

There are also some examples of heterogeneity in behavioral rule modeled in RUM using revealed preferences. For instance using scanner data to model consumer choices in the peanut butter category, loyalty behavior was represented in a choice model capable of identifying hard-core loyal consumers, who always choose the same *SKU*, brand-loyal consumers, and product form loyal consumers, besides utility maximizers. The authors reported that 14% are hard core loyal, i.e., adopting a very simple decision rule that is “repeat the last choice” and 24% of the household are utility maximizers, i.e. behaviorally compatible with the *homo economicus* (Kamakura et al., 1996).

Adamowicz & Swait (2012) have also used scanner data for 25 consumer packaged goods and developed a model to account for utility maximization, pure habit, and variety seeking behaviors. They report results for two categories, concluding that the likelihood of habitual purchase is higher for catsup, while for yogurts consumers are more variety-seekers. But they also report a high incidence of households, 65% in catsup and 90% in yogurt, behaving as utility maximizers.

The revealed preference data has the great advantage of being real behavior, but it does not allow to validate the hypothetical behavior with additional information. Nevertheless, these studies support the idea of decision strategies heterogeneity and confirm that utility maximizing is a relevant behavior, with its deployment being context dependent.

Finally, another class of models, that I will use and detail in section III, is random regret minimization (Chorus, 2010, 2014; Chorus, Rose, & Hensher, 2013; Hensher, Greene, & Chorus, 2013). This is an econometric choice model replacing the utility function by a regret function that emerges from the differences in attribute levels between the focal and every other alternative in the choice task. This model accounts for context effects like attraction and compromise and the empirical tests support a superior model fit than the utility maximization. However, it is still missing a validation of the psychological process and the identification of the environmental conditions to activate the decision rule.

4 EFFECTS OF CONFOUNDING CONSUMERS' CHOICE PROCESS HETEROGENEITY AS TASTE HETEROGENEITY ON THE FIRM DECISION MAKING

4.1 Introduction

This project studies the effects on the firm's performance of misattributing process heterogeneity to tastes as a function of its monopoly power and of the extent of choice process heterogeneity on the demand side.

In marketing, the identification of consumer response heterogeneity is the foundation for segmentation (where to compete) and positioning (how to compete). This implies that the strategic decisions at the marketing level are outcomes of the policy makers' understanding of consumers' preferences that would shape the market structure and competition patterns (Kamakura et al., 1996). The misunderstanding of consumers' preferences and demand structure make firms' performance vulnerable as the statistics on new product failure clearly demonstrate. In accordance to Nielsen's Breakthrough Innovation Report Europe 2014, 76% of new product launches in Western Europe did not achieve 52 weeks of sales and almost 50% did not survive the 26th week.

The basic proposition of this paper is that misattributing choice process heterogeneity to taste is likely to lead firms to be driven away from profit maximization through: (i) first, a wrong understanding of demand will be expressed via the biased location of the utility parameters, and also through the identification of variance in the taste parameters, supporting the existence of heterogeneity in the demand that is not actually present in the marketplace; (ii) second, the biased demand understanding will lead to the inference of wrong policy measures, namely the firms' choices probabilities and the demand attributes' choice elasticities; (iii) then, firms will choose wrong quantities to offer in the marketplace from the biased choice probabilities information and will operate with a suboptimal marginal cost caused by the biased reading of demand price's choice elasticity.

To study the effect of misattributing process heterogeneity to preferences on the firm's strategic decisions, the project unfolds in two steps. In the first step, a demand representation is built from a Monte Carlo simulation based on knowledge of the true data generation process, absent taste heterogeneity. As a result, a true and a misspecified representation of consumers' preferences and market structure will be used to study the market equilibrium resulting from a game theoretical approach. The severity of the consequences will be studied as a function of characteristics that may be observed either on the firm side or on the demand side.

Against the presented motivation, this section of the thesis will be developed as follows: after this introduction I will present a brief sub-section (second) of the relevant literature about choice set heterogeneity, remembering that many of the required concepts are already detailed in the first section of this document. In the third sub-section I will detail the design of the Monte Carlo experiment executed to build the true and the biased demand representation. In the fourth sub-section I will present the results of the demand analysis, including a Monte Carlo error and a bias surface analysis. In the fifth sub-section I will present the effects of misattributing choice set heterogeneity to taste heterogeneity in the market equilibrium using a game theoretical approach. Finally, in the sixth sub-section I will discuss the results.

4.2 Literature review

The choice modeling literature generally attributes heterogeneity in multi-attribute decision-making to different tastes among people, i.e., random utility models assume that people have dissimilar levels of preference for the various attributes or attribute levels involved in the choice but follow a similar behavioral rule, usually random utility maximization, to select one option (Rieskamp et al., 2006). The emphasis on preferences as a source of consumer response heterogeneity is evident in this excerpt from Allenby & Rossi (1998):

> The purpose of marketing is to understand consumer preferences and to help design and deliver appropriate goods and services. Marketers are interested in determining what products to offer, what prices to charge, how the products should be promoted and how to best deliver the products to the consumer. One of the greatest challenges in marketing is to understand the diversity of preferences and sensitivities that exists in the market. Heterogeneity in preferences gives rise to differentiated product offerings, market segments and market niches. Differing sensitivities are the basis for targeted communication programs and promotions. As consumer preferences and sensitivities become more diverse, it becomes less and less efficient to consider the market in the aggregate. (p. 57-58)

However, the behavioral decision theory literature reveals that choice is a complex process driven by individuals' goals and the necessity of balancing the availability of cognitive resources and the requirements of the task. Thus, prior to the resulting observed choice, unobserved latent decisions have been made. These decisions may involve balancing multiple goals, selecting different behavioral rules and choice set formation among other possibilities (Bettman et al., 1998; Dellaert et al., 2017; Swait & Feinberg, 2014; Swait & Marley, 2013; van Osselaer et al., 2005; van Osselaer & Janiszewski, 2012).

The result is that different individuals follow different choice processes that necessarily shape the multi-attribute choice, i.e., process heterogeneity precedes taste heterogeneity. This perception is supported by Desarbo et al. (1997), who define heterogeneity as "result of the individual differences consumers evince with respect to the judgments they make and the processes involved in making such judgments".

In the thesis's literature review I have identified a variety of empirical choice models acknowledging the existence of different sources of consumer response heterogeneity and demonstrating that these models tend to perform better than the ones which attribute heterogeneity only to tastes. The empirical applications comprised both stated and revealed preferences in the exploration of heterogeneity in the choice process and some of these papers were developed considering specific behaviors that are quite meaningful in marketing, like habit, variety seeking and loyalty (Adamowicz & Swait, 2012; Kamakura & Russell, 1993). Other papers implemented decision rules like conjunctive, disjunctive and elimination by aspects proposed in the descriptive theories of judgment and decision-making originated in psychology (Gilbride & Allenby, 2004, 2006). I have also reviewed some papers which proposed that the diversity of decision rules applied on the information available in the choice task emerges either due to cost/benefit consideration (Swait et al., 2016) or as a result of perceptual processes (Hensher & Greene, 2010).

A shared point of view present in the referenced papers is the heterogeneity in choice sets, which results from consumers adopting a dual stage choice process implying, firstly, in reducing the alternatives in the choice scenario and then choosing among the remaining alternatives. This idea is consistent with behavioral decision theories that propose that selectivity is essential to goal based choice resulting from selective attention that operated at the lower level of perceptual processes as well as at the higher level of cognitive processes (Weber & Johnson, 2009). It is also compatible with the marketing construct of consideration set defined as “brands that the consumer considers seriously when making a purchase and/or consumption decision” (Hauser & Wernerfelt, 1990 p. 393), which also imply a dual stage choice process which first reduces the available alternatives for a subsequent evaluation that results in choice (see Chakravarti & Janiszewski, 2003; Kardes, Kalyanaram, Chandrashekaran, & Dornoff, 1993 for a review in variables driving consideration set definition). Swait & Feinberg, (2014) propose that the consideration set is formed among all the alternatives in the universal set of alternatives, *M*, regardless its presence in the choice scenario, meaning that it is a more stable and attitudinal construct, while the choice set is composed by alternatives that have a structurally non-zero probability of being chosen in the purchase occasion. To illustrate, a specific alternative included in the consumer’s consideration may be unavailable in the choice scenario, not being included in the choice set. It may also be part of the consideration set but not be evaluated in a particular choice occasion due to individual, contextual or task characteristics. On the other hand, the choice set is precisely defined as the subset of alternatives available in the choice task considered by the consumer in the specific occasion, formally defined as (Manski, 1977; Swait & Ben-Akiva, 1987):

$$Pr(j \mid B, D, X_n) = Pr(j \mid C, B, X_n) Pr(j \mid C, D, X_n) \tag{4-1}$$

Where j is an alternative, B and D are vectors of parameters, $\mathrm{X_n}$ is a vector describing attributes of the alternatives and individual characteristics, M is the universal set of alternatives, M_n, is the set of feasible alternatives for individual n, C is a choice set and G_n is the set of all nonempty subsets of M_n. The model expresses a choice process in which given an individual universal set of alternatives, a first stochastic process of choice set formation $\mathrm{Pr(j \mid C, B, X_n)}$ is followed by a second discrete choice model $\mathrm{Pr(j \mid C, D, X_n)}$ resulting in all probabilities different from zero.

The identification of the correct consideration set accounts for 78% of the explained variance in probabilistic choice models (Hauser, 1978). Li, Adamowicz, & Swait (2015) used Monte Carlo experiments, in a context of environmental economics, and reported that after ignoring choice set heterogeneity both choice probabilities and welfare measures were severely biased.

Regardless of the variety of choice models investigating consumer response heterogeneity at different levels than preferences (e.g. goals, decision strategies, stochasticity), the adoption of these ideas as an alternative to the simplicity and parsimony of the representation of the *homo economicus* in pure random utility models, is still a challenge to the discipline (Dellaert et al., 2017; Hensher, 2014; Swait & Feinberg, 2014).

Nevertheless, choice models that do not account for process heterogeneity may improperly capture such heterogeneity at the taste level and this issue motivates the following empirical investigation.

4.3 Empirical research

This study progresses in two steps, starting from the development of two market demand representations that are later used to analyze the market equilibrium.

The two demand representations, emulating different market contexts, are portrayed in utility functions detailed by firm-specific constants, generic attributes, and firm-specific prices. One of the demand representations mimics a competitive monopolistic market that describes many fast moving consumer goods categories. The other one describes consumers' experiences with services which also is a very usual domain of consumption. Taking one of the firms as the focal point of analysis, an experimental design will allow studying the effects of misattributing process heterogeneity into taste as a function of the relative size of the focal firm's portfolio, the relative price elasticity of the focal firm, and the extent of choice set formation and how it distributes across firms. Choice set formation will be used as the expression of process heterogeneity since it is an outcome of other instances of process heterogeneity in consumer behavior (like goals, decision strategies or decision rules). Thus, demand will arise through a mixture of choice set formation rules, to be detailed. Furthermore, the true data generation process will have taste homogeneity, not heterogeneity. A mixed logit model, allowing taste heterogeneity, will be used to estimate the consumers' preferences and to identify the market structure. This model, as most of the models described in the literature, will account for taste heterogeneity but not process heterogeneity (here, choice set formation).

In the first step, the demand is built through a Monte Carlo simulation that supports the execution of this study through: (i) permitting the true data generation process to be known; (ii) allowing the robustness of the results to be tested through replication; and (iii) increasing generalizability of the results through the deployment of a multiple-cell experiment compared to an usual single cell real data application (Andrews & Currim, 2005; Li et al., 2015). The true data generation process based on the absence of taste heterogeneity and on the presence of process heterogeneity is induced following the experimental plan to be presented in the next section.

A mixed logit model (Louviere, Hensher, & Swait, 2000 p. 199-205) was fit to the synthetic data allowing the preferences parameter to vary by individual but not accounting for process heterogeneity. For each market context, two representations describing consumers' preferences and market structure were the outcomes of the simulation: (i) a true representation, known since we know the data generation process; (ii) a (wrong) representation arising from the choice model attributing all the heterogeneity in the data to tastes.

In the second step of the study, the demand representation will be used to study market equilibrium using a game theoretical perspective, with firms pursuing to maximize their payoffs, defines as the profits.

4.3.1 Demand data generation process

The demand representation is grounded on consumers making choices based on homogeneous preferences, however, due to the differences in the process of judgment and decision making, each consumer maximizes expected utility within a subjective choice set, leading to a different choice distribution than the one that would result if all the consumers were choosing among all the alternatives.

It is important to notice that we are holding preference homogenous to have a clear picture of the effects of choice process heterogeneity in the firm decision-making. The motivation to do so is that preference heterogeneity is a well-stablished feature of the random utility models, whereas choice set formation, although a well-known phenomenon, is not.

The two data generation processes were planned to reproduce characteristics of a competitive monopoly in the choice scenarios, which is a common competitive structure in the proposed market contexts, and some characteristics of the consumer choice process that may trigger the focal phenomena being studied, i.e., the misattribution of process heterogeneity to tastes.

The first market context represents a fast movable consumer good (*fmcg*), as many of the product categories found in the supermarket. The idea is that, at least in the short term, the environment is constant and the marketing managers have great control over it. It means that consumers find an unchanging assortment in different retail chains, either across points of sales or occasions, and the quality (as described by the levels of their attributes) and the price of the products are stable. The second market context represents many different service consumption experiences that may be exemplified through fast food chains, laundry services or personal services among others. The idea is that any firm plans to deliver a standardized service level, but the consumer experiences varying levels of quality across either occasions or points of sales.

The scenarios were composed by offers from three brands, each one presenting a different number of products (or alternatives). The product attributes (k_s) were: two dummy variables as firm-specific constants defining the three firms (one dummy variable omitted for identification); three different price variables, or firm-specific attributes; and two generic attributes allowing for product differentiation. An overview of this structure is presented in Figure 6.

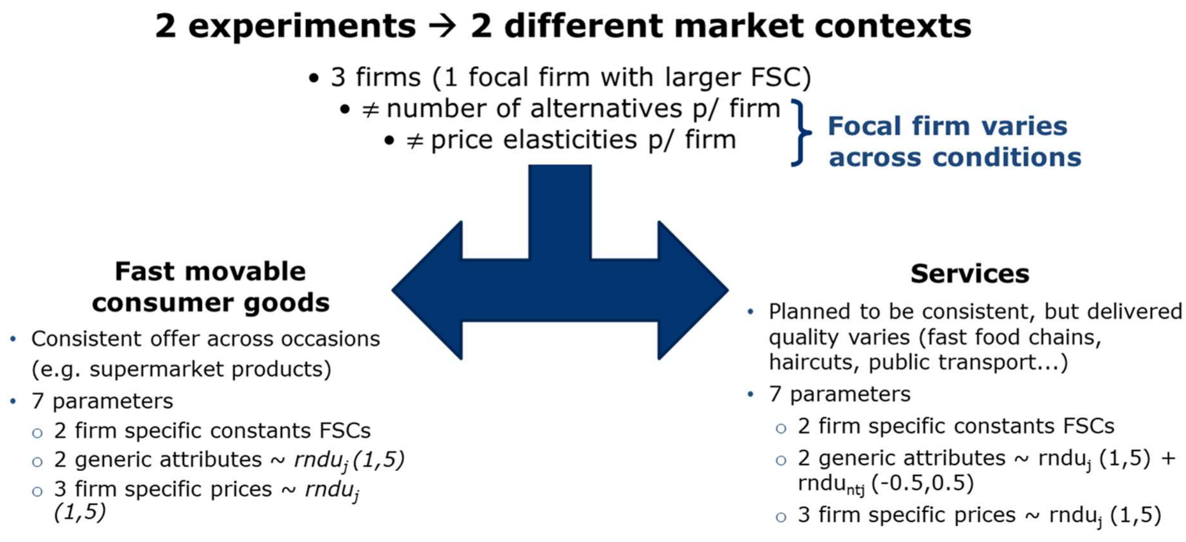

Figure 6 - Overview of the data generation process

The generic and specific attributes levels, for every product, are random variables such that $l_k \sim U(1,6)$. To represent the variation in the level of the service experienced by the consumer, in the second market context, a random variable $v_k \sim U(-0.5,0.5)$ was added to each generic attribute. The draw of l_k represents the level of quality the firm plans to deliver and the draw of v_k represents the variation that he consumer experience across occasions or points of sales.

Finally, notice that the firm-specific prices mean that each firm can set different prices, even for the same product, and that each one has its own demand price elasticities. This arrangement implies that each brand has its own demand curve, in which rests its monopoly power limit, and it can set the price that matches profit maximizing quantity. The true cross-elasticities were set to zero, implying no substitutability among the brands, what will make easier the interpretation of the modeled market structure. For the analytical purposes, firms will be named A, B and C and firm A will the focal one to this study.

The consumer decision-making process was based on a dual stage process, described in the marketing and consumer behavior literature, in which the use of non-compensatory decision rules to reduce the number of alternatives antecedes an evaluative compensatory final stage that results in an observed choice (Gilbride & Allenby, 2004; Hauser, 2014; Hoeffler & Ariely, 1999). At the initial stage, the simplification of the choice task is driven by factors like superior psychological process, individual characteristics, and contextual variables (see Weber & Johnsson, 2009 for a review). At the final stage, the remaining alternatives are fully evaluated in a compensatory fashion consistent with the random utility maximization described above.

Choice set formation (Swait & Ben-Akiva, 1987) is the concept adopted to operationalize multiple stages process heterogeneity, since any of the decision-making processes discussed in the literature will result in the elimination of some alternatives or in the inclusion of others in the evaluation phase ending in choice.

As illustrated in Figure 7, in the first stage of the choice process a consumer selects the firms that will have the products evaluated in the second stage. Since there are three firms, there are seven possible choice sets. In three of them, the consumer is captive to one firm only, meaning that the evaluation in the second stage will consider the products of firm A or firm B or firm C. Other three possible choice sets are the combinations of any two of the three firms. Lastly, any consumer may consider the products of all the three firm in the final stage of the choice process. Once the first stage of choice set formation ends, the products offered by the firms included in the second stage are evaluated through a decision rule that maximizes the expected utility of the chosen alternative.

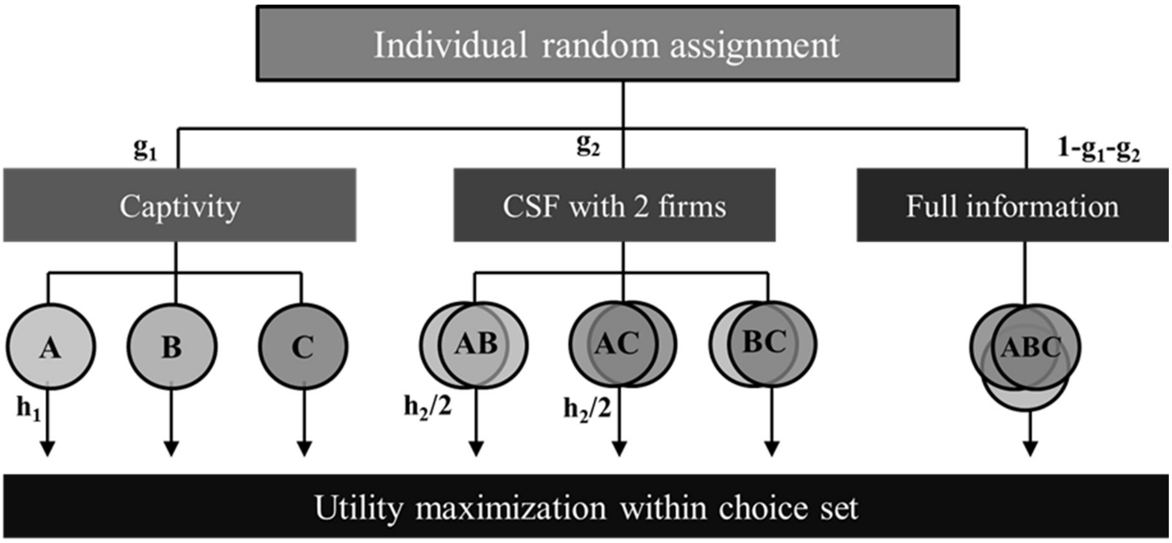

Figure 7 - Description of the true consumer choice process

Now let $0 \leq g_1, g_2 \leq 0.5$ and $0 \leq h_1, h_2 \leq 1$; *A, B* and *C* be three firm offering products in the marketplace and the utility of alternative *i* in each choice task (scenario) to be:

(4-2)

$$U_i = V_i + \varepsilon_i$$

Where:

(4-3)

$$V_i = \sum_{k=1}^{K} \beta_k X_{ikn}$$

$X_{i1n}, \dots, X_{ikn}$ are levels of *K* attributes describing alternative *i*, the subscript *n* applies only to the varying attributes in the service context, $\beta_1, \dots, \beta_k$ are *K* taste parameters that weight the attributes to form the evaluation of alternative *i*, and ε_i is an independent identical distributed stochastic term with Gumbel distribution;

Additionally, let $\widehat{P_i}$ be the expectancy of the MNL choice probabilities estimated from the true model (choice set heterogeneity, preferences homogeneity);

And, let $N\widehat{P_i}$ be the alternative *i* expected demand, such that.:

(4-4)

$$\widehat{\mathrm{NP_A}} = \mathrm{g_1 h_1} \sum_{\mathrm{i \in A}} \frac{\mathrm{e^{U_i}}}{\sum_{\mathrm{j \in A}} \mathrm{e^{U_j}}} + \mathrm{g_2} \frac{\mathrm{h_2}}{2} \sum_{\mathrm{i \in A}} \frac{\mathrm{e^{U_i}}}{\sum_{\mathrm{j \in A,B}} \mathrm{e^{U_j}}} + \mathrm{g_2} \frac{\mathrm{h_2}}{2} \sum_{\mathrm{i \in A}} \frac{\mathrm{e^{U_i}}}{\sum_{\mathrm{j \in A,C}} \mathrm{e^{U_j}}} + (1 - \mathrm{g_1} - \mathrm{g_2}) \sum_{\mathrm{i \in A}} \frac{\mathrm{e^{U_i}}}{\sum_{\mathrm{j \in A,B,C}} \mathrm{e^{U_j}}}$$

Equation (4-4) can be simplified to:

(4-5)

$$\widehat{NP}_A = g_1 h_1 + g_2 \frac{h_2}{2} \sum_{i \in A} \frac{e^{U_i}}{\sum_{j \in A,B} e^{U_j}} + g_2 \frac{h_2}{2} \sum_{i \in A} \frac{e^{U_i}}{\sum_{j \in A,C} e^{U_j}} + (1 - g_1 - g_2) \sum_{i \in A} \frac{e^{U_i}}{\sum_{j \in A,B,C} e^{U_j}}$$

And the expected demand for firms *B* and *C* are estimated in the same way, with the adjustments in the choice set formation parameters

Let's examine why the process just described is compatible with the theories proposed in the literature. Consider that process heterogeneity arises from differences in the individuals' goals when making choices. If the consumer is seeking for variety she will exclude the alternatives already chosen in the recent past from the choice set, making choices among the other ones. When the objective is to choose an alternative that is easy to justify, she eliminates those which seem to be unacceptable to the reference group and choose among the remaining.

When examining choices resulting from reference state dependence that in the marketing context can be expressed, for example, through brand loyalty or habitual purchases, the choice set will be formed among alternatives chosen in the recent past. Additionally, any decision rule described in the judgment and decision-making literatures, as satisficing (H. A. Simon, 1955), conjunctive (Dawes, 1964; Einhorn, 1970), disjunctive (Dawes, 1964), lexicography (Einhorn, 1970) or elimination by aspects (Tversky, 1972b), to name a few, will lead to a first stage when the universal choice set is reduced to smaller ones as the choice process unfolds.

Finally, let a mixed logit model represent the demand generated by the true data generation process, not accounting for the constrained choice set heterogeneity, but allowing for taste heterogeneity, such that demand for every firm is:

(4-6)

$$\widetilde{P_f} = \int_{\Theta} \frac{e^{V_f(x_i|\theta)}}{\sum_j e^{V_j(x_j|\theta)}} f\theta d\theta$$

This probability is a weighted average of a MNL probability, evaluated at different levels of θ (Train, 2009 p. 135), meaning that θ is now a distribution. Now the formula above may be rewritten as:

(4-7)

$$\widetilde{P_A} = \int_{B} \frac{e^{V_A(x_i|\widetilde{\beta_i})}}{\sum_j e^{V_j(x_j|\widetilde{\beta_j})}} \phi(\tilde{\beta}|\tilde{b},\widetilde{W}) d\beta$$

Where $\phi(\tilde{\beta}|\tilde{b},\widetilde{W})$ is the normal density with mean $\tilde{b}$ and covariance $\widetilde{W}$. In our case we are imposing the covariance to be zero and letting the diagonal of this matrix to be free parameters. Notice that the model does impose neither the distribution to be normal nor the covariances to be zero. It means that more flexibility is permitted by the model, but we are trading-off this flexibility (that could include different patterns of cross-substitution) to a simpler representation that allow the problem of the firm decision-making to be tractable.

In the special case of where g_1, g_2, h_1, h_2 are zero the random parameters model should recover the true parameters, including $\widetilde{W} = 0$.

Now, the theoretical difference between choice probabilities derived from the true model in expression (4-5), i.e. choice process heterogeneity and taste homogeneity, and the choice probabilities derived from the biased model in formula (4-7), i.e. choice process homogeneity and taste heterogeneity, is given by:

(4-8)

$$\begin{aligned} N\widehat{P_A} - N\widetilde{P_A} = \Bigg(& g_1 h_1 + g_2 \frac{h_2}{2} \sum_{i \in A} \frac{e^{U_i}}{\sum_{j \in A,B} e^{U_j}} + g_2 \frac{h_2}{2} \sum_{i \in A} \frac{e^{U_i}}{\sum_{j \in A,C} e^{U_j}} \\ & + (1 - g_1 - g_2) \sum_{i \in A} \frac{e^{U_i}}{\sum_{j \in A,B,C} e^{U_j}} \Bigg) - \int_{B} \frac{e^{V_A(x_i|\widetilde{\beta_i})}}{\sum_j e^{V_j(x_j|\widetilde{\beta_j})}} \phi(\tilde{\beta}|\tilde{b},\widetilde{W}) d\beta \end{aligned}$$

The proper derivation of this function could allow the study of the difference in the left side of the expression as a function of the parameters controlling the consumers' choice process. Given that the biased probabilities are expressed as a distribution, the differences could also be represented in terms of its mean and variance.

However, given the complexity of these functions, we conducted a Monte Carlo experiment to map these differences from the choice probabilities computed from the true model and from the estimated (mixed logit model).

4.3.1.1 Experimental plan

The experimental plan is the same for both marketing contexts and it involves characteristics of the choice task and of the consumers' choice process and it is explained from a general overview of the building blocks of the experiment, followed by the description of the variables' manipulation. The building blocks of the experiment were: (a) the utility parameters; (b) the number of SKUs in the choice task, attribute levels and the number of choice tasks; (c) the rules for choice set formation and the attribution of consumers to rules; (d) the experimental design.

a. Utility parameters

Resulting from the proposed demand representation the utility function has seven utility parameters. Two for firm-specific constants (one of the brands is omitted for identification), three for firm-specific prices and two for generic attributes).

The combination of firm-specific constants and firm-specific prices defines different demand curves for one product line of every firm, differentiating its monopoly power. While the firm-specific constants shift the demand curves upward or downward from what would be expected from a pure attribute evaluation strategy, the firm-specific prices change the slope of the demand curves, capturing the essence of monopoly power.

The utility parameters, for the systematic term of the utility function, are presented in Table 1. The firm A (focal brand) has the largest ASC parameter, and firm C is represented by the omitted dummy variable what means that its ASC parameter would be zero, and that firm

B has the smallest FSC. One of the parameters for generic attributes has a negative sign and the other one a positive sign, meaning that one represents a generic cost and the other a generic benefit.

Table 1 - Utility function parameters

Variables	ASCs		Generic attributes		Firm-specific attributes		
	A (focal)	B	I	II	A (focal)	B	C
Parameters β_k	0.5	-0.5	-1.0	1.0	-0.5 -1.0 -1.5	-1.0	-1.0

The only parameter that varies across the experimental conditions is the one related to the firm-specific price for the focal brand. This variable has three levels that represent the focal branding facing price elasticities lower, equal or higher than the competitors, meaning more, equal or less monopoly power. Moreover, the price parameters are negative for the three firms indicating that they face downward sloping demand curves

b. Number of SKUs, attribute levels, and choice task

The experimental design involved the definition of the number of alternatives in the choice set and the profiling of each alternative. The number of alternatives offered by the focal firm varies across the experimental conditions while the competitors offer a fixed number of alternatives, as described in Table 2.

Table 2 - Number of SKUs per firm

FIRM	**A**	**B**	**C**
Product line length	2 4 7 10	4	7

Firm B offers four and Firm C offers seven SKUs, allowing focal firm’s product line length to be: shorter than both competitors, equal to firm B but shorter than firm C, larger that firm B and equal to firm C and larger than both competitors.

It means that the smallest choice scenario in the experiment is composed by 13 alternatives and the largest one is composed by 21 alternatives. The attribute levels for the generic attributes and the specific prices, of the 21 SKUs were drawn as $x_k \sim U(1,6)_j$. These product profiles, defined as the universal choice set (*M*), are kept constant across experimental conditions and replications. Fixed product profiles are consistent with many fast movable consumer goods (fmcg) categories in which consumers face, at least in the short term, a constant set of product offerings.

In the service context the generic attributes were drawn as $x_k \sim U(1,6)_j + U(-0.5,0.5)_{entj}$. The quality levels, varying across consumers *n* and choice occasions *t*, are also consistent with the definition of services. The draw of the of the portion that adds variability to the quality level of the service is done for each experimental condition *e* from *1,…,E*. We acknowledge that the variability could have been included at the replication level, but then it could also be confounded with the disturbance term ε, described below, and to avoid this risk the decision was to keep it at the experimental condition level.

The fixed universal choice set is described in Table 3 and the last column is the value of the systematic part of the utility () for each product profile when the price parameter for the focal firm is set to -1.

Table 3 - The universal choice set (M) for the *fmcg* context

Alternative ID	**Firm-specific constant**		**Generic attributes**		**Firm-specific prices**			**V**
	A	B	I	II	A	B	C	
1	1	0	5.5	4.6	1.6	0	0	-4.7
2	1	0	3.2	4.8	1.4	0	0	-2.3
3	1	0	5.2	1.7	2.3	0	0	-1.2
4	1	0	1.0	4.3	3.1	0	0	-5.7
5	1	0	3.0	3.3	5.8	0	0	-5.1
6	1	0	5.9	5.9	4.5	0	0	-1.7
7	1	0	1.7	5.3	3.6	0	0	-5.6
8	1	0	1.6	2.0	4.5	0	0	-9.3
9	1	0	2.5	5.6	3.0	0	0	0.9
10	1	0	1.2	4.0	3.1	0	0	-3.2
11	0	1	4.6	2.9	0	2.7	0	-1.9
12	0	1	1.9	3.5	0	2.0	0	-5.9
13	0	1	5.5	3.3	0	4.7	0	-1.0
14	0	1	2.9	5.3	0	4.6	0	-3.0
15	0	0	4.8	2.3	0	0	2.4	-2.2

Alternative ID	Firm-specific constant		Generic attributes		Firm-specific prices			V
	A	B	I	II	A	B	C	
16	0	0	3.4	2.8	0	0	3.4	-4.8
17	0	0	3.7	5.4	0	0	1.3	0.0
18	0	0	4.9	3.9	0	0	1.8	-1.1
19	0	0	5.3	2.3	0	0	3.2	-3.0
20	0	0	1.1	5.9	0	0	1.8	-1.8
21	0	0	2.1	4.4	0	0	1.6	-1.4

The focal firm alternatives in the experimental conditions where its product line is shorter than 10, is a subset of M and the alternatives kept are those with the highest V for the condition, i.e., when A offers two products, those with the highest V, in the A’s portfolio, are kept and the others are dropped from the choice task.

To add stochasticity to the utility function, that supports consumers’ choices, the random term was drawn from an identical and independent distributed (i.i.d,) Gumbel distribution with scale equal to one. The random term ε_{erntj} varied across alternatives *j*, choice scenarios *t*, individuals *n*, experimental conditions *e* and replications, acknowledging that unobservable variables may have a different effect on consumers’ choice on different occasions.

The number of choice scenarios per consumer was fixed at eight for every experimental condition and replications. This decision has ecological validity, since: consumer make sequential choices in fixed scenarios in many fmcg categories and varying scenarios in different kind of services, the sequential decision-making allows the introduction of variance at the individual level (trough ε) and eight choice tasks is a common number for discrete choice experiments.

c. Rules for choice set formation

As already discussed, choice set formation may result from several psychological processes or decision rules. It may also result in choice sets formation based on different attributes or any combination of attributes. Examples of choice processes resulting in choice set formation are brand loyalty, the inclusion of alternatives with specific attributes that implies characteristics like sustainability, or the exclusion of others that do not meet certain thresholds like price cutoffs. Given that any of this processes, or any combination of them, will potentially be misattributed to taste heterogeneity when not accounted for in choice models, the choice set formation based on brands was chosen to illustrate the proposed effects.

Considering the three firms in the experiment there are seven possible choice sets. In any real sample of observed choices, through real or stated preferences, the true data generation process will be an unknown mixture of the possible choice sets. The mixture is manipulated in the experiment to understand the severity of the studied phenomena caused by different levels of choice set formation or process heterogeneity.

The experimental manipulation was designed in two steps. A variable G_g defines the relative size of consumer groups allocated to: (i) captivity or choice sets formed by products from only one firm; (ii) choice sets formation including products from two firms; and (iii) full information processing, i.e. consideration of all the three firms' products. This last condition is equivalent to the consumers who are making choice among all the alternatives in the choice task, i.e., fully processing the information. A second variable H_{kh} defines the distribution of consumers allocated to the choice sets formed or not formed by the focal brand within each *g*. Given that for $g = 3$ all the firms are included in the choice set, there is no associated *h* and the choice was fully compensatory. Table 4 describes the allocation process and informs the levels adopted for *g* and *h*.

Table 4 - Assignment to the choice set formation rules

Number of firms in the choice set formation rule	Assignment to CSF rules		CSF rules	Assignment to rules including focal firm		
	Variable	Levels		Variable	Levels	
1	g_1	0 0.17 0.34 0.5	*A*	h_1	0 / 0.25 / 0.75 / 1	
			B		$(1 - h_1) / 2$	
			C		$(1 - h_1) / 2$	
2	g_2	0 0.17 0.34 0.5	*AB*	h_2	0 0.25 0.75 1	$(h_2 / 2)$
			AC			$(h_2 / 2)$
			BC		$(1 - h_2)$	
3	g_3	$(1 - g_1 - g_2)$	*ABC*			

One additional criterion to operationalize the process detailed in Table 4 is that for captivity, i.e. condition (i), consumers assigned to the choice set rules excluding the firm *A* were equally distributed between the other two firms. Likewise, consumers assigned to the choice sets rules including the focal firm in condition (ii), i.e. choice set formation with any two firms, were equally distributed across the two considered firms.

The values of G_g and H_h defined the proportion of the sample across the choice set formation rules and individuals were assigned to a specific rule through a random draw following the desired cumulative uniform distribution.

d. Experimental design

The factors and its levels experimentally manipulated are summarized in Table 5 and the full factorial of this experiment would result in 3072 ($3*4^5$) conditions.

Table 5 - Experimental factors and factor levels

Factors	Levels			
	1	2	3	4
Focal brand number of SKUs	2	4	7	10
Focal brand price parameter	-0.5	-1.0	-1.5	
g_1	0	0.17	0.34	0.5
g_2	0	0.17	0.34	0.5
h_1	0	0.25	0.75	1
h_2	0	0.25	0.75	1

In order to maintain the simulation tractable, from the perspective of the Monte Carlo simulation, a main effects experimental design was used and its 25 conditions are presented in Table 6.

Finally, each data set was generated with 1.000 respondents that is a usual sample size in market research studies among heterogeneous populations. And to obtain a distribution of the parameters (tastes and standard deviations) every experimental condition was replicated 250 times, meaning that the power (E. Koehler, Brown, & Haneuse, 2009) was high enough to detect the relationship between any level of any independent variable and the deviation from the estimated to the true parameter.

Table 6 - Experimental design

Condition	Factor levels					
	Number of SKUs	Price parameter	g_1	g_2	h_1	h_2
1	2	-1.0	0.17	0.17	1	1
2	10	-0.5	0.17	0	0.75	0
3	4	-1.0	0	0.34	0	0
4	10	-1.0	0	0	0	0.25
5	10	-1.5	0.34	0.17	0	0.75
6	7	-0.5	0.17	0.34	0	0.75
7	7	-1.0	0	0.17	0.25	0
8	2	-0.5	0	0	0	0
9	4	-0.5	0	0.5	1	0.75
10	10	-0.5	0	0.5	0.25	1
11	2	-0.5	0.34	0	1	0
12	7	-0.5	0.5	0	0	1
13	2	-1.0	0	0	0.75	0.75
14	2	-0.5	0.34	0.34	0.25	0.25
15	2	-1.0	0.17	0.5	0	0.25
16	2	-0.5	0	0.17	0	0
17	4	-1.5	0.17	0	0.25	0
18	4	-1.0	0.34	0	0	1
19	10	-1.0	0.5	0.34	1	0
20	2	-1.0	0.5	0	0.25	0.75
21	7	-1.0	0.34	0.5	0.75	0
22	2	-1.5	0	0.34	0.75	1
23	2	-1.5	0.5	0.5	0	0
24	7	-1.5	0	0	1	0.25
25	4	-0.5	0.5	0.17	0.75	0.25

Table 7 summarizes the settings for the Monte Carlo simulation. In total, there were 6.250 (the number of experimental conditions times the number of replications) data sets, for each market context, each one with 8.000 choice scenarios (the number of observation per data set times the number of choice tasks per observation). The number of alternatives varied across experimental conditions and it was kept constant across replications.

Table 7 - Setting for the data generation process – for each market context

Variable	Notation	Value
Number of experimental conditions	C	25
Number of replications	R	250
Number of observations per data set	N	1.000
Number of choice tasks per observation	T	8
Number of alternatives per choice task	J	13 / 15 / 18 / 21
Number of taste parameters	K	7

4.3.1.2 Pilot

A pilot, in the *fmcg* context, was conducted before the full roll over of the experiment with two main objectives. The first one was to assure that the mixed logit model can recover the true parameters in a condition without choice set formation. The observations engendered to the pilot implies that choices among products from the three firms are homogenous with respect to the choice process and to tastes. Since choices are only disturbed by the random component of the utility function, the algorithm must be able to recover the true parameters, including the betas and the standard deviations. The second objective was to assure that the mixed logit model could converge even in the adverse experimental conditions planned for this study.

Given these objectives, the specific conditions described in Table 8 were created. The first objective was evaluated through the conditions one to four, in which choices are made among the alternatives offered by all the firms. The number of SKUs from firm *A* varied across this conditions, which means that ability to recover true parameters was assessed across the different choice task sizes defined in the main experiment.

The second objective was evaluated through conditions five to eight, designed to create the maximum entropy in the assignment of observations to the choice set formation rules, i.e., all the seven rules were equally present in the data sets. Across these four conditions, the number of SKUs offered by firm *A* was varied covering the full range planned for the main experiment.

Table 8 - Pilot simulation - experimental conditions

Condition	Factor levels					
	Number of SKUs	Price parameter	g_1	g_2	h_1	h_2
1	2	-1.0	0	0	0	0
2	4	-1.0	0	0	0	0
3	7	-1.0	0	0	0	0
4	10	-1.0	0	0	0	0
5	2	-1.0	0.34	0.34	0.33	0.66
6	4	-1.0	0.34	0.34	0.33	0.66
7	7	-1.0	0.34	0.34	0.33	0.66
8	10	-1.0	0.34	0.34	0.33	0.66

Every condition was replicated five times, providing twenty replications to evaluate the recoverability of the true parameters across the four initial conditions and twenty to reassure the model convergence to a solution across the last four conditions.

There are 14 parameters, seven for tastes and seven standard deviations. The results were averaged across the replications for each one of the objectives. The 20 replications offered the parameters' distributions, which were used to calculate the statistic to compare observed results with true parameters. This statistic is t-distributed with 19 degrees of freedom.

For the conditions where the choices were made among the SKUs from all the three firms, the results are presented in Table 9. As expected, there are no significant differences between the true values and the estimates for the taste parameters. However, the estimates of the standard deviations of the taste parameters are all significantly different from zero, which is the expected value given the homogeneity in preferences that support the data generation process.

The results for maximum entropy in choice set formation are detailed in Table 10. As expected, when heterogeneity in the choice process was not accounted for, the econometric models channeled it into the tastes parameters. The bias affected both the location and the standard deviation of the preferences estimates.

Table 9 - Pilot results - no choice set formation

Variable	FSC I	FASC II	gen1	gen2	price I	price II	price III	sd1	sd2	sd3	sd4	sd5	sd6	sd7
TRUE	**0.5**	**-0.5**	**-1.0**	**1.0**	**-1.0**	**-1.0**	**-1.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
Mean	**0.52**	**-0.57**	**-1.00**	**1.01**	**-1.00**	**-1.01**	**-0.99**	**0.03***	**0.25****	**0.04****	**0.05****	**0.02****	**0.06****	**0.02****
sd	**0.21**	**0.40**	**0.04**	**0.03**	**0.04**	**0.12**	**0.17**	**0.05**	**0.28**	**0.04**	**0.06**	**0.02**	**0.08**	**0.02**
***t* stat**	**0.39**	**0.78**	**0.06**	**0.88**	**0.27**	**0.48**	**0.17**	**2.44**	**3.88**	**4.18**	**3.80**	**4.24**	**3.10**	**3.12**
C1_1	0.475	-0.976	-1.083	1.017	-1.105	-0.891	-1.203	0.005	0.182	0.123	0.107	0.001	0.046	0.001
C1_2	0.541	0.103	-0.997	1.041	-0.988	-1.235	-0.992	0.001	0.020	0.004	0.012	0.013	0.006	0.007
C1_3	0.770	-0.403	-0.933	1.034	-0.982	-0.904	-0.827	0.004	0.218	0.001	0.004	0.067	0.003	0.019
C1_4	0.812	0.200	-0.943	1.029	-0.945	-1.107	-0.737	0.029	0.005	0.009	0.004	0.052	0.009	0.006
C1_5	0.552	-0.207	-0.992	1.000	-0.953	-1.094	-0.890	0.097	0.009	0.025	0.146	0.023	0.009	0.009
C2_1	0.218	-0.651	-1.033	1.050	-1.044	-1.023	-1.251	0.008	0.009	0.025	0.014	0.006	0.017	0.044
C2_2	0.768	-0.198	-0.956	0.990	-0.972	-1.077	-0.757	0.007	0.727	0.074	0.005	0.003	0.012	0.011
C2_3	0.594	-0.620	-0.955	1.049	-0.917	-0.927	-0.819	0.002	0.650	0.094	0.006	0.007	0.119	0.003
C2_4	0.321	-0.280	-1.038	0.981	-1.052	-1.223	-1.212	0.001	0.177	0.095	0.114	0.008	0.178	0.008
C2_5	0.206	-1.033	-1.003	0.989	-1.005	-0.893	-1.152	0.011	0.029	0.004	0.050	0.026	0.014	0.059
C3_1	0.783	-0.146	-0.973	0.993	-0.996	-1.094	-0.836	0.002	0.590	0.049	0.067	0.026	0.007	0.007
C3_2	0.570	-0.791	-0.986	1.033	-0.994	-0.920	-0.953	0.050	0.685	0.070	0.174	0.009	0.036	0.092
C3_3	0.400	-0.798	-1.035	1.005	-0.999	-1.122	-1.099	0.197	0.411	0.060	0.122	0.007	0.203	0.019
C3_4	0.477	-1.058	-1.017	0.954	-1.031	-0.870	-1.053	0.003	0.009	0.001	0.002	0.008	0.004	0.004
C3_5	0.141	-1.195	-1.034	0.995	-1.037	-0.974	-1.264	0.146	0.651	0.043	0.012	0.032	0.007	0.008
C4_1	0.425	-0.662	-1.003	1.009	-1.003	-1.135	-1.033	0.006	0.420	0.000	0.002	0.023	0.282	0.004
C4_2	0.689	-0.538	-1.009	0.955	-1.002	-0.924	-0.882	0.000	0.008	0.001	0.048	0.001	0.012	0.004
C4_3	0.376	-1.095	-1.013	1.027	-0.991	-0.879	-1.047	0.006	0.096	0.004	0.060	0.024	0.129	0.009
C4_4	0.817	-0.580	-0.987	0.964	-1.031	-0.890	-0.831	0.021	0.007	0.014	0.013	0.006	0.013	0.011
C4_5	0.444	-0.502	-1.021	1.005	-1.003	-1.081	-1.034	0.008	0.015	0.029	0.003	0.001	0.039	0.002

* p-value <= 0.05 // ** p-value <= 0.01

Table 10 - Pilot results - maximum entropy in choice set formation

Variable	FSC I	FSC II	gen1	gen2	price I	price II	price III	sd1	sd2	sd3	sd4	sd5	sd6	sd7
TRUE	**0.5**	**-0.5**	**-1.0**	**1.0**	**-1.0**	**-1.0**	**-1.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
Mean	**0.20**	**-8.20**	**-1.03**	**1.01**	**-1.05**	**-1.00**	**-1.31**	**3.70**	**11.61**	**0.04**	**0.10**	**0.08**	**0.08**	**0.55**
sd	**0.47**	**1.54**	**0.06**	**0.04**	**0.07**	**0.05**	**0.44**	**0.45**	**1.95**	**0.03**	**0.05**	**0.04**	**0.05**	**0.47**
t stat	**2.76****	**21.80****	**2.31***	**1.26**	**3.08****	**0.24**	**3.04****	**35.83****	**25.90****	**5.97****	**8.22****	**8.88****	**6.84****	**5.11****
C5_1	0.450	-6.396	-1.075	0.915	-1.104	-0.908	-1.167	4.470	10.042	0.003	0.037	0.049	0.070	0.389
C5_2	0.383	-7.041	-0.963	1.015	-0.960	-1.015	-0.899	4.850	10.872	0.033	0.128	0.058	0.047	0.264
C5_3	-0.218	-5.697	-1.085	0.959	-1.146	-0.931	-1.781	3.879	8.375	0.030	0.055	0.056	0.096	1.268
C5_4	-0.564	-6.513	-1.155	1.034	-1.198	-0.993	-1.981	3.949	9.583	0.003	0.184	0.063	0.008	0.875
C5_5	-1.202	-5.912	-1.198	1.078	-1.247	-0.955	-2.812	2.641	6.892	0.031	0.187	0.141	0.071	1.929
C6_1	0.209	-7.403	-1.036	1.007	-1.038	-0.993	-1.178	3.947	11.410	0.072	0.087	0.144	0.000	0.291
C6_2	0.393	-8.052	-1.032	0.971	-1.054	-1.105	-1.154	3.815	11.926	0.015	0.090	0.061	0.158	0.317
C6_3	0.567	-8.385	-0.963	1.048	-0.991	-1.049	-1.106	3.703	12.091	0.027	0.105	0.097	0.081	0.040
C6_4	0.036	-7.959	-1.040	1.002	-1.038	-1.048	-1.321	3.482	11.239	0.050	0.146	0.038	0.061	0.826
C6_5	0.679	-7.207	-0.986	0.981	-0.996	-0.943	-0.939	3.522	10.438	0.106	0.036	0.076	0.105	0.187
C7_1	0.296	-11.004	-1.011	1.034	-1.056	-1.043	-1.178	3.710	12.773	0.017	0.157	0.078	0.015	0.124
C7_2	0.220	-8.986	-1.027	1.026	-1.023	-1.016	-1.243	3.372	12.737	0.044	0.097	0.025	0.125	0.499
C7_3	0.395	-9.954	-1.022	1.027	-1.091	-1.022	-1.254	3.157	14.186	0.088	0.151	0.149	0.065	0.893
C7_4	0.698	-8.822	-0.977	1.025	-0.958	-1.021	-1.001	3.562	12.549	0.031	0.067	0.131	0.151	0.582
C7_5	0.691	-9.692	-0.998	1.007	-1.012	-1.084	-1.022	3.838	15.635	0.058	0.064	0.096	0.187	0.114
C8_1	0.108	-9.042	-1.041	1.047	-1.062	-1.054	-1.424	3.757	11.968	0.016	0.178	0.050	0.129	0.715
C8_2	0.695	-8.900	-0.998	0.985	-1.039	-0.911	-1.009	3.638	12.426	0.038	0.031	0.070	0.075	0.146
C8_3	-0.247	-10.235	-1.034	1.079	-1.034	-0.983	-1.458	3.606	12.750	0.112	0.115	0.163	0.045	0.548
C8_4	0.091	-9.900	-0.997	0.994	-1.010	-1.017	-1.231	3.373	13.105	0.049	0.016	0.075	0.029	0.781
C8_5	0.315	-6.817	-0.993	0.992	-0.994	-0.969	-1.030	3.731	11.222	0.027	0.089	0.050	0.140	0.140

* p-value <= 0.05 // ** p-value <= 0.01

Some interesting patterns could be noticed, after observing that given the assignment of a choice set formation rule to an observation (first choice stage) the choice among the remaining alternatives (second stage) were purely compensatory. If the initial stage process made clear that the bias would be distributed among FSCs and attribute specific attributes, the second stage of the choice process could have provided information for the correct estimation of the generic attributes. However, not only the standard deviation of both generic attributes was biased but the location of gen1 was also different from the true value.

Additionally, all the remaining standard deviations, the two firm-specific constants and two of three firm-specific attributes were biased. Considering that heterogeneity in the choice process was induced by choice sets formed based on firms, this pattern was expected.

It's also worth noticing that severe FSC bias helped to attenuate the corresponding price specific bias as can be observed in the pair FSC II and price II. The FSC set as the base level, i.e. zero, could not be adjusted leading the corresponding alternative-specific attribute (price III) to be severely biased. Finally, when the bias was not strong enough in one of the variables of the pair it was spread in both as can be observed in FSC I and price I.

4.3.1.3 Number of draws in the mixed logit estimation

The number of draws in the estimation of mixed logit models is an important variable that affects the parameters' significance (Hensher & Greene, 2003; Louviere et al., 2000 p. 204). To make sure that the results from the Monte Carlo would not be an artifact of the number of draws selected for estimating the model, the condition 4 (replication 5) of the pilot was used for estimation with a different number of draws.

The evaluation criteria to select the number of draws was the stabilization of the value of the log-likelihood function at the convergence as reported in **Erro! Fonte de referência não encontrada.**. Given the necessity of balancing precision and estimation time, the full experiment was set with 300 draws of the parameters for each replication.

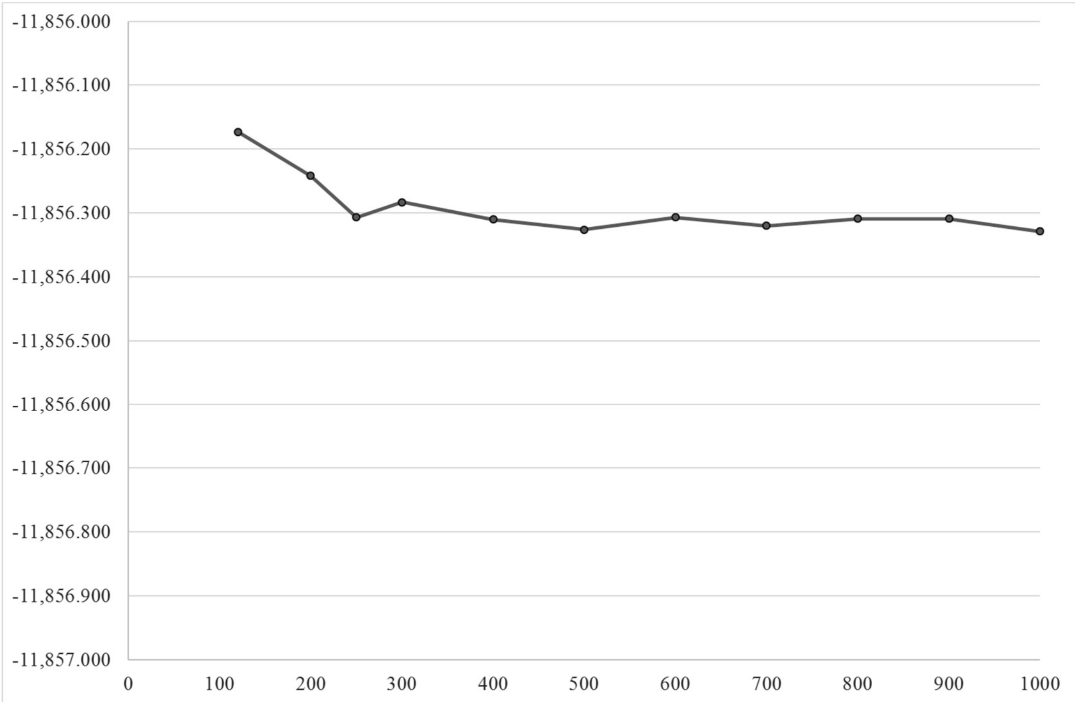

Figure 8- Value of the Log Likelihood function for different number of draws

4.3.2 Demand analysis

The analysis of the demand structure was conducted in two steps that allowed the assessment of the effects of misattributing process heterogeneity to tastes. First, the taste parameters were analyzed in comparison to the true parameters. Second, the policy measures suggested by the true and by the estimate taste parameters were compared, specifically the focal firm's choice probabilities and the focal firm demand attributes' choice elasticities.

4.3.2.1 Taste parameters

Even though model parameters are not the final measures used as inputs in the firm's decision-making, they are used to estimate the policy measures. Hence, the analysis of taste parameters is fundamental to address two issues, one methodological and one substantive.

From the methodological perspective, Monte Carlo simulations are suitable for studying the behavior of statistical methods under experimental conditions. However, it is also subject to uncertainties that should be reported and minimized (E. Koehler et al., 2009). The assessment of the Monte Carlo uncertainties was based on the three measures proposed by Koehler et al., 2009, which are: (a) the percent bias, (b) the coverage probability, and (c) the power to detect an association between the effects and the experimental variables.

From the substantive perspective, the following analysis identified the extent that model estimates were biased, how the bias distributed across parameters and how it responded to the experimental conditions.

a. Percent bias

The percent bias is the relative difference, per experimental condition, between the estimated and the true parameters across replications. It is estimated as:

(4-9)

$$\hat{\phi}_R^b = \frac{1}{R}\sum_{r=1}^{R}\frac{\hat{\beta}_X^r - \beta_X}{\beta_X}\ x\ 100$$

The results portrayed in Figure 9 demonstrate that the biases spread over all the taste parameters, although the percent biases were larger for the firm-specific constants. Specifically, the firm *B* specific constant was, averaged across experimental conditions, ten times larger than the true parameter. The signal of the firm *A* specific constant was reversed, eliminating the qualitative property of higher market power implicit in its positive value.

There was also a difference between the pattern of the biases in the two market contexts studied. For the *fmcg* setting the biases, averaged across experimental condition, were smaller and all the taste parameters (except for the firm *A* specific constant) were overestimated. It is important to notice that the largest attribute bias was observed in the firm *C* specific price. Given that the firm *C* specific constant is omitted for identification, the choice set formation structure needed to be all accounted for in its specific price attribute. And the biases in the other firms' specific prices also seem to be related to their alternative-specific constant such that the largest is the bias in the latter (see firm *B*) the smaller is the bias in the first, and vice-versa. This is the same pattern observed in the pilot study.

Lastly, the magnitude of the bias increased for every attribute in the service context, when variability was added to the generic attributes. Moreover, the direction of the bias was reversed for four out of the five attributes (not considering firm-specific constants), such that the overestimation observed in the *fmcg* context became underestimation in the service context.

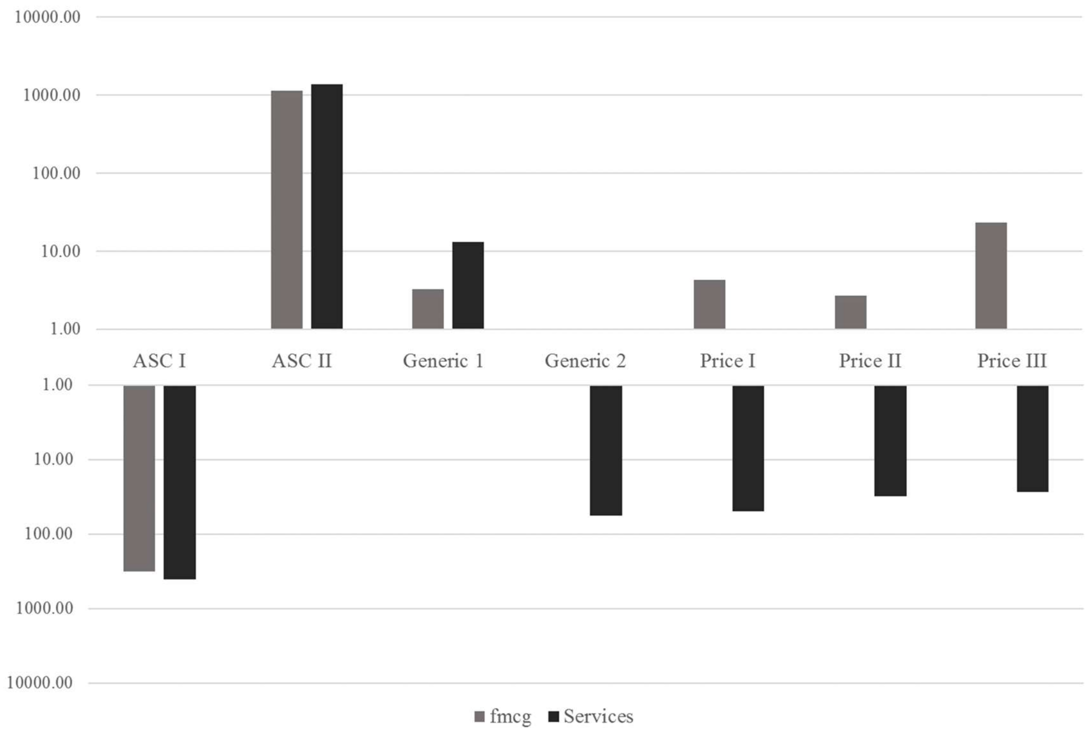

Figure 9 - Monte Carlo experiment - Parameters' biases

The detailed results for percent bias in Table 11 confirmed that the pattern observed in the pilot study spread across every experimental condition. To account for choice set formation, the mixed logit model imposed strongly biased firm-specific constants and firm-specific prices parameters. Since one firm-specific constant (*C*) is omitted for identification, for this firm the bias was entirely in the firm-specific price. For firms *A* and *B* the bias was distributed such that the stronger was the distortion in the specific constant the weaker it was in the specific price, and vice-versa.

Moreover, the estimates for the generic attributes parameters were also biased, which implies that the compensatory stage of the choice process was not enough to provide robust information about these product characteristics when the choice set formation was not accounted for.

Finally, the biases in the standard deviations tended to infinite because the absence of taste heterogeneity implies that these parameters equal to zero in the true data generation process and the estimates from the model are all spurious.

Table 11 – *fmcg* context - Monte Carlo Simulation - Percent bias

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
sku (2)	-694.8	875.5	5.9	-0.4	6.4	0.8	38.8	INF	INF	INF	INF	INF	INF	INF
sku (4)	-32.2	1218.9	1.2	0.9	3.5	3.8	9.8	INF	INF	INF	INF	INF	INF	INF
sku (7)	-272.3	2020.3	1.0	0.7	1.6	3.3	6.8	INF	INF	INF	INF	INF	INF	INF
Sku (10)	87.8	500.4	1.5	0.7	1.6	5.4	16.1	INF	INF	INF	INF	INF	INF	INF
pp (-0.5)	-42.5	1603.9	1.7	0.5	1.4	4.7	13.2	INF	INF	INF	INF	INF	INF	INF
pp (-1.0)	-129.0	694.1	2.9	0.2	4.1	2.7	19.5	INF	INF	INF	INF	INF	INF	INF
pp (-1.5)	-1263.3	894.7	6.5	0.0	8.6	-0.7	44.8	INF	INF	INF	INF	INF	INF	INF
g_1 (0)	-82.2	76.0	3.4	0.2	2.6	4.2	25.8	INF	INF	INF	INF	INF	INF	INF
g_1 (0.17)	-142.7	1363.1	4.9	0.0	7.4	1.0	36.5	INF	INF	INF	INF	INF	INF	INF
g_1 (0.34)	-293.4	1261.4	2.5	0.5	1.3	2.7	14.4	INF	INF	INF	INF	INF	INF	INF
g_1 (0.5)	-1005.9	2714.2	1.4	0.5	5.6	2.0	7.6	INF	INF	INF	INF	INF	INF	INF
g_2 (0)	-42.5	1344.2	1.2	0.3	0.1	3.7	5.6	INF	INF	INF	INF	INF	INF	INF
g_2 (0.17)	-66.7	861.7	4.5	0.3	5.6	1.9	34.7	INF	INF	INF	INF	INF	INF	INF
g_2 (0.34)	-64.1	1191.0	5.7	0.2	6.4	0.8	40.2	INF	INF	INF	INF	INF	INF	INF
g_2 (0.5)	-1390.5	749.5	2.9	0.3	7.4	4.0	24.2	INF	INF	INF	INF	INF	INF	INF
h_1 (0)	-699.6	1534.4	2.4	0.3	3.5	2.3	12.9	INF	INF	INF	INF	INF	INF	INF
h_1 (0.25)	-127.9	1172.4	1.5	0.3	2.9	3.7	13.1	INF	INF	INF	INF	INF	INF	INF
h_1 (0.75)	-214.5	1147.1	5.6	0.2	6.1	1.3	39.5	INF	INF	INF	INF	INF	INF	INF
h1 (1)	135.3	102.4	3.8	0.2	3.6	4.3	31.8	INF	INF	INF	INF	INF	INF	INF

	Taste parameters (*β*s)							Standard deviation of *β*s (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
h2 (0)	-608.6	518.6	1.4	0.3	0.9	4.6	5.8	INF	INF	INF	INF	INF	INF	INF
h2 (0.25)	-118.2	1218.3	2.8	0.0	5.4	2.0	14.7	INF	INF	INF	INF	INF	INF	INF
h2 (0.75)	-102.0	1345.1	2.1	0.8	3.2	3.2	18.0	INF	INF	INF	INF	INF	INF	INF
h2 (1)	-168.9	1890.0	7.9	0.0	9.0	-0.2	66.0	INF	INF	INF	INF	INF	INF	INF

The results for the services context, in Table 12, show that the underestimation of taste parameters happened in most of the experimental conditions (except for firm *B* specific constant and generic attribute 1). In the service context simulation, the biases were larger in magnitude than in the *fmcg* context and this difference was noticeable for every kind of attribute. It means that when the quality level of the consumers' experience varied, the biases caused by not accounting for choice process heterogeneity increased. And such increase spread throughout all taste parameters, not being limited to the ones which caused the variability in the consumers' experience. Moreover, there was also a substantive difference between the two market contexts, since while the *fmcg* contexts induced to overestimation, the services context lead to underestimation of parameters.

Table 12 – Services context - Monte Carlo Simulation - Percent bias

	Taste parameters (*β*s)							Standard deviation of *β*s (across observations)						
	Firm-specific constants		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
sku (2)	-1060.8	1220.9	31.8	-65.0	-124.5	-44.9	3.3	INF	INF	INF	INF	INF	INF	INF
sku (4)	-53.0	1403.0	3.8	-54.4	-35.4	-24.7	-45.1	INF	INF	INF	INF	INF	INF	INF
sku (7)	-193.8	2191.6	5.3	-50.3	-3.6	-25.2	-48.0	INF	INF	INF	INF	INF	INF	INF
Sku (10)	180.6	640.7	-2.1	-50.9	-5.8	-19.2	-47.8	INF	INF	INF	INF	INF	INF	INF
pp (-0.5)	-554.0	1836.9	6.1	-60.4	-173.8	-29.2	-39.2	INF	INF	INF	INF	INF	INF	INF

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constants		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
pp (-1.0)	-12.4	927.7	15.6	-55.4	18.8	-32.1	-28.2	INF	INF	INF	INF	INF	INF	INF
pp (-1.5)	-1055.0	1147.7	27.3	-54.1	16.1	-36.3	0.6	INF	INF	INF	INF	INF	INF	INF
g_1 (0)	-283.6	324.0	16.4	-58.8	-77.0	-32.5	-18.9	INF	INF	INF	INF	INF	INF	INF
g_1 (0.17)	-6.6	1560.4	11.3	-54.8	13.9	-29.1	-23.1	INF	INF	INF	INF	INF	INF	INF
g_1 (0.34)	-731.7	1569.4	19.9	-60.2	-142.4	-36.8	-30.7	INF	INF	INF	INF	INF	INF	INF
g_1 (0.5)	-882.3	2899.1	6.7	-53.2	-11.4	-27.9	-42.6	INF	INF	INF	INF	INF	INF	INF
g_2 (0)	-229.3	1634.1	16.5	-59.3	-68.6	-34.9	-33.9	INF	INF	INF	INF	INF	INF	INF
g_2 (0.17)	-271.9	1085.1	15.4	-59.3	-80.3	-34.4	-18.3	INF	INF	INF	INF	INF	INF	INF
g_2 (0.34)	-205.9	1396.5	14.3	-55.2	-64.5	-31.9	-13.1	INF	INF	INF	INF	INF	INF	INF
g_2 (0.5)	-1251.5	927.2	7.9	-52.6	-11.8	-22.8	-35.1	INF	INF	INF	INF	INF	INF	INF
h_1 (0)	-850.3	1804.0	17.5	-57.4	-71.9	-34.3	-32.9	INF	INF	INF	INF	INF	INF	INF
h_1 (0.25)	-281.6	1440.8	15.3	-56.9	-58.9	-31.5	-29.2	INF	INF	INF	INF	INF	INF	INF
h_1 (0.75)	-132.5	1313.6	8.9	-54.4	-4.7	-28.7	-16.5	INF	INF	INF	INF	INF	INF	INF
h_1 (1)	-73.1	314.6	11.4	-59.5	-86.4	-30.0	-22.7	INF	INF	INF	INF	INF	INF	INF
h_2 (0)	-914.2	800.5	16.3	-58.6	-115.2	-33.0	-34.5	INF	INF	INF	INF	INF	INF	INF
h_2 (0.25)	-318.5	1460.9	15.4	-57.4	-73.7	-33.9	-32.2	INF	INF	INF	INF	INF	INF	INF
h_2 (0.75)	-13.4	1561.3	9.8	-55.9	-3.4	-28.8	-39.2	INF	INF	INF	INF	INF	INF	INF
h_2 (1)	-27.6	2053.7	12.9	-55.2	13.6	-30.2	6.1	INF	INF	INF	INF	INF	INF	INF

In summary, as proposed, the misattribution of process heterogeneity to tastes led not only to the emergence of preference heterogeneity but also affected the location of taste parameters, including the generic attributes. And the magnitude and the direction of the biases were context dependent.

b. Coverage probability

The coverage probability is the frequency that the 95% confidence interval of estimates includes the true parameters. It can be interpreted as the probability of obtaining estimates that are statistically equivalent to the true parameters and it is given by:

(4-10)

$$\widehat{\emptyset}_R^c = \frac{1}{R}\sum_{r=1}^{R} I\left[\hat{\beta}_X^r - 1.96\widehat{se}\left(\hat{\beta}_X^r\right) \leq \beta_X \leq \hat{\beta}_X^r + 1.96\widehat{se}\left(\hat{\beta}_X^r\right)\right]$$

In general, the parameters' empirical coverage probabilities were low in both simulated contexts (Figure 10). Surprisingly, the coverage probabilities were lower for parameters of generics attributes and firm-specific constants than they were for alternative-specific constants, confirming that firm-specific constants were not enough to capture the effect of omitting choice process heterogeneity in the econometric model.

The exceptions were parameters describing the distribution of tastes for the specific price of firm *C* (sd7), which were more similar in magnitude to distributions of the tastes for alternative-specific constants (sd1, sd2). As it can be noticed in the coverage probability formula, the chance of obtaining a parameter estimate that includes the true value in its confidence interval is affected both by the estimate and its variance across replications. Since there is no firm *C* specific constant, the variance of the parameter estimating the distribution of this taste increased and so did the empirical coverage probability.

More important is the observation that adding variability to the generic attributes in the services context reduced the coverage probability in general, especially for the taste parameters. It means that adding variability to the *Xs*, reduced the chance of obtaining valid estimates for the betas when choice process heterogeneity was not accounted for.

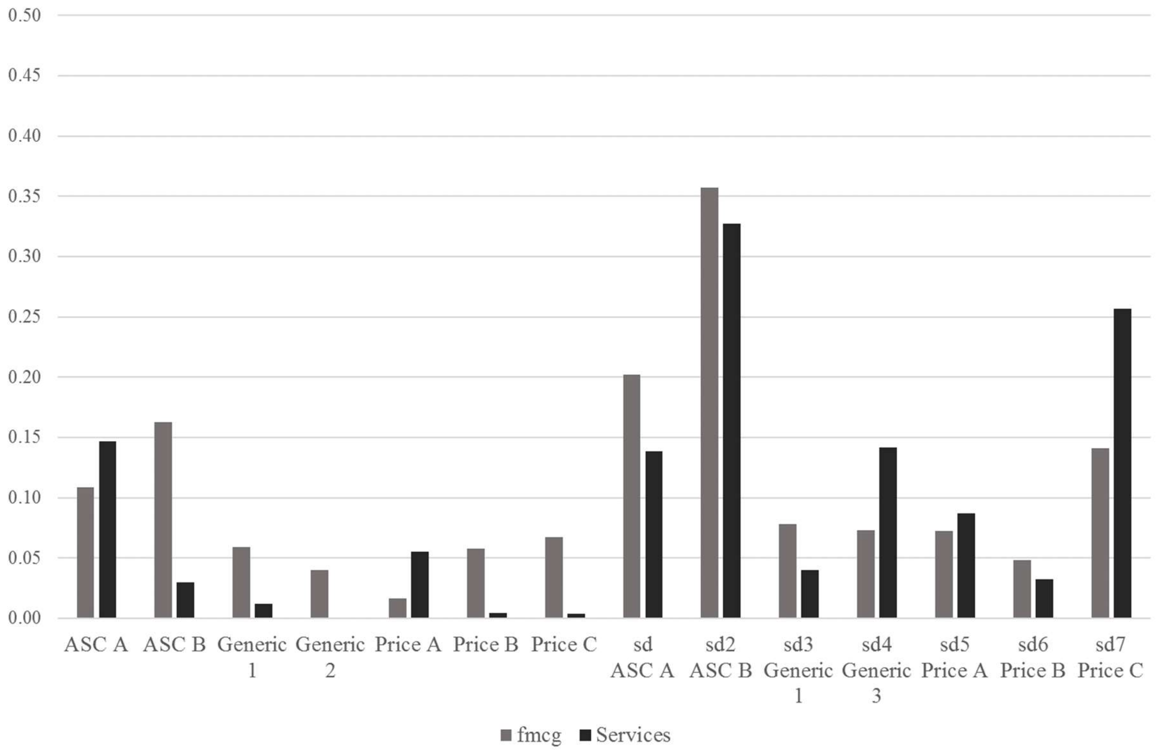

Figure 10 - Monte Carlo experiment - Coverage probability

The results for the coverage probability in Table 13 reveal that, in general, the probabilities of recovering true parameters, from the model that allowed taste heterogeneity without accounting for process heterogeneity, were very low either for the location or the standard deviation of parameters. Surprisingly, this observation was valid even for the generic attributes and it is also valid when the magnitude of the biases were small. It happened because the magnitude of the parameters is closely associated with its standard deviation, meaning tighter confidence intervals for parameters with small biases and wider confidence interval for parameters with larger biases.

Table 13 – *fmcg* context - Monte Carlo simulation - Coverage probability

	Taste parameters (*βs*)							Standard deviation of *βs* (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
sku (2)	0.21	0.11	0.04	0.03	0.01	0.04	0.07	0.29	0.50	0.07	0.07	0.09	0.04	0.20
sku (4)	0.02	0.20	0.04	0.04	0.00	0.06	0.05	0.00	0.31	0.06	0.05	0.03	0.04	0.05
sku (7)	0.10	0.37	0.04	0.04	0.00	0.06	0.04	0.19	0.40	0.11	0.09	0.06	0.07	0.14

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
Sku (10)	0.14	0.17	0.06	0.04	0.00	0.07	0.06	0.36	0.40	0.08	0.11	0.09	0.05	0.18
pp (-0.5)	0.04	0.25	0.03	0.03	0.06	0.06	0.03	0.12	0.46	0.05	0.05	0.08	0.03	0.08
pp (-1.0)	0.07	0.17	0.05	0.04	0.05	0.04	0.07	0.19	0.46	0.08	0.08	0.05	0.04	0.21
pp (-1.5)	0.33	0.08	0.11	0.04	0.12	0.06	0.14	0.41	0.38	0.09	0.08	0.06	0.06	0.18
g_1 (0)	0.05	0.06	0.07	0.03	0.01	0.04	0.08	0.29	0.14	0.12	0.12	0.10	0.07	0.38
g_1 (0.17)	0.03	0.00	0.06	0.04	0.00	0.05	0.08	0.00	0.17	0.03	0.03	0.04	0.02	0.06
g_1 (0.34)	0.03	0.05	0.05	0.05	0.01	0.06	0.05	0.01	0.15	0.06	0.04	0.03	0.03	0.04
g_1 (0.5)	0.14	0.16	0.05	0.04	0.02	0.08	0.05	0.20	0.20	0.05	0.03	0.04	0.04	0.03
g_2 (0)	0.05	0.28	0.04	0.03	0.01	0.06	0.04	0.30	0.48	0.09	0.10	0.09	0.05	0.17
g_2 (0.17)	0.01	0.21	0.04	0.04	0.00	0.06	0.08	0.01	0.49	0.08	0.06	0.06	0.04	0.16
g_2 (0.34)	0.03	0.24	0.11	0.05	0.00	0.06	0.14	0.20	0.54	0.07	0.06	0.06	0.06	0.14
g_2 (0.5)	0.36	0.09	0.06	0.04	0.00	0.06	0.04	0.34	0.29	0.07	0.05	0.05	0.04	0.01
h_1 (0)	0.22	0.23	0.04	0.03	0.00	0.05	0.05	0.20	0.39	0.06	0.07	0.07	0.04	0.17
h_1 (0.25)	0.02	0.20	0.04	0.04	0.02	0.07	0.04	0.09	0.31	0.08	0.06	0.10	0.05	0.06
h_1 (0.75)	0.14	0.10	0.12	0.04	0.02	0.05	0.14	0.39	0.35	0.09	0.09	0.11	0.07	0.24
h_1 (1)	0.10	0.06	0.08	0.06	0.00	0.05	0.08	0.19	0.11	0.12	0.10	0.07	0.05	0.19
h_2 (0)	0.21	0.13	0.04	0.03	0.00	0.04	0.03	0.20	0.43	0.08	0.06	0.06	0.04	0.08
h_2 (0.25)	0.11	0.25	0.06	0.05	0.00	0.05	0.06	0.33	0.34	0.10	0.13	0.20	0.06	0.29
h_2 (0.75)	0.04	0.20	0.04	0.03	0.02	0.07	0.05	0.15	0.33	0.08	0.08	0.07	0.04	0.15
h_2 (1)	0.03	0.15	0.10	0.05	0.00	0.08	0.08	0.19	0.60	0.06	0.05	0.05	0.06	0.04

In summary, the probabilities of obtaining wrong parameters from the model that ignored process heterogeneity and allowed it to be channeled into tastes were higher than 80% for most of the parameters.

The empirical probabilities of recovering valid parameters in the services context simulation were also very low, as it can be observed in Table 14. Adding variability to represent the consumers' experiences in a service context not only reduced the coverage probabilities of the generic attributes, but also the chance of recuperating the true parameters for other types of attributes.

Table 14 – Services context - Monte Carlo simulation - Coverage probability

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
sku (2)	0.07	0.00	0.00	0.00	0.00	0.00	0.01	0.16	0.46	0.00	0.06	0.06	0.03	0.34
sku (4)	0.04	0.05	0.00	0.00	0.01	0.01	0.00	0.00	0.25	0.00	0.07	0.02	0.03	0.04
sku (7)	0.24	0.03	0.04	0.00	0.02	0.00	0.00	0.18	0.40	0.00	0.11	0.04	0.04	0.06
Sku (10)	0.01	0.04	0.00	0.00	0.15	0.01	0.00	0.24	0.37	0.18	0.21	0.18	0.05	0.29
pp (-0.5)	0.10	0.05	0.03	0.00	0.04	0.01	0.00	0.06	0.42	0.09	0.16	0.15	0.02	0.11
pp (-1.0)	0.08	0.02	0.00	0.00	0.07	0.00	0.00	0.15	0.42	0.00	0.12	0.03	0.03	0.42
pp (-1.5)	0.55	0.00	0.00	0.00	0.01	0.00	0.01	0.25	0.36	0.00	0.10	0.05	0.04	0.46
g_1 (0)	0.08	0.01	0.00	0.00	0.00	0.01	0.00	0.14	0.09	0.05	0.22	0.09	0.05	0.57
g_1 (0.17)	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.08	0.16	0.10	0.02	0.36
g_1 (0.34)	0.10	0.00	0.00	0.00	0.13	0.00	0.01	0.00	0.14	0.00	0.06	0.03	0.02	0.04
g_1 (0.5)	0.35	0.11	0.07	0.00	0.17	0.00	0.00	0.18	0.20	0.00	0.08	0.07	0.02	0.02
g_2 (0)	0.09	0.00	0.02	0.00	0.00	0.00	0.00	0.19	0.47	0.02	0.14	0.06	0.03	0.05
g_2 (0.17)	0.22	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.42	0.00	0.11	0.07	0.04	0.40

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
g_2 (0.34)	0.18	0.06	0.00	0.00	0.21	0.00	0.01	0.16	0.51	0.00	0.18	0.07	0.03	0.60
g_2 (0.5)	0.16	0.06	0.00	0.00	0.04	0.01	0.00	0.22	0.23	0.13	0.20	0.15	0.02	0.12
h_1 (0)	0.25	0.00	0.02	0.00	0.01	0.00	0.00	0.12	0.38	0.00	0.10	0.03	0.02	0.07
h_1 (0.25)	0.35	0.05	0.00	0.00	0.00	0.00	0.00	0.02	0.30	0.13	0.22	0.22	0.03	0.18
h_1 (0.75)	0.03	0.00	0.01	0.00	0.05	0.00	0.00	0.32	0.31	0.07	0.18	0.20	0.05	0.65
h_1 (1)	0.04	0.02	0.00	0.00	0.20	0.01	0.00	0.19	0.09	0.00	0.15	0.06	0.05	0.56
h_2 (0)	0.19	0.01	0.00	0.00	0.13	0.01	0.00	0.18	0.38	0.02	0.10	0.04	0.03	0.04
h_2 (0.25)	0.10	0.01	0.00	0.00	0.01	0.00	0.01	0.25	0.30	0.00	0.16	0.10	0.04	0.17
h_2 (0.75)	0.04	0.06	0.00	0.00	0.00	0.01	0.00	0.10	0.28	0.00	0.11	0.06	0.03	0.08
h_2 (1)	0.03	0.10	0.07	0.00	0.00	0.00	0.00	0.05	0.60	0.13	0.27	0.12	0.03	0.27

c. Power

The power is the ability of the design to detect an effect when the effect is true and it is estimated as:

(4-11)

$$\hat{\phi}_R^b = \frac{1}{R}\sum_{r=1}^{R} I\left[\left|\frac{\hat{\beta}_X^r}{\widehat{se}(\hat{\beta}_X^r)}\right| > 1{,}96\right]$$

I[.] is an indicator taking value 1 when the argument is true an 0 otherwise, and $\widehat{se}$ is the standard error of the parameter estimate from the mixed logit model.

The power to detect an association between the experimental design and the results, described in Figure 11, was very high (the number of replications was chosen to produce high power) in both market contexts. Notice also the power relating the design to the location parameters was slightly higher than the power linking the design and the distribution (standard deviations) of tastes.

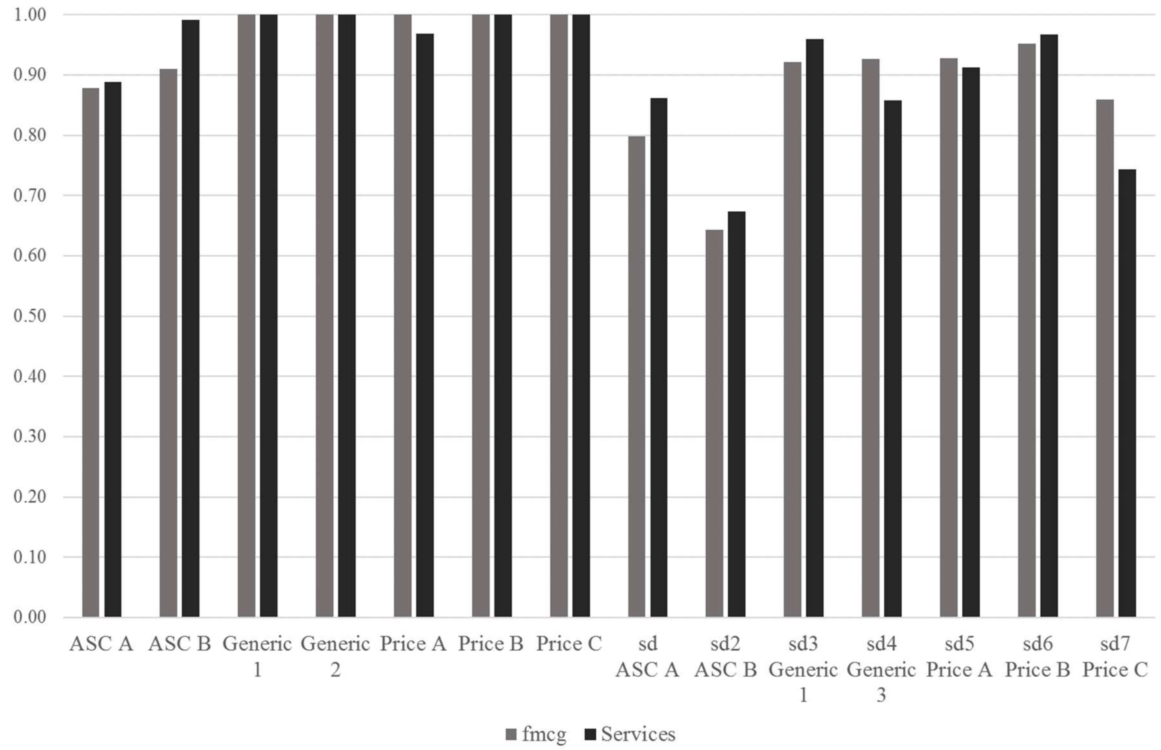

Figure 11 - Monte Carlo experiment - Power to detect an association

The *fmcg* detailed results for the power analysis are presented in Table 15. The power to detect an association between the design and the results approached 1 to every generic attribute or firm-specific price. It was also high for the alternative-specific constant of firm ***B*** (minimum of 0.73) as it was for the focal firm (minimum of 0.61). Given that the most severe adjustments were placed on these parameters, it was expected some reduction in power given the large variance of the estimates across conditions and replications.

The same pattern is observed for the standard deviation of the *βs* with higher power observed among generic attributes and firm-specific prices when compared to alternative-specific constants. It is also noticeable that the power is smaller for the dispersion parameters than for the location ones.

Table 15 – *fmcg* context - Monte Carlo simulation - Power to detect an association

	Taste parameters (βs) Firm-specific constant		Generic attributes		Firm-specific prices			Standard deviation of βs (across observations)						
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
sku (2)	0.74	0.94	1.00	1.00	1.00	1.00	1.00	0.71	0.50	0.93	0.93	0.91	0.96	0.80
sku (4)	0.92	0.90	1.00	1.00	1.00	1.00	1.00	1.00	0.69	0.94	0.95	0.97	0.96	0.95
sku (7)	0.89	0.73	1.00	1.00	1.00	1.00	1.00	0.81	0.60	0.89	0.91	0.94	0.93	0.86
Sku (10)	0.98	0.92	1.00	1.00	1.00	1.00	1.00	0.64	0.60	0.92	0.89	0.91	0.95	0.82
pp (-0.5)	0.98	0.85	1.00	1.00	1.00	1.00	1.00	0.88	0.54	0.95	0.95	0.92	0.97	0.92
pp (-1.0)	0.89	0.92	1.00	1.00	1.00	1.00	1.00	0.81	0.54	0.92	0.92	0.95	0.96	0.79
pp (-1.5)	0.61	0.97	1.00	1.00	1.00	1.00	1.00	0.59	0.62	0.91	0.92	0.94	0.94	0.82
g_1 (0)	0.97	0.97	1.00	1.00	1.00	1.00	1.00	0.71	0.86	0.88	0.88	0.90	0.93	0.62
g_1 (0.17)	0.91	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.83	0.97	0.97	0.96	0.98	0.94
g_1 (0.34)	0.93	0.96	1.00	1.00	1.00	1.00	1.00	0.99	0.85	0.94	0.96	0.97	0.97	0.96
g_1 (0.5)	0.75	0.90	1.00	1.00	1.00	1.00	1.00	0.80	0.80	0.95	0.97	0.96	0.96	0.97
g_2 (0)	0.97	0.82	1.00	1.00	1.00	1.00	1.00	0.70	0.52	0.91	0.90	0.91	0.95	0.83
g_2 (0.17)	0.92	0.90	1.00	1.00	1.00	1.00	1.00	0.99	0.51	0.92	0.94	0.94	0.96	0.84
g_2 (0.34)	0.92	0.87	1.00	1.00	1.00	1.00	1.00	0.80	0.46	0.93	0.94	0.94	0.94	0.86
g_2 (0.5)	0.74	0.96	1.00	1.00	1.00	1.00	1.00	0.66	0.71	0.93	0.95	0.95	0.96	0.99
h_1 (0)	0.70	0.85	1.00	1.00	1.00	1.00	1.00	0.80	0.61	0.94	0.93	0.93	0.96	0.83
h_1 (0.25)	0.92	0.92	1.00	1.00	1.00	1.00	1.00	0.91	0.69	0.92	0.94	0.90	0.95	0.94
h_1 (0.75)	0.98	0.94	1.00	1.00	1.00	1.00	1.00	0.61	0.65	0.91	0.91	0.89	0.93	0.76
h_1 (1)	0.91	0.97	1.00	1.00	1.00	1.00	1.00	0.81	0.89	0.88	0.90	0.93	0.95	0.81

	Taste parameters (*βs*) Firm-specific constant		Generic attributes		Firm-specific prices			Standard deviation of *βs* (across observations)						
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
h2 (0)	0.69	0.93	1.00	1.00	1.00	1.00	1.00	0.80	0.57	0.92	0.94	0.94	0.96	0.92
h2 (0.25)	0.98	0.87	1.00	1.00	1.00	1.00	1.00	0.67	0.66	0.90	0.87	0.80	0.94	0.71
h2 (0.75)	0.97	0.88	1.00	1.00	1.00	1.00	1.00	0.85	0.67	0.92	0.92	0.93	0.96	0.85
h2 (1)	0.94	0.92	1.00	1.00	1.00	1.00	1.00	0.81	0.40	0.94	0.95	0.95	0.94	0.96

The power to detect an association between the experimental design and the results was also high for the simulation in the services context, as it can be observed in Table 16. Considering only taste parameters, it was above 0.9 form every *β* except for the firm-specific constant, when the price parameter was set at -1.5 (0.42). The power was also high for the distribution of the *βs*, although it was smaller than for taste parameters.

Table 16 - Services context - Monte Carlo simulation - Power to detect an association

	Taste parameters (*βs*) Firm-specific constant		Generic attributes		Firm-specific prices			Standard deviation of *βs* (across observations)						
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
sku (2)	0.91	1.00	1.00	1.00	1.00	1.00	1.00	0.84	0.54	1.00	0.94	0.94	0.97	0.66
sku (4)	0.97	0.98	1.00	1.00	0.89	1.00	1.00	1.00	0.75	1.00	0.93	0.98	0.97	0.96
sku (7)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.82	0.60	1.00	0.89	0.96	0.96	0.94
Sku (10)	0.94	0.98	1.00	1.00	1.00	1.00	1.00	0.76	0.63	0.82	0.79	0.82	0.95	0.71
pp (-0.5)	0.98	0.98	1.00	1.00	0.95	1.00	1.00	0.94	0.58	0.91	0.84	0.85	0.98	0.89
pp (-1.0)	0.96	1.00	1.00	1.00	1.00	1.00	1.00	0.85	0.58	1.00	0.88	0.97	0.97	0.58
pp (-1.5)	0.42	1.00	1.00	1.00	1.00	1.00	1.00	0.75	0.64	1.00	0.90	0.95	0.96	0.54
g1 (0)	0.90	1.00	1.00	1.00	0.96	1.00	1.00	0.86	0.91	0.95	0.78	0.91	0.95	0.43
g1 (0.17)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.86	0.92	0.84	0.90	0.98	0.64

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constant		Generic attributes		Firm-specific prices									
	A	B	1	2	A	B	C	1	2	3	4	5	6	7
g_1 (0.34)	0.86	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.86	1.00	0.94	0.97	0.98	0.96
g_1 (0.5)	0.76	0.97	1.00	1.00	0.95	1.00	1.00	0.82	0.80	1.00	0.92	0.93	0.98	0.98
g_2 (0)	0.99	1.00	1.00	1.00	1.00	1.00	1.00	0.81	0.53	0.98	0.86	0.94	0.97	0.95
g_2 (0.17)	0.88	1.00	1.00	1.00	0.91	1.00	1.00	1.00	0.58	1.00	0.89	0.93	0.96	0.60
g_2 (0.34)	0.71	0.99	1.00	1.00	1.00	1.00	1.00	0.84	0.49	1.00	0.82	0.93	0.97	0.40
g_2 (0.5)	0.79	0.98	1.00	1.00	0.94	1.00	1.00	0.78	0.77	0.87	0.80	0.85	0.98	0.88
h1 (0)	0.76	1.00	1.00	1.00	1.00	1.00	1.00	0.88	0.62	1.00	0.90	0.97	0.98	0.93
h1 (0.25)	0.98	0.98	1.00	1.00	1.00	1.00	1.00	0.98	0.70	0.87	0.78	0.78	0.97	0.82
h1 (0.75)	0.92	1.00	1.00	1.00	0.94	1.00	1.00	0.68	0.69	0.93	0.82	0.80	0.95	0.35
h1 (1)	0.96	0.99	1.00	1.00	0.89	1.00	1.00	0.81	0.91	1.00	0.85	0.94	0.95	0.44
h2 (0)	0.87	1.00	1.00	1.00	1.00	1.00	1.00	0.82	0.62	0.98	0.90	0.96	0.97	0.96
h2 (0.25)	0.99	1.00	1.00	1.00	0.91	1.00	1.00	0.75	0.70	1.00	0.84	0.90	0.96	0.83
h2 (0.75)	0.97	0.98	1.00	1.00	0.94	1.00	1.00	0.90	0.72	1.00	0.89	0.94	0.97	0.92
h2 (1)	0.96	0.97	1.00	1.00	1.00	1.00	1.00	0.95	0.40	0.87	0.73	0.88	0.97	0.73

All in all, the power was high and tended to be associated with the percent bias such that the larger the magnitude of the bias the smaller was the power, meaning that as the biases increased so did the variance of parameters' estimates across replications. In general, the specific simulation plan was robust enough to allow the substantive analysis and the understanding design variables' effects on the performance of the mixed logit model.

d. Surface analysis

To answer the substantive question of how the biases in the mixed logit parameters were related to the experimental conditions, a series of generalized linear models were run. The mean differences between estimated and true parameters for the *fmcg* context simulation, per experimental factor and across replications, are reported in Table 17. It can be noticed that the differences were larger for the standard deviations than for the taste parameters. Like in the percent bias analysis, the differences were larger in the firms specific constants and attributes compared to the generic attributes. Moreover, the larger was the difference in firm-specific constants the smaller they were in firm-specific prices. And since the firm *C* specific constant was omitted for identification, the largest difference among the firm-specific prices was the one of firm *C*.

Table 17 – *fmcg* context - Differences between true and biased parameters - means across replication

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constants		Generic attributes		Firm-specific prices									
	A	*B*	1	2	*A*	*B*	*C*	1	2	3	4	5	6	7
sku (2)	-3.47	-4.38	-0.06	0.00	-0.09	-0.01	-0.39	1.33	6.79	0.05	0.07	0.08	0.07	0.54
sku (4)	-0.16	-6.09	-0.01	0.01	-0.02	-0.04	-0.10	1.79	7.82	0.05	0.07	0.11	0.09	0.25
sku (7)	-1.36	-10.10	-0.01	0.01	-0.01	-0.03	-0.07	3.28	16.07	0.04	0.06	0.07	0.08	0.18
sku (10)	0.44	-2.50	-0.01	0.01	-0.02	-0.05	-0.16	3.35	3.58	0.04	0.06	0.05	0.10	0.34
pp (-0.5)	-0.21	-8.02	-0.02	0.01	-0.01	-0.05	-0.13	1.62	11.23	0.05	0.06	0.11	0.10	0.25
pp (-1.0)	-0.64	-3.47	-0.03	0.00	-0.04	-0.03	-0.19	3.52	5.28	0.04	0.06	0.06	0.08	0.33
pp (-1.5)	-6.32	-4.47	-0.06	0.00	-0.13	0.01	-0.45	0.79	8.03	0.04	0.07	0.06	0.07	0.66
g_1 (0)	-0.41	-0.38	-0.03	0.00	-0.04	-0.04	-0.26	0.38	0.19	0.04	0.05	0.05	0.08	0.33
g_1 (0.17)	-0.71	-6.82	-0.05	0.00	-0.07	-0.01	-0.36	1.68	7.11	0.04	0.07	0.08	0.09	0.57
g_1 (0.34)	-1.47	-6.31	-0.02	0.00	-0.02	-0.03	-0.14	3.93	9.98	0.05	0.07	0.12	0.09	0.34

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constants		Generic attributes		Firm-specific prices									
	A	*B*	1	2	*A*	*B*	*C*	1	2	3	4	5	6	7
g_1 (0.5)	-5.03	-13.57	-0.01	0.00	-0.06	-0.02	-0.08	4.68	23.58	0.05	0.07	0.09	0.07	0.25
g_2 (0)	-0.21	-6.72	-0.01	0.00	-0.01	-0.04	-0.06	1.00	11.22	0.05	0.06	0.06	0.08	0.13
g_2 (0.17)	-0.33	-4.31	-0.05	0.00	-0.06	-0.02	-0.35	1.85	5.14	0.05	0.06	0.10	0.08	0.53
g_2 (0.34)	-0.32	-5.96	-0.06	0.00	-0.08	-0.01	-0.40	4.03	6.41	0.05	0.07	0.10	0.08	0.55
g_2 (0.5)	-6.95	-3.75	-0.03	0.00	-0.08	-0.04	-0.24	3.19	7.06	0.04	0.07	0.08	0.10	0.49
h_1 (0)	-3.50	-7.67	-0.02	0.00	-0.05	-0.02	-0.13	1.03	13.35	0.05	0.06	0.07	0.08	0.29
h_1 (0.25)	-0.64	-5.86	-0.01	0.00	-0.02	-0.04	-0.13	1.85	8.89	0.05	0.07	0.10	0.09	0.28
h_1 (0.75)	-1.07	-5.74	-0.06	0.00	-0.07	-0.01	-0.39	3.45	5.26	0.05	0.07	0.08	0.09	0.54
h_1 (1)	0.68	-0.51	-0.04	0.00	-0.04	-0.04	-0.32	3.71	0.21	0.04	0.06	0.08	0.09	0.44
h_2 (0)	-3.04	-2.59	-0.01	0.00	-0.02	-0.05	-0.06	3.23	4.20	0.05	0.06	0.07	0.09	0.14
h_2 (0.25)	-0.59	-6.09	-0.03	0.00	-0.04	-0.02	-0.15	2.07	7.36	0.05	0.06	0.12	0.08	0.30
h_2 (0.75)	-0.51	-6.73	-0.02	0.01	-0.03	-0.03	-0.18	1.67	9.87	0.04	0.06	0.08	0.08	0.39
h_2 (1)	-0.84	-9.45	-0.08	0.00	-0.11	0.00	-0.66	0.86	15.42	0.04	0.07	0.06	0.07	0.85

The service context's absolute differences between true and biased parameters, detailed in Table 18, reproduced the same qualitative pattern observed in the *fmcg* context. However, the magnitude of the differences was larger for every parameter and standard deviation indicating that the distance increased not only for generic attributes parameters that incorporated the variability in its planned levels.

Table 18 – Services context - Differences between true and biased parameters - means across replication

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constants		Generic attributes		Firm-specific prices									
	A	*B*	1	2	*A*	*B*	*C*	1	2	3	4	5	6	7
sku (2)	-5.30	-6.10	-0.32	-0.65	0.48	0.45	-0.03	1.49	7.12	0.26	0.04	0.10	0.10	0.33
sku (4)	-0.26	-7.01	-0.04	-0.54	0.16	0.25	0.45	1.89	7.99	0.22	0.01	0.20	0.14	0.03
sku (7)	-0.97	-10.96	-0.05	-0.50	0.00	0.25	0.48	3.38	16.21	0.17	0.01	0.07	0.12	0.02
sku (10)	0.90	-3.20	0.02	-0.51	0.02	0.19	0.48	3.66	3.84	0.11	0.01	0.05	0.14	0.12
pp (-0.5)	-2.77	-9.18	-0.06	-0.60	0.87	0.29	0.39	1.76	11.33	0.17	0.02	0.14	0.14	0.07
pp (-1.0)	-0.06	-4.64	-0.16	-0.55	-0.19	0.32	0.28	3.69	5.54	0.20	0.02	0.08	0.11	0.15
pp (-1.5)	-5.27	-5.74	-0.27	-0.54	-0.24	0.36	-0.01	1.00	8.57	0.27	0.03	0.09	0.10	0.40
g_1 (0)	-1.42	-1.62	-0.16	-0.59	0.30	0.32	0.19	0.49	0.27	0.20	0.02	0.08	0.10	0.24
g_1 (0.17)	-0.03	-7.80	-0.11	-0.55	-0.18	0.29	0.23	1.96	7.44	0.17	0.03	0.07	0.14	0.27
g_1 (0.34)	-3.66	-7.85	-0.20	-0.60	0.69	0.37	0.31	4.15	10.43	0.22	0.03	0.16	0.13	0.05
g_1 (0.5)	-4.41	-14.50	-0.07	-0.53	0.00	0.28	0.43	4.81	23.89	0.22	0.01	0.13	0.13	0.03
g_2 (0)	-1.15	-8.17	-0.17	-0.59	0.28	0.35	0.34	1.09	11.39	0.20	0.02	0.07	0.12	0.03
g_2 (0.17)	-1.36	-5.43	-0.15	-0.59	0.32	0.34	0.18	2.09	5.36	0.23	0.03	0.15	0.12	0.27
g_2 (0.34)	-1.03	-6.98	-0.14	-0.55	0.25	0.32	0.13	4.15	6.60	0.20	0.03	0.12	0.11	0.38
g_2 (0.5)	-6.26	-4.64	-0.08	-0.53	-0.01	0.23	0.35	3.48	7.54	0.19	0.02	0.10	0.15	0.13
h_1 (0)	-4.25	-9.02	-0.18	-0.57	0.31	0.34	0.33	1.24	13.74	0.21	0.02	0.09	0.12	0.04
h_1 (0.25)	-1.41	-7.20	-0.15	-0.57	0.23	0.32	0.29	1.99	9.18	0.20	0.02	0.12	0.12	0.12
h_1 (0.75)	-0.66	-6.57	-0.09	-0.54	-0.10	0.29	0.17	3.55	5.33	0.21	0.03	0.12	0.12	0.38
h_1 (1)	-0.37	-1.57	-0.11	-0.60	0.37	0.30	0.23	3.87	0.31	0.19	0.03	0.10	0.12	0.25

	Taste parameters (βs)							Standard deviation of βs (across observations)						
	Firm-specific constants		Generic attributes		Firm-specific prices									
	A	*B*	1	2	*A*	*B*	*C*	1	2	3	4	5	6	7
h_2 (0)	-4.57	-4.00	-0.16	-0.59	0.56	0.33	0.34	3.31	4.42	0.21	0.02	0.09	0.12	0.03
h_2 (0.25)	-1.59	-7.30	-0.15	-0.57	0.30	0.34	0.32	2.26	7.56	0.25	0.02	0.18	0.11	0.05
h_2 (0.75)	-0.07	-7.81	-0.10	-0.56	-0.10	0.29	0.39	1.96	10.33	0.19	0.02	0.10	0.14	0.04
h_2 (1)	-0.14	-10.27	-0.13	-0.55	-0.20	0.30	-0.06	1.06	15.55	0.16	0.03	0.06	0.12	0.69

The root mean square error (RMSE) reported is a usual measure to evaluate the reported differences (Andrews, Ansari, & Currim, 2002; Andrews & Currim, 2005) and is defined as:

(4-12)

$$RMSE(\beta_r) = \sqrt{\sum_{x=1}^{X} \frac{\left(\hat{\beta}_x^r - \beta_x^r\right)^2}{X}}$$

It is important to notice that the root mean square error is a measure influenced by observations with larger deviation from the means, given that the deviations are squared before being averaged.

The coding scheme used for the independent variables is the described below and was the same for every model used to conduct the surface analyses, either for parameter estimates or for policy measures. The independent variables were coded as orthogonal polynomial contrasts (OPC) allowing the estimation of linear and quadratic effects between each IV and the DV (Street & Burgess, 2007). The polynomial vectors were normalized, following the procedure described by Pihlens (2008 p. 128-130), such that the support for the OPCs comes from the unit circle permitting the empirical comparison of the magnitude of the attribute effects.

The number of alternatives offered by the focal firm and its price parameter are interpreted as the change in the dependent variable caused by changing the level of the independent variable. At this point, it is important to remember that these variables are descriptors of the market structure. The number of alternatives offered by the focal firm has effect neither on the choice set formation nor on the product evaluation, but given the multinomial logit model as the number of alternatives offered by the focal firm increases its choice probabilities will not decrease. The price parameter has no influence on the choice set formation but it has on the evaluative stage of the decision-making.

Table 19 - Normalized Orthogonal Polynomial Contrasts for Surface Analysis

Factors		Levels			
		1	2	3	4
Focal brand number of alternatives (NALT)	Experimental levels	2	4	7	10
	(1) Linear effect	$-3/\sqrt{20}$	$-1/\sqrt{20}$	$1/\sqrt{20}$	$3/\sqrt{20}$
	(2) Quadratic effect	$1/2$	$-1/2$	$-1/2$	$1/2$
Focal brand price parameter (pp)	Experimental levels	-0.5	-1.0	-1.5	
	(3) Linear effect	$-1/\sqrt{2}$	0	$1/\sqrt{2}$	
	(4) Quadratic effect	$1/\sqrt{6}$	$-2/\sqrt{6}$	$1/\sqrt{6}$	
g_1 and g_2	Experimental levels	0	0.17	0.34	0.5
	(5) Linear effect	$-3/\sqrt{20}$	$-1/\sqrt{20}$	$1/\sqrt{20}$	$3/\sqrt{20}$
	(6) Quadratic effect	$1/2$	$-1/2$	$-1/2$	$1/2$
h_1 and h_2	Experimental levels	0	0.25	0.75	1
	(7) Linear effect	$-3/\sqrt{20}$	$-1/\sqrt{20}$	$1/\sqrt{20}$	$3/\sqrt{20}$
	(8) Quadratic effect	$1/2$	$-1/2$	$-1/2$	$1/2$

The g_1, g_2, h_1 and h_2 are variables that describe the choice process, specifically the structure and extent of choice set formation. The g_1 is the percentage of the sample doing choice set formation with the alternatives of only one firm included in the choice set. The associated parameters are the effects, in the dependent variable, of increasing the choice set formation in favor of any one brand by one percentage point. The h_1 is interpreted as the linear and quadratic effects, in the dependent variable, caused by a one percent point variation in the share of the focal firm's captive consumers, given the choice set formation favoring only one brand. Likewise, g_2 is interpreted as the effects on the *RMSE(β)* of a one percent point variation in the share of choice set formation rules benefitting two firms. And the h_2 represents the effect of including the focal firm in the choice set formation rule, given a choice set formation with any two firms. It is expected that increasing choice set formation should increase bias and increasing choice set formation through rules data exclude the focal firm should lead to even more bias given that the data generation process was based on the focal firm associated with larger firm-specific constants compared to the competitors.

A generalized linear model (GENLIN) was fit to a gamma distribution with a log link, accounting for the extreme values and non-normality of the dependent variable, and the formal model is expressed as:

$$RMSE(\beta) = c + \gamma_1 nalt'_A + \gamma_2 nalt'^2_A + \gamma_3 pp'_A + \gamma_4 pp'^2_A + \gamma_5 g'_1 + \gamma_6 g'^2_1 + \gamma_7 h'_1 + \gamma_8 h'^2_1 + \gamma_9 g'_2 + \gamma_{10} g'^2_2 + \gamma_{11} h'_2 + \gamma_{12} h'^2_2 + scale \tag{4-13}$$

In this equation c is the regression intercept, γ_1 to γ_8 are the regression coefficients describing the relationship between the eight factors coded in Table 19 and the dependent variable and the scale is the parameter the informs the variance of the heteroscedastic stochastic part of the model.

This model outcome is presented for the *fmcg* context in Table 20 and for the services context in Table 21, revealing that all the effects were significant. The presence of non-linear effects imposes some difficulty to comprehend the substantive results only from the model parameters and also to compare the results between contexts. Also, the direct comparison of the magnitude of coefficients is also impaired by the difference in the implicit scales of each model, as can be observed in the last lines of the focal tables. To understand this results, Figure 1Figure 12 portrays the graphics allowing the visualization of the predicted value of *RMSE(β)* as a function of independent variables.

Starting with the *fmcg* context (Table 20), the significant linear effect (Wald $\chi^2[1] = 9.2$, $p > 0.01$) of the number of alternatives offered by the focal firm reveals that the bias increased as the value of independent variable increased. This finding was qualified by a significant quadratic effect (Wald $\chi^2[1] = 118.6$, $p \leq 0.01$) that led the value of the dependent variable to be, surprisingly, reduced after peaking at intermediate levels of the independent variable, as can be observed in Figure 12 (a).

The price parameter of the focal firm also explained part of the bias in the set of parameters estimated by the mixed logit model. The combination of a significant linear effect (Wald $\chi^2[1] = 251.0$, $p \leq 0.01$) with an also significant quadratic effect (Wald $\chi^2[1] = 192.3$, p

≤ 0.01) implies that the bias in the taste parameters increased as the price parameter of the focal firm increases, but an inflexion makes the bias minimum at the intermediate level of the independent variable. This pattern is well described by the U-shaped line in Figure 12 (b). Notice that when the price parameter increased, it also rose the marketing power of the focal firm in the data generation process and the adjustment of the estimated parameters when choice set formation excluded the focal firm must be more severe. Likewise, when the marketing power of the focal firm was reduced through the price parameters in the data generation process, the adjustment of the estimated parameters needed to be less severe when the focal firm was excluded from choice set formation. In other words, when the market power of the focal firm was larger (relatively to the competitors) the parameters were more severely biased to accommodate the choice set formation when it happened against firm *A*. Likewise, when its market power is reduced (compared to the other firms), the biases imposed to the parameters were larger to account for choice set formation when it benefited the focal firm.

Table 20 – *fmcg* context - GLM results for RMSE(β)

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	Df	Sig.
Intercept	0.658	0.010	4691.343	1	0.000
NALT - linear	0.055	0.018	9.166	1	0.002
NALT – quadratic	-0.154	0.014	118.598	1	0.000
pp - linear	0.279	0.018	251.000	1	0.000
pp - quadratic	0.149	0.011	192.261	1	0.000
g_1	2.767	0.012	49884.284	1	0.000
g_1 - quadratic	-0.957	0.013	5547.792	1	0.000
h_1	-1.283	0.023	3026.712	1	0.000
h_1_qdt	-0.628	0.022	833.223	1	0.000
g_2	0.371	0.017	468.115	1	0.000
g_2 - quadratic	-0.311	0.016	390.816	1	0.000
h_2	0.594	0.024	605.862	1	0.000
h_2 - quadratic	0.738	0.029	670.108	1	0.000
Scale	0.234	0.004			

The positive linear effect of increasing the proportion of captivity (g_1) was significant (Wald $\chi^2[1] = 49884.3$, $p \leq 0.01$) revealing that the bias in the *RMSE(β)* was increasing in g_1.

This effect was qualified by a significant quadratic effect (Wald $\chi^2[1] = 5547.8, p \leq 0.01$) that suggests that the bias was more sensitive to increases in g_1 at its intermediate levels. This pattern is clear in Figure 12 (c). The Wald Chi-square statistic indicated that the proportion of captivity was the independent variable with the strongest driver in the dependent variable.

The linear effect of increasing the captivity proportion benefiting the focal firm, given choice set formation with one firm only, was also significant (Wald $\chi^2[1] = 3026.7, p \leq 0.01$) and decreasing in h_1, meaning that it attenuated the effect of g_1. Given that the focal brand has the largest firm-specific constant, the effect was expected because when captivity favors firm *A* it matched the market structure reducing the necessity of adjustment in the parameters of the mixed logit model. There was also a quadratic effect (Wald $\chi^2[1] = 833.2, p \leq 0.01$) indicating that as the choice set formation toward the focal firm increased, the *RMSE(β)* first increased and then decreased (see Figure 12 (d)). This effect is explained by the nested structure of g_1 and h_1, meaning that the first level of h_1 represents two different conditions in the consumers' choice process structure. In the first condition, there was no captivity (g_1 equals zero) implying that h_1 equals zero and no need of adjustment in the true parameters was needed to account for choice set formation. The second condition is the extreme situation in which captivity occurred, but it was always against the focal firm imposing the necessity of a strong bias in the preference parameters to contradict the largest firm-specific constant.

Finally, the *RMSE(β)* was also increasing in g_2, the choice set formation with rules including any two brands (Wald $\chi^2[1] = 468.1$, $p \leq 0.01$). This effect was qualified by a significant quadratic effect (Wald $\chi^2[1] = 390.8, p \leq 0.01$) that driven to a steeper bias response as g_2 reached higher levels. This pattern can be observed in Figure 12 (e), as well as the fact the fact sensitivity of the *RMSE(β)* to g_2 was much weaker than to g_1 (this can also be noticed by the differences in the Wald Chi-square statistic).

Following the pattern observed in the captivity rule, *RMSE(β)* was decreasing in h_2, or the inclusion of firm *A* in the choice set with two firms (Wald $\chi^2[1] = 605.9, p \leq 0.\ 01$. Once again, given choice set formation and a larger firm *A* specific constant, the inclusion of the focal firm in the choice set attenuated the biases in the taste parameters, as it can be noticed in Figure 12 (f).

The results for the services context, presented in Table 21, show that all the effects were significant and substantively similar to the *fmcg* context. One observation is that the sensitivity of the *RMSE(β)* to the manipulation of the focal firm price parameters was higher in the services context, as suggested by the larger Wald Chi-square statistics. However, Figure 12 (b) describes a similar profile of the surface response between the two contexts, with just a steeper sensitivity of the bias to the price parameter when the market power of firm *A* was increases. It means that variations in the *X*s describing generic attributes drove to larger biases in preference parameters when the price elasticity of the focal firm was reduced.

Table 21 – Services context - GLM results for RMSE(β)

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	Df	Sig.
Intercept	0.613	0.006	10575.186	1	0.000
NALT - linear	-0.285	0.012	588.902	1	0.000
NALT - quadratic	0.065	0.010	47.494	1	0.000
pp - linear	0.444	0.009	2244.047	1	0.000
pp - quadratic	0.341	0.009	1363.279	1	0.000
g_1	2.046	0.008	59701.258	1	0.000
g_1 - quadratic	-0.763	0.010	5429.568	1	0.000
h_1	-0.988	0.021	2266.779	1	0.000
h_1 - quadratic	-0.146	0.019	59.741	1	0.000
g_2	0.088	0.009	98.189	1	0.000
g_2 - quadratic	-0.333	0.013	647.453	1	0.000
h_2	-0.145	0.016	77.359	1	0.000
h_2 - quadratic	0.513	0.015	1139.460	1	0.000
Scale	0.132	0.002			

Finally, as can be seen in Figure 12 the average *RMSE(β)* was larger in the services context, compared to the *fmcg* one. It could be expected, since we have already described a larger average bias in the services context, meaning that bias in the taste parameters may increase as the consumer choice context presents more variation across occasions. On the other hand, the surface of the bias was very similar in the two contexts confirming that the effects of the experimental factors implemented in the simulations tend to generalize across contexts.

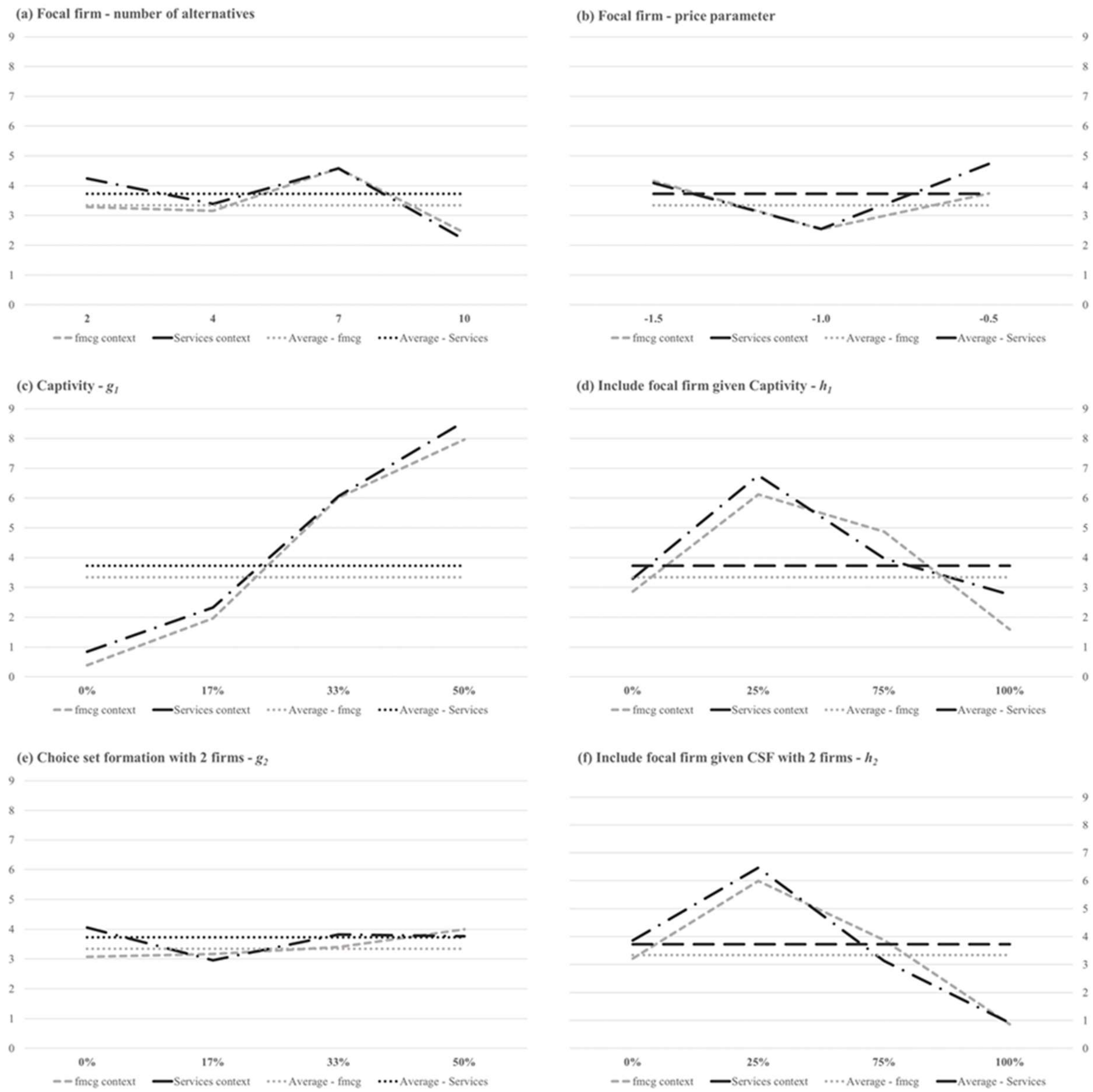

Figure 12 - RMSE(β) as a function of the independent variables

4.3.2.2 Policy measures

After examining the mixed logit parameters, it is important to understand the effects of biased parameters on the final policy measures that will support the firm's decision-making, once it is possible that the biases in the parameters are not large enough to cause large distortions in such policy measures. Moreover, since the largest biases are observed in the firm-specific constants, one could argue that the firm decision-making based on price and other attributes elasticities would result in products with characteristics as demanded by the consumers. In order to understand the potential effects of the taste parameters biases on the firm decision-making, through policy measures, I conducted the analysis of the firm's choice probabilities and the own demand attribute elasticities and their relationship to the experimental design.

For the analysis of the policy measures the variables were aggregated at the firm level and the reporting of the elasticities is based on the results for the focal firm. There are three main motivations for the aggregation. First the choice process heterogeneity was generated at the firm level, conditioning the firms' specific constants coding and implying the appearance of the parameters' biases at this level. Second, the firm is expected to conduct analysis at the product level but to make decisions based on the aggregated performance. Finally, given that the number of alternatives was one of the variables of the experimental design, meaning that not all the alternatives of the focal brand were present in every experimental condition, the firm aggregation allows the analysis of the policy measures across the experimental factors.

a. Choice probabilities

The true choice probabilities for each firm were estimated from the data generation process as:

(4-14)

$$P_f^r = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{S_e}\frac{exp(U_s)}{\sum_{j=1}^{J_{csfn}} exp(U_j)}$$

Where N indexed the 1.000 observations from the dgp, T indexed the eight choice scenarios, S_e indexed the number of alternatives offered by the firm f as defined by the experimental condition, and J_{csf_n} indexed the alternatives included in the choice set of individual n.

Notice that the true probabilities were estimated from U instead of V, meaning that even if the analyst had known the underlying choice set formation, there would still be unobserved variables influencing what the firm understand to be its demand through ε. In other words, these choice probabilities accounted for preference homogeneity and choice set formation but are still subject to other unobserved sources of variation present in any data generation process.

The procedure for estimating the biased choice probabilities took as inputs the taste parameters and their distributions in the population given, as normally distributed, by the mixed logit model illustrated in **Figure 13**. These taste parameters, or random coefficients, can be represented as, $g(\beta|\theta)$ where θ is the vector of the parameters (means and variances) of the tastes distributions.

Still unanswered was the position of each observation of the sample in the population's tastes distributions, $g(.)$, estimated by the model. The choice sequence for each individual or observation in the sample revealed a difference in tastes that drove a distribution, $h(\beta|y,x,\theta)$, for a subsample that made a sequence of choices y, when facing the choice situations described by x. To compute the choice probabilities at the individual level, the location of each individual in the distribution of the population tastes was inferred from the sequence of observed choices following the method described in Train (2009 p.259-265).

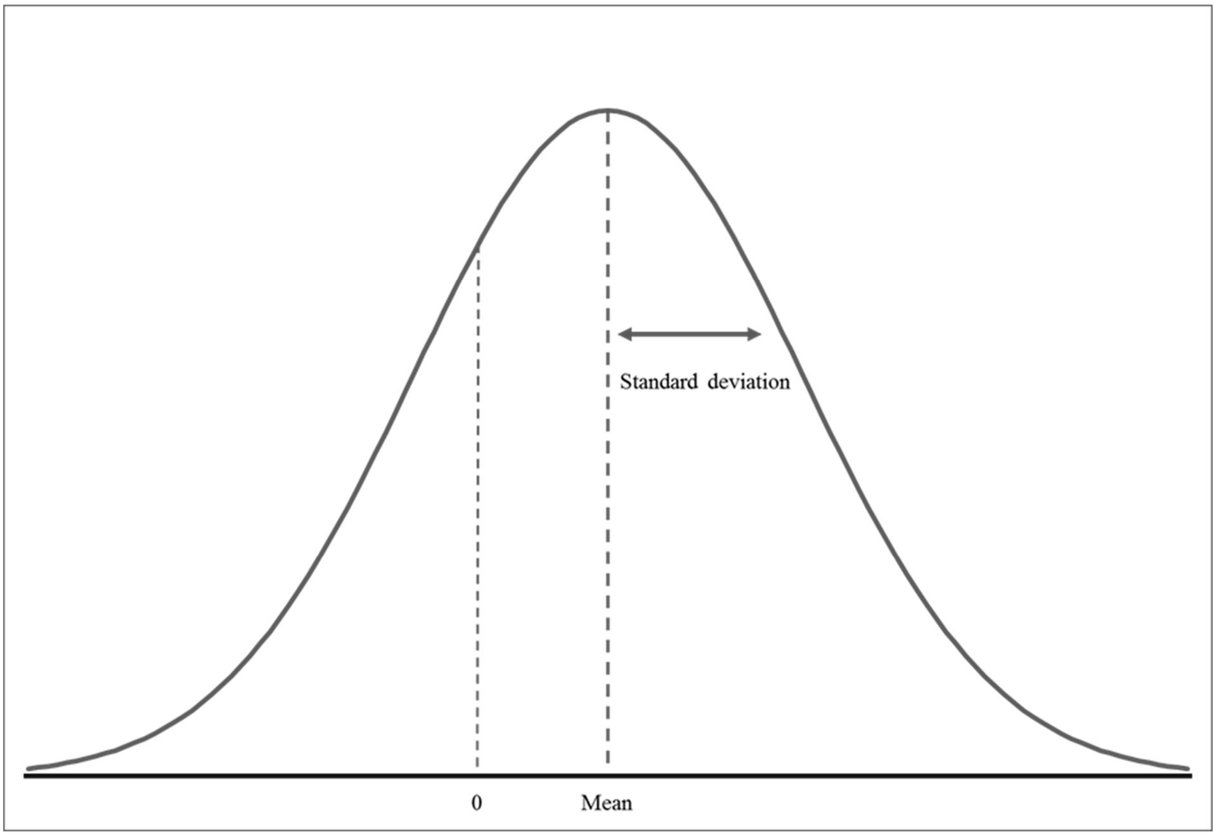

Figure 13 **- Distribution of a mixed logit coefficient in a given sample**
Source: Train, 2009 p. 260

The process involved drawing replications of taste parameters per individual, such that:

(4-15)

$$\theta_n^r = \hat{\beta} + chol(\hat{\Sigma}).n^r$$

Where $\hat{\beta}$ was the vector of estimated means of the random coefficients, $chol(\hat{\Sigma})$ is the Cholesky decomposition of a diagonal matrix with the estimated variances of the random coefficients and $n^r \sim MVN(\hat{\beta}, \hat{\Sigma})$.

The likelihood of the observed sequence of choices per individual is estimated as $ll_n^r = \prod_{t=1}^{T} P_{it}\ (\theta_n^r)$, where P_{it} is the choice probability, given by the true parameters, of the chosen alternative, T is eight and N is 1.000 from the data generation process, and *r* is also 1.000. Finally, the positioning of every individual in $g(.)$ was given by the individuals' *r* draws weighted by the likelihood of the observed choice sequence, or:

(4-16)

$$\hat{\beta}_n = \frac{\sum \theta_n^r\, ll_n^r}{\sum_r ll_n^r}$$

From the individual taste parameters, the choice probabilities for every firm were estimated as:

(4-17)

$$\hat{P}_f^r = \frac{1}{N}\sum_{s=1}^{s_f} \frac{exp(V_s)}{\sum_{j=i}^{J} exp(V_j)}$$

Where $s_1, \ldots\ s_f$, indexed the number of alternatives of the firm, as defined by the experimental condition, and the expression means that the choice probability of any firm was the summation over the choice probabilities of its alternatives. The true and the biased choice probabilities and difference between them, in the *fmcg* context, are presented in Table 22.

Considering only the focal firm, the relative error in the choice probability varied from -7.8% in the condition where the price parameter was -1.0 to +2.5 in the condition in which the price parameter was -0.5. In relative terms, the mistake about the share of choice of the focal firm ranged from a prediction of a demand 36% smaller than it really was (5.3% / 14.6% when the price parameter is -1.5) to a demand that was 5% larger than it really should have been (2.5% / 52.1% when the price parameter was -0.5).

The choice probabilities for firm *B* indicated that from being the smallest player in the true demand representation (2.9% to 14.1%, depending on the experimental condition) it basically disappeared from the marketplace (0.1% to 1.5%) when choice set formation was not accounted for in the choice model. And the winner of this demand misrepresentation was firm *C* that became much more dominant in the biased representation than it was in the true one.

Table 22 – *fmcg* context - Choice probabilities

	(1) True choice probabilities			(2) Biased choice probabilities			(2) – (1)		
	A	B	C	A	B	C	A	B	C
sku (2)	29.7%	10.0%	60.3%	27.7%	0.9%	71.4%	-2.0%	-9.2%	11.1%
sku (4)	36.5%	8.6%	54.9%	33.3%	0.3%	66.5%	-3.3%	-8.3%	11.6%
sku (7)	33.3%	10.2%	56.6%	26.2%	0.3%	73.4%	-7.0%	-9.8%	16.9%
sku (10)	49.9%	6.1%	44.0%	48.3%	0.6%	51.1%	-1.6%	-5.5%	7.1%
pp (-0.5)	52.1%	7.3%	40.6%	54.7%	0.2%	45.2%	2.5%	-7.1%	4.6%
pp (-1.0)	30.1%	8.1%	61.8%	22.3%	0.5%	77.2%	-7.8%	-7.6%	15.4%
pp (-1.5)	14.6%	14.1%	71.3%	9.3%	1.5%	89.2%	-5.3%	-12.6%	17.9%
g_1 (0)	37.0%	3.6%	59.4%	31.3%	1.1%	67.6%	-5.7%	-2.5%	8.3%
g_1 (0.17)	38.0%	8.1%	53.9%	36.6%	0.4%	63.0%	-1.4%	-7.7%	9.1%
g_1 (0.34)	32.6%	12.8%	54.6%	28.2%	0.4%	71.3%	-4.4%	-12.4%	16.7%
g_1 (0.5)	34.5%	16.9%	48.7%	35.7%	0.1%	64.2%	1.3%	-16.8%	15.5%
g_2 (0)	35.0%	9.5%	55.5%	30.3%	0.3%	69.3%	-4.7%	-9.2%	13.8%
g_2 (0.17)	37.3%	7.7%	55.0%	34.2%	0.9%	64.8%	-3.0%	-6.8%	9.8%
g_2 (0.34)	36.1%	8.0%	55.9%	34.4%	1.2%	64.3%	-1.7%	-6.7%	8.4%
g_2 (0.5)	35.6%	10.3%	54.1%	33.8%	0.2%	66.0%	-1.8%	-10.1%	11.9%
h_1 (0)	28.0%	12.9%	59.1%	23.6%	0.5%	75.9%	-4.4%	-12.4%	16.8%
h_1 (0.25)	35.0%	10.1%	54.9%	31.0%	0.2%	68.8%	-4.0%	-9.9%	13.9%
h_1 (0.75)	40.2%	6.2%	53.6%	37.9%	1.1%	61.0%	-2.3%	-5.0%	7.4%

	(1) True choice probabilities			(2) Biased choice probabilities			(2) – (1)		
	A	B	C	A	B	C	A	B	C
h_1 (1)	47.8%	2.9%	49.3%	47.0%	0.7%	52.3%	-0.8%	-2.2%	3.0%
h_2 (0)	37.4%	6.6%	56.1%	33.1%	0.3%	66.6%	-4.3%	-6.2%	10.5%
h_2 (0.25)	32.1%	8.6%	59.3%	26.6%	0.4%	73.0%	-5.5%	-8.2%	13.7%
h_2 (0.75)	34.4%	11.5%	54.1%	32.7%	0.6%	66.6%	-1.7%	-10.8%	12.5%
h_2 (1)	37.7%	11.8%	50.5%	37.7%	1.3%	61.0%	0.0%	-10.5%	10.5%

The results for the services context, in Table 23, exposes the same pattern observed in the *fmcg* context, including the fact that firms *A* and *B* tended to have reduced choice probabilities in the biased demand representation while the choice probabilities of firm *C* tended to increase. But the magnitude of differences in the firms' choice probabilities were slightly larger in the services context, as it could be expected since the biases in the taste parameters were also larger in this context.

Table 23 – Services context - Choice probabilities

	(1) True choice probabilities			(2) Biased choice probabilities			(2) – (1)		
	A	B	C	A	B	C	A	B	C
sku (2)	29.7%	10.0%	60.3%	20.2%	0.6%	79.2%	-9.5%	-9.5%	19.0%
sku (4)	36.5%	8.6%	54.9%	30.7%	0.2%	69.1%	-5.8%	-8.4%	14.3%
sku (7)	33.2%	10.2%	56.6%	26.1%	0.2%	73.7%	-7.1%	-9.9%	17.1%
sku (10)	49.9%	6.1%	44.0%	48.1%	0.3%	51.7%	-1.9%	-5.9%	7.7%
pp (-0.5)	52.1%	7.3%	40.6%	47.0%	0.1%	52.9%	-5.2%	-7.2%	12.3%
pp (-1.0)	30.1%	8.1%	61.8%	21.7%	0.3%	78.0%	-8.4%	-7.8%	16.2%
pp (-1.5)	14.6%	14.1%	71.3%	7.9%	0.9%	91.2%	-6.7%	-13.2%	19.9%
g_1 (0)	37.0%	3.6%	59.4%	27.4%	0.8%	71.8%	-9.6%	-2.8%	12.4%

	(1) True choice probabilities			(2) Biased choice probabilities			(2) – (1)		
	A	**B**	**C**	**A**	**B**	**C**	**A**	**B**	**C**
g_1 (0.17)	37.9%	8.1%	53.9%	37.0%	0.1%	62.9%	-1.0%	-8.0%	9.0%
g_1 (0.34)	32.6%	12.8%	54.7%	17.7%	0.1%	82.2%	-14.9%	-12.7%	27.6%
g_1 (0.5)	34.5%	16.8%	48.7%	35.7%	0.1%	64.2%	1.3%	-16.8%	15.5%
g_2 (0)	35.0%	9.5%	55.5%	26.8%	0.2%	72.9%	-8.2%	-9.2%	17.4%
g_2 (0.17)	37.3%	7.7%	55.0%	28.9%	0.3%	70.8%	-8.4%	-7.4%	15.8%
g_2 (0.34)	36.1%	8.0%	55.9%	30.1%	0.9%	68.9%	-6.0%	-7.1%	13.1%
g_2 (0.5)	35.5%	10.3%	54.1%	32.5%	0.1%	67.4%	-3.0%	-10.2%	13.3%
h_1 (0)	28.0%	12.9%	59.1%	19.8%	0.2%	79.9%	-8.2%	-12.6%	20.8%
h_1 (0.25)	35.0%	10.2%	54.9%	26.6%	0.1%	73.2%	-8.3%	-10.0%	18.3%
h_1 (0.75)	40.2%	6.2%	53.7%	37.8%	0.9%	61.3%	-2.4%	-5.3%	7.7%
h_1 (1)	47.8%	2.9%	49.3%	41.1%	0.4%	58.6%	-6.7%	-2.5%	9.2%
h_2 (0)	37.4%	6.5%	56.1%	28.1%	0.2%	71.6%	-9.2%	-6.3%	15.5%
h_2 (0.25)	32.1%	8.6%	59.3%	20.7%	0.2%	79.1%	-11.4%	-8.3%	19.8%
h_2 (0.75)	34.4%	11.5%	54.1%	30.9%	0.2%	68.9%	-3.5%	-11.3%	14.8%
h_2 (1)	37.7%	11.8%	50.5%	37.3%	0.9%	61.8%	-0.4%	-10.9%	11.3%

One important and unexpected detail, that has not been considered so far, is that the biases in the choice probabilities may be related to the parametrization of the firm-specific constants. Notice that, as a function of the variables controlling choice set formation, biased choice probabilities always penalize firms *A* and *B* in both market contexts simulated. There was one exception in the *fmcg* context (difference in probabilities for the focal firm is zero when h_2 is one) and one in the services context (difference for the focal firm is +1.3% when g_1=0.5).

Examining the parametrization, since one FSC must be omitted for identification, firms *A* and *B* were described by the remaining dummy variables and the true taste parameters 0.5 and -0.5 were assigned to each of them. These parameters mean that firm *A* has some degree of market power, since that if both offer identical products the larger FSC would raise the choice probability of the alternative offered by the focal firm. The inverse argument applies to firm *B*, and the sources of such market power (or weakness) would be explained, in any empirical context, by unobserved variables whose effects are captured by these dummies. Notice that to omit the firm *C* specific constant means to assign a value of zero to its taste parameter, implying that its market power is between *A* and *B* and also implying that the firms are ordered with respect to the firm-specific constants (or to market power). Given that the true data generation process imposes independence between choice set formation and parametrization there is no reason to expect firm *C* to be a systematic winner in the biased demand representation and firms *A* (largest FSC parameter) and *B* (smallest FSC parameter) to be the systematic loser.

We would need to rerun all the models to confirm the relationship between parametrization and choice probabilities bias, what it was not a feasible effort in the timeframe of this thesis. However, we can speculate that, although the amount of bias is not driven by the parametrization, its distribution across the alternatives could be affected. Even not being relevant to our objective at this point, which is to demonstrate that not considering choice process heterogeneity causes the analyst to misattribute it to tastes, the possibility speculated here is disturbing. Once the parametrization in any empirical study is an arbitrary decision made by the analyst, that should be neutral to the substantive findings, the presence of biases introduced by not modeling choice set formation may result in dependence between parametrization and choice probabilities. This hypothesis should be investigated in future studies.

After this digression, taking expressions (4-14) and (4-17) in consideration, the root mean square error of the choice probabilities was defined as:

(4-18)

$$RMSE(P_r) = \sqrt{\sum_{f=1}^{3} \frac{\left(\hat{P}_f^r - P_f^r\right)^2}{3}}$$

To study the effects of the experimental design on the firms' probabilities bias a generalized linear model (GENLIN) was fit to a normal distribution on the square root of the *RMSE*(P_r) with identity link, so the model was formalized as:

$$RMSE(P_r) = c + \gamma_1 nalt'_A + \gamma_2 nalt'^2_A + \gamma_3 pp'_A + \gamma_4 pp'^2_A + \gamma_5 g'_1 + \gamma_6 g'^2_1 + \gamma_7 h'_1 + \gamma_8 h'^2_1 + \gamma_9 g'_2 + \gamma_{10} g'^2_2 + \gamma_{11} h'_2 + \gamma_{12} h'^2_2 + scale \tag{4-19}$$

The results for the *fmcg* context are presented in Table 24 and for the services context in Table 25. Also, Figure 14 depicts the relationship between the *RMSE(P)* and the experimental factors to help the interpretation of the model coefficients and the comparison between contexts

The linear effect of the number of alternatives offered by the focal firm was significant (Wald $\chi^2[1] = 29.4$, $p \leq 0.\ 01$) and it indicates that the value of the dependent variable tended to increase in the number of alternatives of the focal firm. Moreover, there was a significant quadratic effect (Wald $\chi^2[1] = 96.7$, $p \leq 0.\ 01$) implying that the *RMSE*(P_r) peaked at the intermediate levels of the number of alternatives offered by the focal firm, as it can be observed in Figure 14 (a).

The linear effect of the price parameter was significant (Wald $\chi^2[1] = 1005.0$, $p \leq 0.\ 01$), revealing that the higher the focal firm's market power the smaller was the bias in the distribution of the choice probabilities across firms. This effect was qualified by a significant quadratic effect (Wald $\chi^2[1] = 958.7$, $p \leq 0.\ 01$) which reveals that, starting from the intermediate level in which the price parameter is equal (-1.0) for all the firms, the decrease in the value of the dependent variable was steeper when the market power of the focal firm is enlarged (price parameter of *A* equal to -0.5) than the decrease when it was reduced (price parameter of *A* equal to -1.5), as depicted in Figure 14 (b). The explanation for this quadratic effect may rest in the higher focal firm-specific constant and in the fact that if the price parameter moved in conformance with the firm-specific constant, the deviation of the set of choice probabilities was smaller than if the focal firm-specific constant and it price parameter contradicted each other.

The positive linear effect of g_1 (Wald $\chi^2[1] = 7926.5.6$, $p \leq 0.01$) confirms that the stronger was the choice set formation involving one brand only, the stronger was the bias on the set of probabilities. The quadratic effect of g_1 was significant (Wald $\chi^2[1] = 1444.4, p \leq 0.1$) revealing that the increase in bias decelerated when g_1 approached its highest levels (see Figure 14 (c)). Given the level of choice set formation including one firm only, the significant negative linear effect of h_1 (Wald $\chi^2[1] = 5824.7$, $p \leq 0.01$) implies that the more the focal firm was benefited the more the bias was attenuated. This was expected since the larger firm-specific constant associated with the focal firm means that choice set formation against it will require stronger adjustments in the model parameters to fit the data. This linear effect was qualified by a significant quadratic effect (Wald $\chi^2[1] = 478.1, p \leq 0.01$), detected in Figure 14 (d), that can be explained by the nested structure between g_1 and h_1 as described in the analysis of the *RMSE(β)*.

Table 24 – *fmcg* context - GENLIN results for RMSE(P_r)

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	Df	Sig.
Intercept	0.261	0.001	60370.486	1	0.000
NALT - linear	0.008	0.001	29.375	1	0.000
NALT - quadratic	-0.014	0.001	96.712	1	0.000
pp - linear	-0.038	0.001	1004.990	1	0.000
pp - quadratic	-0.036	0.001	958.694	1	0.000
g_1	0.126	0.001	7926.456	1	0.000
g_1 - quadratic	-0.061	0.002	1444.440	1	0.000
h_1	-0.181	0.002	5824.678	1	0.000
h_1 - quadratic	-0.044	0.002	478.080	1	0.000
h_2	-0.006	0.002	15.819	1	0.000
h_2 - quadratic	0.012	0.001	60.955	1	0.000
h_2	-0.027	0.002	199.068	1	0.000
h_2 - quadratic	0.032	0.002	249.272	1	0.000
Scale	0.003	0.000			

Finally, there was a significant linear effect of g_2 (Wald $\chi^2[1] = 15.8, p \leq 0.01$) qualified by a significant quadratic effect (Wald $\chi^2[1] = 61.0, p \leq 0.01$) meaning that the extreme values of choice set formation including two firms tended to increase the biases in the firm's choice probabilities (see Figure 14 (e)). The significant negative linear effect of h_2 (Wald $\chi^2[1] = 199.1$, $p \leq 0.\ 01$) means that given the choice set formation including any two firms, the inclusion of the focal firm helped to attenuate the bias in the set of probabilities. Moreover, the significant positive quadratic effect (Wald $\chi^2[1] = 249.3$, $p \leq 0.01$) implies that the attenuation effect became more important at higher levels of h_2, what is partially explained by the nested structured between g_2 and h_2.

Table 25 describes the results of the services context surface analysis of *RMSE*(P_r). The patterns were similar once again, with a minor difference in the relationships between the two first factors and the dependent variable, as can be noticed in Figure 14 (a) and (b). The effect of the number of alternatives offered by firm *A* seems to conform to a linear relationship attenuated at the intermediate levels of the independent variable. And the effect of price parameter differs only in the quadratic effect that now attenuated the reduction in the *RMSE*(P_r) when the price parameter increased. These small difference are induced by the introduction of variation in the *Xs* of the generic attributes and did not change the substantive results.

Table 25– Services context - GENLIN results for *RMSE*(Pr)

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	df	Sig.
Intercept	0.252	0.001	42627.497	1	0.000
NALT - linear	-0.028	0.002	217.206	1	0.000
NALT - quadratic	0.040	0.002	493.257	1	0.000
pp - linear	0.001	0.002	0.506	1	0.477
pp - quadratic	0.014	0.001	91.506	1	0.000
g_1	0.090	0.002	3498.926	1	0.000
g_1 - quadratic	-0.107	0.002	3276.652	1	0.000
h_1	-0.179	0.003	3914.482	1	0.000
h_1 - quadratic	-0.017	0.003	36.831	1	0.000
h_2	-0.002	0.002	1.435	1	0.231
h_2 - quadratic	-0.038	0.002	390.469	1	0.000
h_2	-0.113	0.002	2546.375	1	0.000
h_2 - quadratic	0.043	0.002	393.982	1	0.000
Scale	0.005	0.000			

In conclusion, the overall results suggest that the *RMSE*(P_r) bias's surface were quite similar across the two studied contexts, as it was in the previous dimensions analyzed. Additionally, the magnitude of the bias was larger in the services context than in the *fmcg* one as indicated by the horizontal lines in the Figure 14 graphs. This difference was expected since it follows the larger Monte Carlo experiment bias and the larger *RMSE(β)*. In order words, when the consumers' experience is affected by a larger variability imposed by the market context, the biases, introduced by ignoring choice process heterogeneity and allowing it to be confounded with preferences, were higher both in the preference parameters and in the choice probabilities.

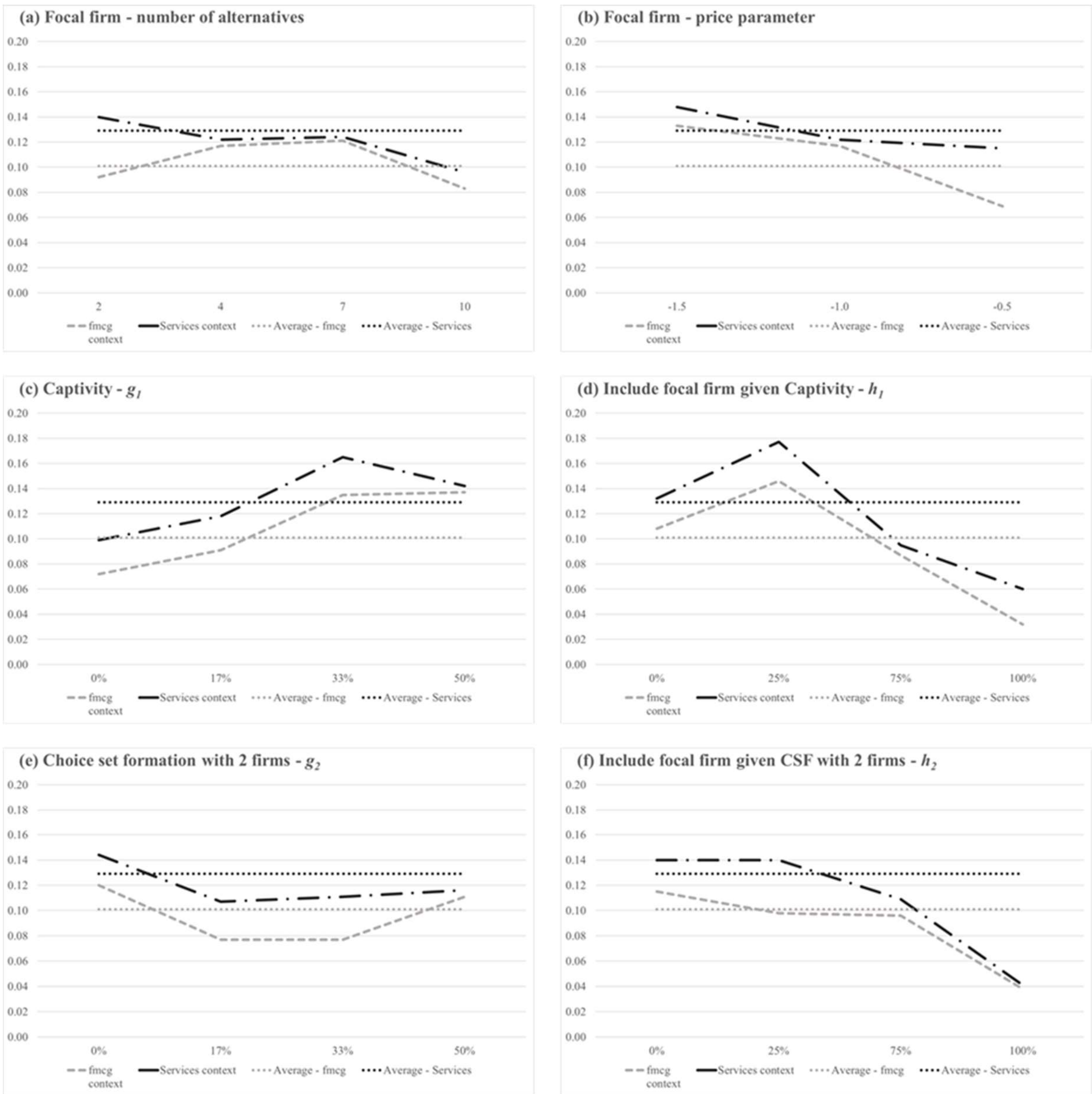

Figure 14 - RMSE(P) as a function of the independent variables

It is important to notice that there were some unexpected nonlinearities, in both contexts, like the ones drove by the number of alternatives of the focal firm, by g_1 and by g_2. A detailed examination of Table 22 and Table 23 suggests that these patterns may have been caused by the deviation in the focal firm's choice probabilities. A possible explanation is that the context manipulation through the number of alternatives offered by firm *A* and through its price parameter may be causing some interaction between the factors and the ones that describe the consumers' choice process. The interaction could have been introduced in the design by the nested structure of g_1 and h_1 and of g_2 and h_2.

Now I turn the attention to the analysis of the bias in the focal firm's choice probabilities to understand the first piece of information that a focal firm's decision-maker would use. Table 26 and Table 27 detail the results of generalized linear models (GENLIN) fitted to the absolute difference between true and biased focal firm's choice probabilities, using a normal distribution and an identity link. The results can be interpreted as the change in biased choice probabilities minus the true probabilities as a function of varying the levels of the independent variables, or more precisely:

(4-20)

$$\hat{P} - \tilde{P} = c + \gamma_1 nalt'_A + \gamma_2 nalt'^2_A + \gamma_3 pp'_A + \gamma_4 pp'^2_A + \gamma_5 g'_1 + \gamma_6 g'^2_1 + \gamma_7 h'_1 + \gamma_8 h'^2_1 + \gamma_9 g'_2 + \gamma_{10} g'^2_2 + \gamma_{11} h'_2 + \gamma_{12} h'^2_2 + scale$$

In this equation $\hat{P}$ is true focal firm's choice probability, i.e. considering choice process heterogeneity and taste homogeneity and $\tilde{P}$ is the biased focal firm's choice probability, i.e. assuming choice process homogeneity and allowing for taste heterogeneity. The independent variables are the same described for the previous regressions and the bias surface analysis results are presented in Table 26 for the *fmcg* context and in Table 27 for the services context. Once again a series of graphs, in Figure 15, are used to support the interpretation of the model parameters and the comparison between contexts.

The combination of a significant linear effect of focal firm's portfolio size (Wald $\chi^2[1] = 71.3$, $p \leq 0.01$) with a significant quadratic effect (Wald $\chi^2[1] = 114.4$, $p \leq 0.01$) reveals that the largest differences in the firm's *A* choice probabilities were observed at the intermediate levels of the independent variable (see Figure 15 (a)).

The combination of a significant linear effect of the price parameter (Wald $\chi^2[1] = 3331.7$, $p \leq 0.01$) with a significant quadratic effect (Wald $\chi^2[1] = 2615.0$, $p \leq 0.01$) reveals that the focal firm's absolute difference in choice probabilities peaked when the firm's A price parameter was equal to the competitors. This effect can be observed in Figure 15 (b) through the U-shaped response curve (notice that the negative deviations imply that the largest difference is observed in the low side of the graph). A possible explanation for this pattern is that when the focal firm price parameter was different from the competitors price parameter, a model that did not account for choice set formation received a strong signal form the firm's *A* specific price parameters and needed to overcome it in of two ways: if the firm's *A* price parameters was smaller (and price elasticity is larger) than the competitors' price parameter, a strong adjustment was necessary to accommodate the choice set formation that benefited firm *A*. Likewise, when the focal firm price parameter was larger (and price elasticity is smaller), a strong adjustment was now necessary to accommodate the choice set formation that excluded firm *A*.

Table 26 – *fmcg context* - GENLIN results for the focal firm choice probabilities (biased - true)

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	df	Sig.
Intercept	-0.008	0.001	97.348	1	0.000
NALT - linear	-0.009	0.001	71.260	1	0.000
NALT - quadratic	0.013	0.001	114.428	1	0.000
pp - linear	0.053	0.001	3331.664	1	0.000
pp - quadratic	0.045	0.001	2615.007	1	0.000
g_1	0.052	0.001	1692.127	1	0.000
g_1 - quadratic	0.020	0.001	230.539	1	0.000
h_1	0.031	0.002	421.800	1	0.000
h_1 - quadratic	0.023	0.002	192.752	1	0.000
g_2	0.001	0.002	0.474	1	0.491
g_2 - quadratic	0.015	0.001	196.693	1	0.000
h_2	0.063	0.001	2636.107	1	0.000
h_2 - quadratic	0.001	0.002	0.322	1	0.570
Scale	0.002	0.000			

The significant linear effects of g_1 (Wald $\chi^2[1] = 1692.1$, $p \leq 0.01$) implies that the absolute difference moved in the same direction of the adoption of this choice rule and g_1 - *quadratic* (Wald $\chi^2[1] = 230.5$, $p \leq 0.01$) imposed a less steep curve at the intermediate levels of the adoption of captivity, meaning that the biased demand of the focal firm was smaller than the true one at lower levels of g_1 and it was larger at the higher levels of the independent variable

(Figure 15 (c)). The same can be observed from the significant effects of h_1 (Wald $\chi^2[1] = 421.8$, $p \leq 0.01$) and h_1 - *quadratic* (Wald $\chi^2[1] = 192.8$, $p \leq 0.01$) since that when h_1 was at zero, choice set formation may be absent or may be happening but excluding firm *A*. Then absolute bias tended to zero when *h1* moved to its intermediate level and grew positive when captivity predominantly included the focal firm (Figure 15 (d)).

Finally, the results reveal that the response of the dependent variable to the variables describing choice set formation with two firms had similar effects to the ones modeling choice set formation with one firm. In other words, when the incidence of choice set formation with two firms increased the absolute difference in firm's *A* choice probabilities was reduced (Figure 15 (e)) and if the inclusion of the focal firm in the CSF rule got more frequent the absolute differences became positive (see Figure 15 (f)).

Table 27 – Services context - GENLIN results for the focal firm choice probabilities (biased - true)

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	df	Sig.
Intercept	-0.012	0.001	82.385	1	0.000
NALT - linear	0.043	0.002	393.617	1	0.000
NALT - quadratic	-0.021	0.002	91.499	1	0.000
pp - linear	0.016	0.002	81.408	1	0.000
pp - quadratic	-0.004	0.002	5.719	1	0.017
g_1	0.077	0.002	1996.227	1	0.000
g_1 - quadratic	0.058	0.002	727.671	1	0.000
h_1	0.009	0.003	11.525	1	0.001
h_1 - quadratic	-0.019	0.003	31.785	1	0.000
g_2	-0.004	0.002	4.196	1	0.041
g_2 - quadratic	0.051	0.002	529.776	1	0.000
h_2	0.144	0.002	4425.684	1	0.000
h_2 - quadratic	0.026	0.002	112.026	1	0.000
Scale	0.006	0.000			

The results for the services context, in Table 27, are similar to the *fmcg* one. The linear effect of the focal firm's number of alternatives was significant (Wald $\chi^2[1] = 393.6$, $p \leq 0.01$) as well the quadratic effect (Wald $\chi^2[1] = 91.5$, $p \leq 0.01$). But the combined effect implies that when the number of alternatives increased from its smallest level the absolute difference moves from negative to neutral, at intermediate levels of the dependent variable the response was less steep, and at higher levels the absolute difference in the focal firm's choice probabilities became positive as it can be observed in Figure 15 (a).

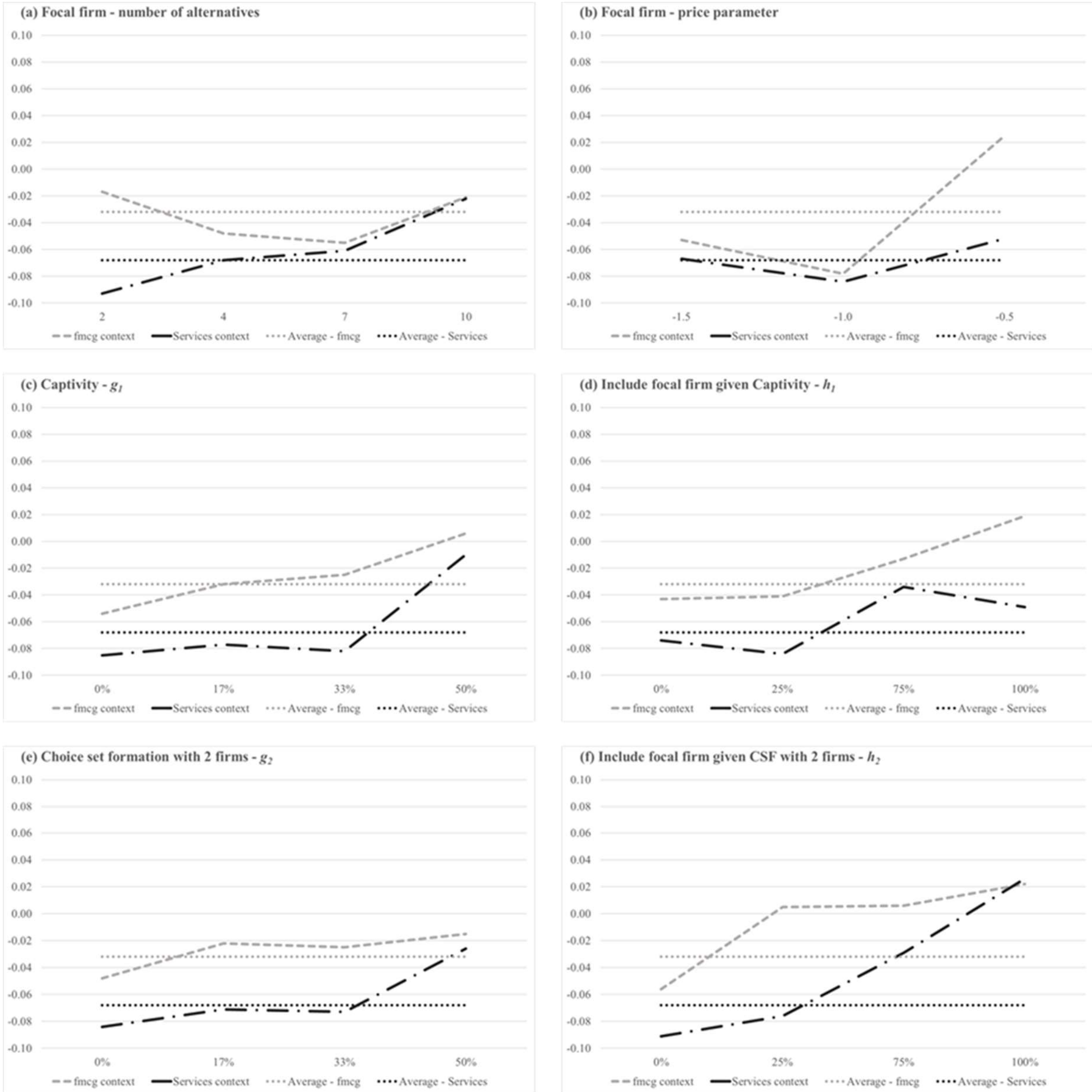

Figure 15 - Absolute difference in focal firm's choice probability as a function of the independent variables

The focal firm choice probabilities' response to its price parameter manipulation was similar in both contexts. In the services context the linear effect was significant (Wald $\chi^2[1] = 81.4$, $p \leq 0.01$) and so was the quadratic effect (Wald $\chi^2[1] = 5.7$, $p \leq 0.05$). However, the quadratic effect was less pronounced as it can be noticed from a less steep response of the independent variable to the firm's *A* price parameter when it moved from -1.0 to -0.5 (see Figure 15 (b)).

The parameters describing the relationship between choice set formation and the absolute difference in the focal firm's choice probability were all significant and substantively similar to results found in the *fmcg* context, as can be noticed in Figure 15 (c) to (f). Once again, it is important to notice that the absolute difference in the focal firm's choice probability, averaged across experimental conditions, was largest in the services context. See in the horizontal lines in Figure 15 that in the *fmcg* the average deviation was around 3% while in the services context it was around 7%.

b. Demand elasticities of choice probabilities

The demand elasticity measures the expected relative change in the product's probability of being chosen as a function of a relative change in any of its attributes. More specifically, it informs, for instance, the relative change in choice probability given a 1% change in the product's price and it is estimated as (Louviere et al., 2000, p. 59):

(4-21)

$$E_{X_{ixn}}^{P_{in}} = \beta_{ixn} X_{ixn} (1 - P_{in})$$

Where P denotes the choice probability, X is the vector of attributes, i indexes the alternatives in the choice scenario, x indexes the attribute in X and n indexes the individual. It means that choice elasticities are estimated per alternative, per attribute and at the individual level. Given that alternative-specific constants and price parameters varied by firm and that other parameters were generic, the choice elasticities were also aggregated at the firm level allowing for comparison across experimental conditions. The aggregation was done through weighting each choice elasticities by the relative share of each alternative within the total share of the focal firm, or:

(4-22)

$$P_i^s = \frac{P_i}{\sum_{j \in f} P_j}$$

Where P_i^s is the rescaled probability of each alternative, adding up to one, offered by the focal firm *f*. Finally, the focal firm choice elasticities are given by:

$$E_{X_{xn}}^{P_{fn}} = \sum_{i \in f} P_i^s \, . \, E_{X_{ixn}}^{P_{i_s n}} \tag{4-23}$$

The expression above states that the firm *f*'s attributes choice elasticities are the attributes choice elasticities of its individual products weighted by their rescaled probabilities of choice. Table 28 presents the focal firm's true and biased choice elasticities for the *fmcg* context, as well as the differences between them. In general the biased choice elasticities, as seen by the analyst, had the right signal but were much larger that the true ones (in average twice as large as can be seen in the last line of the table), what would let the firm to set prices lower than the maximizing prices or to improve quality (through generic attributes) expecting that consumers would pay more for it than they were actually willing to pay. The firm-specific constant was the exception for the correct signaling of the choice elasticities. Given the highest focal firm-specific constant in the true data generation process, the true firm *A* specific constant choice elasticity should be always positive. However, in most of the experimental conditions, the FSC choice elasticity was not only lower than the true ones, but they were negative. It means that the analyst would view the consumers demanding a discount to consume the focal firm's products when they were actually willing to pay a premium.

Table 28 – *fmcg* context - Focal firm's choice elasticities

	(1) True choice elasticities				(2) Biased choice elasticities				(2) – (1)			
	FSC	gen1	gen2	price	FSC	gen1	gen2	price	FSC	gen1	gen2	price
sku (2)	0.27	-0.98	2.47	-1.14	-4.07	-2.13	4.87	-2.41	-4.34	-1.15	2.40	-1.27
sku (4)	0.29	-1.14	2.82	-1.49	0.41	-2.24	5.50	-2.94	0.13	-1.10	2.68	-1.45
sku (7)	0.27	-1.14	2.60	-1.47	-2.35	-2.26	5.53	-2.79	-2.62	-1.12	2.93	-1.32
sku (10)	0.31	-1.30	2.99	-1.51	1.38	-2.19	5.25	-2.65	1.06	-0.89	2.26	-1.15
pp (-0.5)	0.27	-0.88	2.56	-0.85	0.11	-1.60	4.87	-1.76	-0.16	-0.72	2.32	-0.92
pp (-1.0)	0.28	-1.18	2.67	-1.44	-0.85	-2.35	5.34	-2.92	-1.13	-1.17	2.67	-1.48
pp (-1.5)	0.30	-1.42	2.91	-2.18	-7.23	-3.05	5.61	-3.85	-7.54	-1.63	2.70	-1.67
g_1 (0)	0.36	-1.40	3.38	-1.75	0.08	-2.12	5.09	-2.59	-0.27	-0.72	1.71	-0.84

	(1) True choice elasticities				(2) Biased choice elasticities				(2) – (1)			
	FSC	**gen1**	**gen2**	**price**	**FSC**	**gen1**	**gen2**	**price**	**FSC**	**gen1**	**gen2**	**price**
g_1 (0.17)	0.30	-1.26	2.86	-1.41	-0.79	-2.36	5.24	-2.57	-1.09	-1.10	2.37	-1.16
g_1 (0.34)	0.22	-0.82	2.13	-1.15	-2.80	-2.08	5.29	-2.87	-3.02	-1.25	3.16	-1.73
g_1 (0.5)	0.17	-0.65	1.60	-0.70	-5.29	-2.28	5.31	-2.59	-5.46	-1.62	3.70	-1.89
g_2 (0)	0.33	-1.31	3.14	-1.66	0.23	-2.12	5.12	-2.53	-0.10	-0.81	1.97	-0.88
g_2 (0.17)	0.29	-1.13	2.79	-1.44	-0.10	-2.12	5.11	-2.77	-0.40	-0.99	2.32	-1.33
g_2 (0.34)	0.26	-1.08	2.48	-1.30	0.48	-2.24	5.31	-2.71	0.22	-1.17	2.83	-1.42
g_2 (0.5)	0.19	-0.72	1.80	-0.71	-9.54	-2.36	5.37	-2.66	-9.73	-1.64	3.57	-1.95
h_1 (0)	0.24	-0.93	2.35	-1.14	-4.03	-2.27	5.37	-2.68	-4.27	-1.34	3.01	-1.53
h_1 (0.25)	0.29	-1.18	2.76	-1.46	-0.37	-2.17	5.28	-2.68	-0.66	-0.99	2.53	-1.22
h_1 (0.75)	0.31	-1.28	2.88	-1.39	-1.84	-2.28	5.20	-2.60	-2.15	-1.00	2.32	-1.20
h_1 (1)	0.32	-1.23	3.02	-1.62	1.57	-1.97	4.80	-2.58	1.25	-0.74	1.79	-0.97
h_2 (0)	0.26	-0.95	2.49	-1.22	-3.53	-2.11	5.18	-2.63	-3.79	-1.16	2.70	-1.41
h_2 (0.25)	0.28	-1.13	2.68	-1.54	-0.67	-2.15	5.23	-2.84	-0.95	-1.02	2.55	-1.30
h_2 (0.75)	0.29	-1.18	2.75	-1.33	-0.35	-2.23	5.21	-2.58	-0.64	-1.05	2.46	-1.25
h_2 (1)	0.31	-1.33	2.96	-1.44	-0.62	-2.36	5.22	-2.52	-0.93	-1.03	2.26	-1.07
mean	**0.28**	**-1.11**	**2.66**	**-1.36**	**-1.75**	**-2.22**	**5.23**	**-2.68**	**-2.03**	**-1.10**	**2.57**	**-1.32**

The firm *A* choice elasticities for the services context are detailed in Table 29 and some differences, compared to the *fmcg* context, are noticeable. Adding variability to the levels of the generic attribute broke the regularity observed between true and biased choice elasticities reported in the *fmcg* context for the generic attributes and the firm-specific prices. A comparison of the last lines of Table 28 and Table 29 reveals that in the services context I observed a larger bias in firm-specific choice elasticity (the signal was still reversed). The generic 2 choice elasticity was now underestimated (the signal was still correct) and the overestimation of the specific price choice elasticity was attenuated. The pattern for the remaining attribute (*gen1*) was similar in both contexts.

Table 29 – Services context - Focal firm's choice elasticities

	(1) True choice elasticities				(2) Biased choice elasticities				(2) – (1)			
	FSC	gen1	gen2	Price	FSC	gen1	gen2	price	FSC	gen1	gen2	price
sku (2)	0.27	-0.98	2.47	-1.14	-7.09	-2.67	1.68	-0.26	-7.36	-1.69	-0.80	0.88
sku (4)	0.29	-1.14	2.82	-1.49	0.46	-2.23	2.19	-2.49	0.17	-1.10	-0.63	-1.00
sku (7)	0.27	-1.14	2.60	-1.47	-1.53	-2.21	2.38	-2.92	-1.80	-1.07	-0.22	-1.45
sku (10)	0.31	-1.30	2.99	-1.51	2.24	-2.00	2.21	-2.66	1.93	-0.70	-0.78	-1.15
pp (-0.5)	0.27	-0.88	2.55	-0.84	-3.78	-1.74	1.75	1.71	-4.05	-0.86	-0.80	2.56
pp (-1.0)	0.28	-1.18	2.67	-1.44	0.30	-2.49	2.13	-3.53	0.02	-1.30	-0.54	-2.09
pp (-1.5)	0.30	-1.42	2.91	-2.18	-6.04	-3.34	2.36	-4.95	-6.35	-1.92	-0.55	-2.76
g_1 (0)	0.36	-1.40	3.38	-1.75	-1.16	-2.35	1.95	-1.46	-1.51	-0.95	-1.44	0.29
g_1 (0.17)	0.30	-1.26	2.86	-1.41	0.55	-2.38	2.09	-3.01	0.25	-1.12	-0.77	-1.60
g_1 (0.34)	0.22	-0.82	2.13	-1.14	-6.55	-2.47	2.01	0.02	-6.77	-1.65	-0.12	1.16
g_1 (0.5)	0.17	-0.65	1.60	-0.70	-4.70	-2.22	2.14	-2.67	-4.87	-1.57	0.54	-1.97
g_2 (0)	0.33	-1.31	3.14	-1.66	-0.59	-2.36	1.92	-1.72	-0.92	-1.06	-1.23	-0.06
g_2 (0.17)	0.29	-1.13	2.79	-1.44	-1.51	-2.33	1.92	-1.23	-1.80	-1.20	-0.87	0.21
g_2 (0.34)	0.26	-1.08	2.49	-1.30	-1.78	-2.43	2.15	-1.17	-2.04	-1.35	-0.33	0.13
g_2 (0.5)	0.19	-0.72	1.80	-0.70	-8.54	-2.30	2.23	-2.75	-8.73	-1.59	0.43	-2.04
h_1 (0)	0.24	-0.93	2.35	-1.14	-4.80	-2.48	2.07	-1.66	-5.04	-1.55	-0.28	-0.52
h_1 (0.25)	0.29	-1.18	2.76	-1.46	-2.51	-2.45	2.06	-1.34	-2.79	-1.28	-0.70	0.12
h_1 (0.75)	0.31	-1.28	2.88	-1.39	-1.26	-2.28	2.12	-2.79	-1.57	-1.01	-0.76	-1.40
h_1 (1)	0.32	-1.23	3.02	-1.62	0.35	-2.09	1.81	-1.13	0.04	-0.85	-1.21	0.49

	(1) True choice elasticities				(2) Biased choice elasticities				(2) – (1)			
	FSC	**gen1**	**gen2**	**Price**	**FSC**	**gen1**	**gen2**	**price**	**FSC**	**gen1**	**gen2**	**price**
h2 (0)	0.26	-0.95	2.49	-1.22	-5.47	-2.34	1.97	-0.78	-5.72	-1.39	-0.52	0.44
h2 (0.25)	0.28	-1.13	2.68	-1.54	-3.08	-2.44	2.09	-1.31	-3.36	-1.31	-0.59	0.23
h2 (0.75)	0.29	-1.18	2.75	-1.33	0.63	-2.33	2.05	-2.85	0.34	-1.15	-0.69	-1.52
h2 (1)	0.31	-1.33	2.96	-1.44	0.36	-2.33	2.06	-2.87	0.05	-1.00	-0.89	-1.43
mean	**0.28**	**-1.11**	**2.66**	**-1.36**	**-2.41**	**-2.36**	**2.06**	**-1.91**	**-2.69**	**-1.25**	**-0.60**	**-0.54**

In summary, the descriptive analysis of the attributes' choice elasticities indicated that the biases in the preference parameters, originated by not accounting for choice set formation, were strong enough to affect the magnitude of all the choice elasticities and to cause a reversal in the signal in the firm-specific constant choice elasticity.

From the definition of attribute choice elasticity, I can formalize the root mean square error of the focal firm choice elasticities as:

(4-24)

$$RMSE(E_r^A) = \sqrt{\sum_{k=1}^{4} \frac{\left(\hat{E}_{X_{ixn}}^{P_{An}} - E_{X_{ixn}}^{P_{An}}\right)^2}{4}}$$

The root mean square of the focal firms choice elasticities was analyzed through a generalized linear model (GENLIN) fit to a gamma distribution, due to the extreme values observed in some conditions, with a log link. The model is expressed as:

(4-25)

$$\begin{aligned} RMSE(E_r^A) = c + \gamma_1 nalt'_A + \gamma_2 nalt'^2_A + \gamma_3 pp'_A + \gamma_4 pp'^2_A + \gamma_5 g'_1 + \gamma_6 g'^2_1 \\ + \gamma_7 h'_1 + \gamma_8 h'^2_1 + \gamma_9 g'_2 + \gamma_{10} g'^2_2 + \gamma_{11} h'_2 + \gamma_{12} h'^2_2 + scale \end{aligned}$$

This model result is presented in Table 30 for the *fmcg* context and in Table 31 for the services context. The graphs that support the interpretation of the statistical effects and the comparison between the two contexts are ordered in Figure 16.

The combined significant linear (Wald $\chi^2[1] = 36.7, p \leq 0.01$) and quadratic effect (Wald $\chi^2[1] = 19.0$, $p \leq 0.01$) effects of the portfolio size generated a peculiar bi-modal response of the dependent variable, peaked at the first and third levels of the focal brand number of alternatives, as can be observed in Figure 16 (a). An examination of Table 28 tells that this pattern was driven by the firm A specific constant while the other attributes choice elasticities tended to peak at the intermediate levels of this independent variable. There is no substantive reason to expect this behavior from the FSC choice elasticity that I may, speculatively, attribute to the larger bias and variance in the estimated parameter across replications. The change in the firm-specific constant's choice elasticity across the different levels of the number of alternatives seems to support this explanation, that would mean that the strong bias in the FSC prevents the establishment of a proper location to this attribute's choice elasticity.

Table 30 – *fmcg* context - Results for $RMSE(E_r^A)$

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	df	Sig.
Intercept	0.405	0.006	5367.656	1	0.000
NALT - linear	-0.045	0.007	36.685	1	0.000
NALT - quadratic	0.030	0.007	18.959	1	0.000
pp - linear	-0.279	0.007	1478.717	1	0.000
pp - quadratic	0.061	0.007	80.966	1	0.000
g_1	0.930	0.007	16910.893	1	0.000
g_1 - quadratic	-0.240	0.007	1084.916	1	0.000
h_1	-0.344	0.012	823.310	1	0.000
h_1 - quadratic	-0.003	0.011	0.055	1	0.814
g_2	1.032	0.007	19877.588	1	0.000
g_2 - quadratic	-0.183	0.007	620.700	1	0.000
h_2	-0.503	0.011	2107.707	1	0.000
h_2 - quadratic	0.289	0.011	714.986	1	0.000
Scale	0.069	0.001			

The sign of the significant linear effect of the price parameter (Wald $\chi^2[1] = 1478.7$, $p \leq 0.01$) implies that decreasing the own demand price elasticity of the focal firm (or increasing its price parameter) decreased the bias introduced by the design in the focal firm elasticities.

This effect was attenuated at the lower level of the dependent variable by a significant quadratic effect (Wald $\chi^2[1] = 81.0$, $p \leq 0.01$), as can be seen in Figure 16 (b). This result suggests that the larger was the market power, expressed as own demand price elasticity, the smaller was the bias in the choice elasticities generated by the misattribution of choice process heterogeneity into preferences.

The significant linear effect of g_1 (Wald $\chi^2[1] = 16910.9$, $p \leq 0.01$) discloses that increasing choice set formation in favor of any one firm only increased the bias in the focal firm choice elasticities. The significant quadratic effect of g_1 (Wald $\chi^2[1] = 1085.0$, $p \leq 0.01$) indicated that the effect increased as it did the level of g_1 (see Figure 16 (c)). However, the negative significant effect of h_1 (Wald $\chi^2[1] = 823.3$, $p \leq 0.01$) reveals that favoring the focal firm, given choice set formation with one firm only, attenuated the increase in the *RMSE(E)*. Figure 16 (d) discloses mixed evidence of the relationship between h1 and the dependent variable and the details in Table 28 confirms that while the difference between true and biased FSC and *gen1* choice elasticities decreased in h_1, the same difference for *gen2* increased and there is no clear signal for the focal firm-specific price.

The effects of choice set formation with any two firms on the own elasticities of the focal firm followed the same patterns of the choice set formation with one firm only, with a positive significant effect of g_2 (Wald $\chi^2[1] = 19877.6$, $p \leq 0.01$) revealing that as choice set formation increased so did the bias in the own elasticities of the focal firm. Moreover, the quadratic effect of g_2 was also significant (Wald $\chi^2[1] = 620.7$, $p \leq 0.01$) indicating that the higher the level of g_2 the more steep is the effect, as observable in Figure 16. The significant negative linear effect of h_2 (Wald $\chi^2[1] = 2107.7$, $p \leq 0.01$) reveals that when the choice set formation benefited the focal firm, given rules including any two firms, the biases created by the design were attenuated. There was also a significant quadratic effect of h_2 (Wald $\chi^2[1] = 715.0$, $p \leq 0.01$) most likely created by the nested structure between g_2 and h_2.

Turning to the services context (Table 31) the linear (Wald $\chi^2[1] = 2594.9$, $p \leq 0.01$) and quadratic (Wald $\chi^2[1] = 6064.7$, $p \leq 0.01$) effects of the focal firm's number of alternatives were significant, but Figure 16 (a) suggests a great drop in the $RMSE(E_r^A)$ from the first to the second level of the independent variable, and a stabilization after this point. A close examination of Table 29 indicates that this drop was driven by the firm-specific constant and that the evidence of the relationship among deviations in attributes choice elasticities and the independent variable were mixed.

The linear effect of the price parameter was significant (Wald $\chi^2[1] = 831.7, p \leq 0.01$) and so was the quadratic effect (Wald $\chi^2[1] = 3339.2, p \leq 0.01$). The combination of this effect resulted in a U-shaped (see Figure 16 (b)) relationship between dependent and independent variables. It means that adding variability to the generic attributes minimized the $RMSE(E_r^A)$ when the firm A price parameter equaled the competitors. It was an expected pattern since that when the focal firm price parameter was small, the bias in the elasticities to accommodate choice set formation including firm A needed to be larger. And so it was when it was necessary to accommodate choice set formation excluding firm A, if its price parameter was larger than the ones of the competitors.

Table 31 – Services context - Results for $RMSE(E_r^A)$

Parameter	γ	Std. Error	Hypothesis Test		
			Wald Chi-Square	df	Sig.
Intercept	0.086	0.008	108.851	1	0.000
NALT - linear	-0.566	0.011	2594.866	1	0.000
NALT - quadratic	0.864	0.011	6064.663	1	0.000
pp - linear	-0.247	0.009	831.717	1	0.000
pp - quadratic	0.492	0.009	3339.178	1	0.000
g_1	0.144	0.009	254.065	1	0.000
g_1 - quadratic	-0.478	0.013	1410.179	1	0.000
h_1	0.028	0.010	7.587	1	0.006
h_1 - quadratic	0.121	0.016	58.236	1	0.000
g_2	1.019	0.011	8777.922	1	0.000
g_2 - quadratic	-0.329	0.010	984.342	1	0.000
h_2	-1.430	0.010	21264.331	1	0.000
h_2 - quadratic	0.434	0.017	677.506	1	0.000
Scale	0.135	0.002			

The linear effect of g_1 was significant (Wald $\chi^2[1] = 254.1, p \leq 0.01$), as well as the quadratic effect (Wald $\chi^2[1] = 1410.2, p \leq 0.01$) and they imply that the $RMSE(E_r^A)$ was increasing in captivity, and the more captivity in the data the steeper was the response of the dependent variable (see Figure 16 (c)). As in the *fmcg* context, the linear effect of h_1 (Wald $\chi^2[1] = 7.6$, $p \leq 0.01$) was significant and so was the quadratic effect (Wald $\chi^2[1] = 58.2$, $p \leq 0.01$). But as it can be noticed in Figure 16 (d), no clear pattern emerged due to the mixed evidence observed from the different attributes in Table 29.

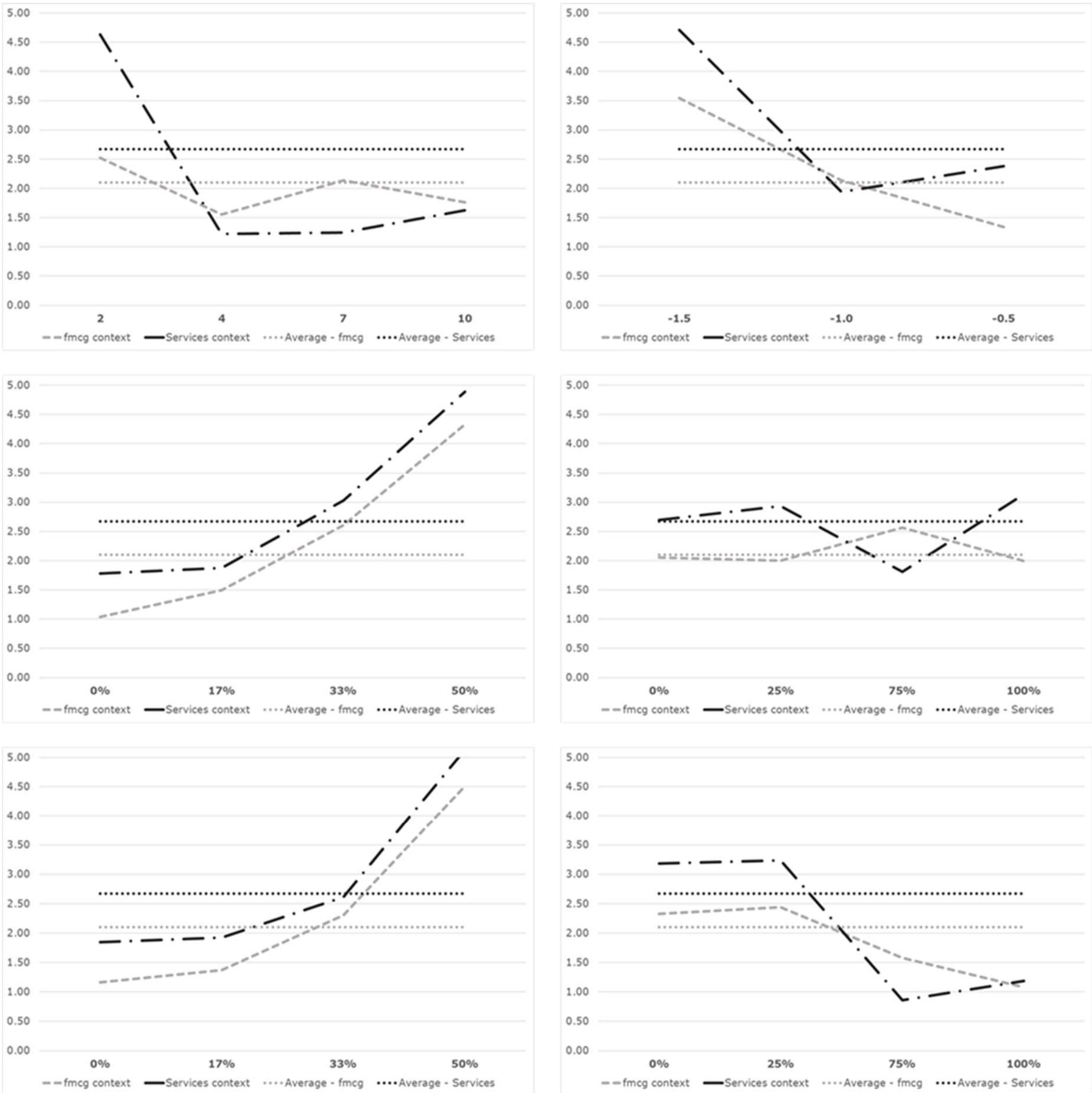

Figure 16 - RMSE(E) as a function of the independent variables

Lastly, the linear effect of choice set formation with any two firms, g_2, was significant (Wald $\chi^2[1] = 8778.0$, $p \leq 0.01$) as it was the quadratic effect (Wald $\chi^2[1] = 984.3$, $p \leq 0.01$). The observation of Figure 16 (e) confirms that the $RMSE(E_r^A)$ increased in g_1 and the most often was choice set formation with any two firms the steeper was the response of the dependent variable. The linear effect of h_2 was significant (Wald $\chi^2[1] = 21.264.3$, $p \leq 0.01$) and the inclusion of the focal firm in the choice set formation reduced the $RMSE(E_r^A)$. There was also a quadratic effect caused by the nested structure of g_2 and h_2 (see Figure 16 (f)).

This pattern of the choice set formation effects on the $RMSE(E_r^A)$ reveals that not accounting for the choice process increased the bias in the own elasticities of the focal firm, but this spurious effect was diminished if choice set formation benefited this firm. There are some situations when the relationship between the dependent variable and the experimental factors were not clear, but this fact seems to be caused by the mixed experimental effects over the choice elasticities of different attributes. But, at the bottom line, all the elasticities are affected. Finally, the effects were similar in both contexts and, once again, the magnitude of the bias was larger in the service context.

4.3.3 Market equilibrium

To analyze the market equilibrium under the true and the biased demand representations, I need some assumptions about the firms' behavior and about the market. At this point, I bases this analysis in a statistic game of complete information. It means that players simultaneously choose actions and then they receive the payoff depending on the combination of chosen actions; moreover, the players' payoff functions are common knowledge among all the players. The players are also rational, meaning that each maximizes its own payoff, given its choice and the other players' choices. The solution for this game should be a Nash Equilibrium (NE), understood as the set of strategies in which each player's strategy is the best response to others players' strategies, and this is a condition to prevent any player to deviate from the theory's prediction. Finally, in this game, the firm will choose quantities to offer from the common knowledge information that is available. It means that I am examining the Nah Equilibrium to a Cournot game (Gibbons, 1992). Notice that this the simplest form of the game, but since it is enough to demonstrate the effects that I want to study, i.e. the effect of misattributing choice set heterogeneity to tastes, I will avoid, by now, the more complex forms of this game.

The logic adopted was that there are two sets of solutions for this game. The first, a discrete solution, is the true one, i.e. had the firms known the true demand functions their behaviors would inform the individual and the aggregate payoffs. The second one is a set of solutions, each one corresponding to one choice set formation structure and to one of the experimental conditions attributed to the focal firm (in terms of the number of alternatives offered and price parameters).

The NE means that choice probabilities predicted to the firms represent the solution in which each firm is maximizing its payoff and each firm maximize its payoff when the marginal cost equals the marginal revenue. Given that prices are known from the choice task and the demand price elasticities were estimated using equation (4-23), it is possible to compute the marginal cost as (Pindyck & Rubinfeld, 2013 p.363):

(4-26)

$$MC_i = MR_i = pr_i + pr_i\left(\frac{1}{E_i}\right)$$

In expression (4-26), *MC* is marginal cost, *MR* is marginal revenue, *pr* is price and E_i is the demand price elasticity of alternative i. Remember that in equation (4-23) the choice elasticities were aggregated at the firm level, and this was necessary to run the simulation since the number of products offered by the focal firm varies preventing comparison across experimental conditions.

Hence, I must aggregate equation (4-26) at the firm level, to be able to use the firms' level choice elasticities, and this step requires to weight prices by firm's choice probabilities. Given that prices are constant across choice tasks and individuals, this aggregation can be expressed at the sample level as:

(4-27)

$$pr_f = \sum_{i\in f} P_i^s \,.\, pr_i$$

Therefore, all the information needed to estimate the marginal cost at the firm level, i.e. averaged across alternatives, is available from the true data generation process, and from the Monte Carlo simulation the price elasticities choice probabilities are available as a function of the experimental factors and can be estimated for any choice set formation structure, as:

(4-28)

$$MC_f = MR_f = pr_f + pr_f\left(\frac{1}{E_{pr}^{P_f}}\right)$$

Observe that in equation (4-28) the index *n* is absent, indicating that the elasticities were averaged across respondents. Also, the subscript E_x was replaced by *pr* since the only choice elasticity used, at this point, was the price elasticity. Notice also that the same equation (4-28) informs that the larger is the demand price elasticity the higher the marginal cost will be, and since the price is given, a lower expected profit result from increasing elasticities. A simple examination of Table 28 and of Table 29 disclose that demand prices elasticities tend to increase in the biased demand representation for the firm *A*, suggesting that at least firm *A* should obtain a smaller profit per unit if playing the biased game. The marginal costs were estimated from the demand Monte Carlo simulation, to understand the overall impact of the bias. And, finally, the firms' payoff, or profit, is given by:

(4-29)

$$\pi_f = N * P_f * \left(pr_f - MC_f\right)$$

Where π_f is firm *f* profit, *N* is the number of consumers in the market, P_f is firm *f* choice probability, pr_f is firm *f* weighted price and MC_f is firm *f* marginal cost. Notice that *N* is unknown, since we have created the conditional demand functions in the data generation process. But, dropping *N* from equation (4-29) it is possible to understand the dynamic of any firm's payoff, and now:

(4-30)

$$\widetilde{\pi_f} = P_f * \left(pr_f - MC_f\right)$$

As in equation (4-29), $\tilde{\pi}_f$ is increasing both in the choice probability as it is in the unit price $(pr_f - MC_f)$, and it is capturing the net effect of variation in the choice probabilities and in the marginal costs across experimental conditions. The conditional market level profit is the sum of the firms' unit profits weighted by the respective choice probabilities, i.e.:

(4-31)

$$\tilde{\pi} = \sum_f P_f * \left(pr_f - MC_f\right)$$

The resulting firm's *A* and market level payoffs in the *fmcg* context are presented in Figure 17. The bottom (darker) solid area in each graph depicts firm's A payoff, when misattributing demand choice process heterogeneity into tastes. The shaded (diagonal lines) area adds to the solid one to inform the payoff of correct modeling the demand, i.e. choice process heterogeneity and taste homogeneity. It means that the shaded are only informs the firm's A lost payoff due demand misspecification. The next solid area in the middle of the graph tell about the aggregated payoff when the demand function is biased, and the shaded area above is the true market level payoff.

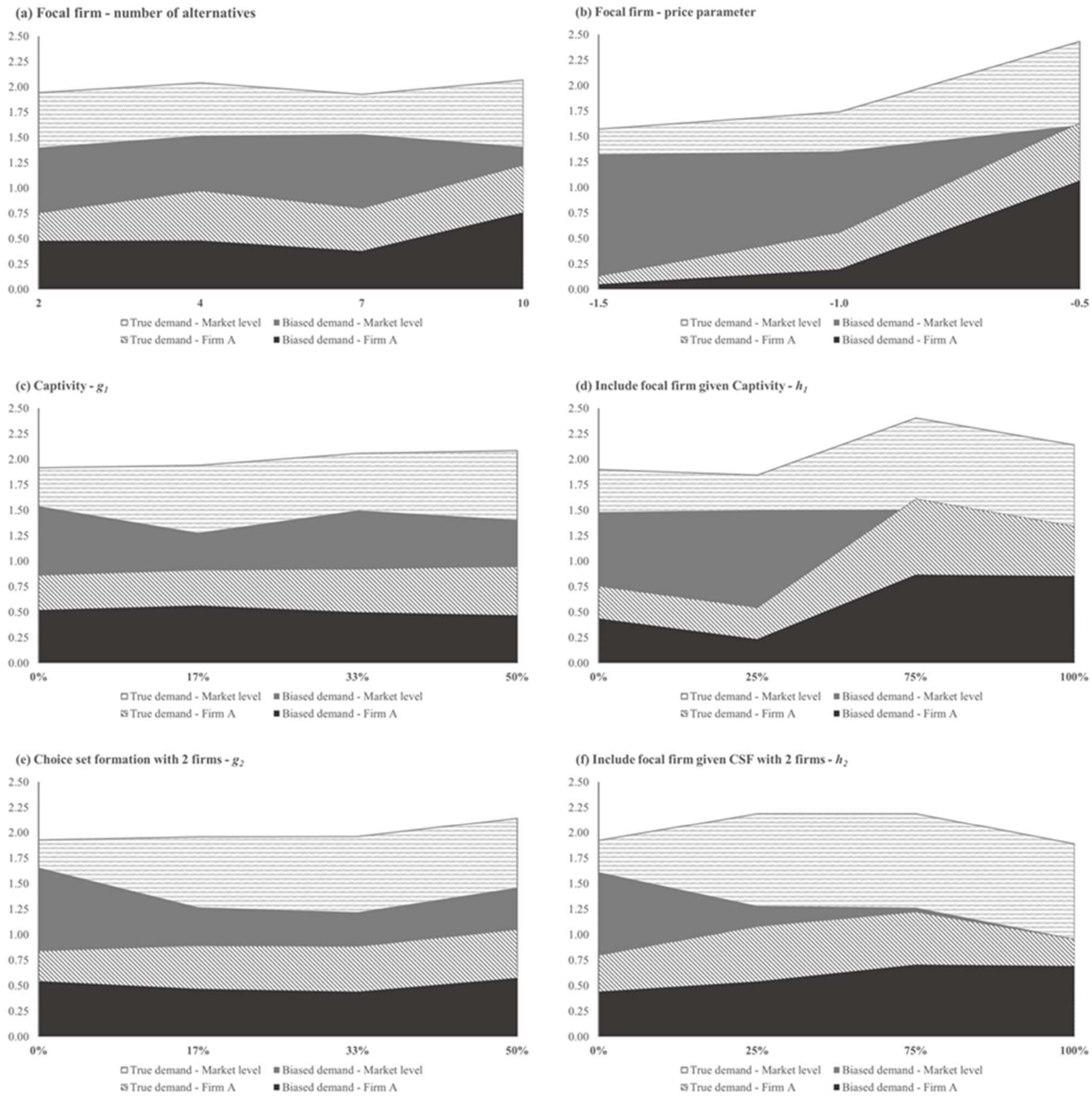

Figure 17 - *fmcg* context - Focal firm and market level payoffs

It possible to conclude that misattributing choice process heterogeneity, here choice set formation, to tastes, caused a profit reduction in the *fmcg* context. These losses were quite regular over the experimental conditions. Firm's A payoff was reduced approximately by 25% to a 65% as a function of the parameters controlling choice set formation, see panels (c) and (e). Even when its price parameter reduced its elasticity, see the last column of graphic (b), its payoff was only around 65% of what it could be. And the market level payoff was also notably under-realized when demand representation is biased, with reduction ranging from around 15% to around 50%.

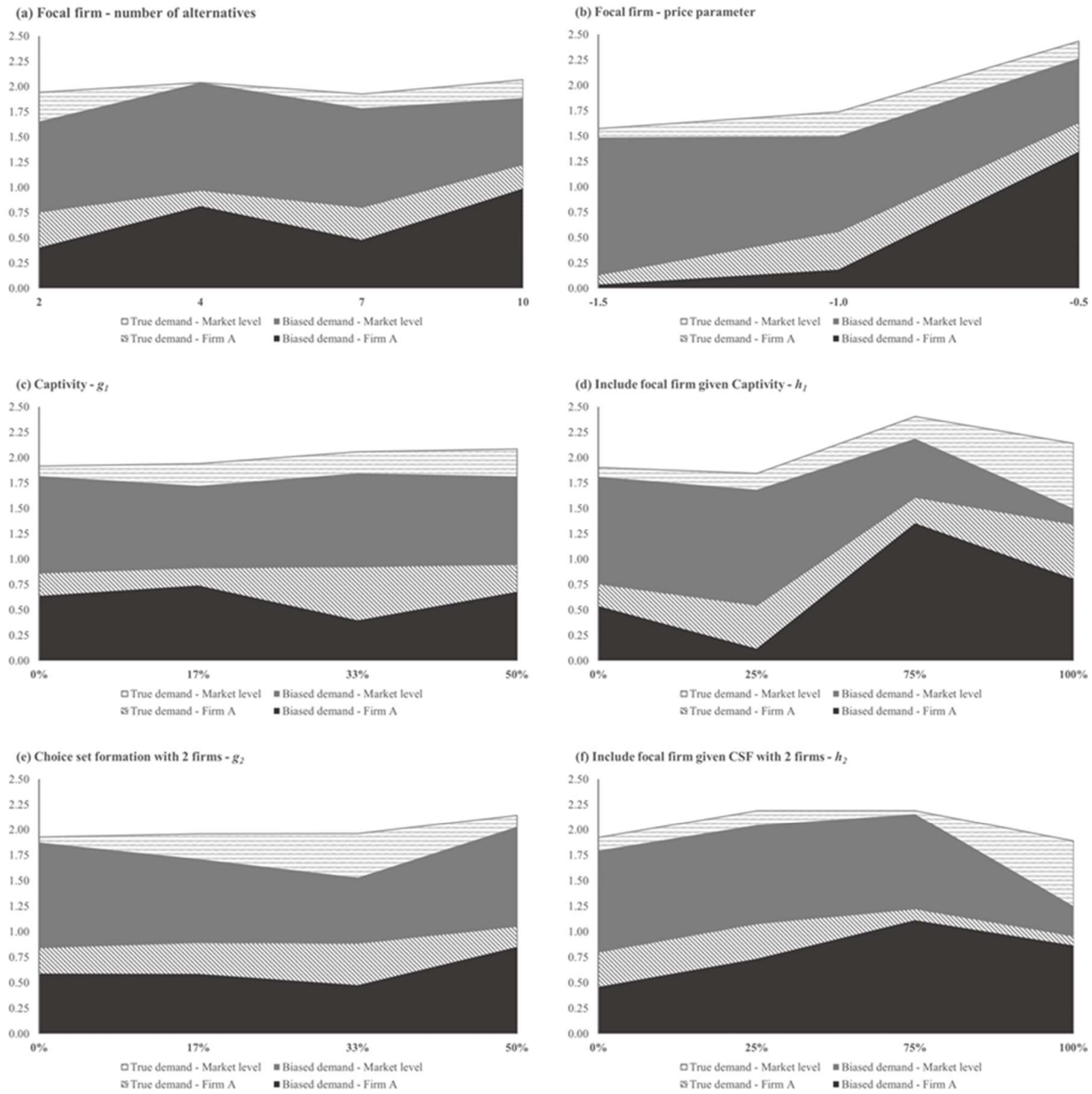

Figure 18 - Services context - Focal firm and market level payoffs

The information is presented in Figure 18, now for the services context. The focal firm payoff reduction was even more dramatic in this context, ranging from approximately 10% to 80% for the focal. However it was less intense at the market level, ranging from approximately 0% to 33%. A possible explanation is that the variation added to the generic attributes, to configure the consumer experience variability in the context service, may have protected the price attributes to accommodate the bias, allowing competitors to benefits from the focal firm problems. If this was the case, other consequences would be caused by larger bias in the generic attributes choice elasticities, but this not my focus here.

4.3.4 Final considerations

Behavioral decision theories have a long tradition in the study of heterogeneity in the choice process, as a source of violations in the normative rationality imposed to the *homo economicus*. Almost every kind of heterogeneity examined in the literature results in choice set heterogeneity.

Although this is a well-known phenomenon among choice modelers, it is still challenging the DCM community. The most precise way to model the phenomenon is by enumeration at the choice task level, but the set of possible choice set (2^J-1) increases exponentially with the cardinality of the choice task, and this is one of the reasons preventing this brute force solution to widespread in applied choice models.

In this section I have conducted a Monte Carlo experiment to study the effects, on the firm decision-making, of misattributing choice process heterogeneity, here choice set formation, to tastes. After generating a true and a biased representation I have used these results to study the Nash Equilibrium in a static game of complete information. The results reveal that firms will be unequivocally driven away from profit-maximization and this damage will affect specific firms and the aggregated industry. The intensity of the phenomenon is context dependent, but its occurrence is not. The process will happen through: (i) generalized biased estimation of preference parameters; (ii) biased estimation of policy measures, here choice probabilities and attributes' choice elasticities; (iii) firm decision-making producing a Nash Equilibrium in which the focal and the aggregated market ends worse-off than if the choice set formation had been accounted for.

As it is any enterprise ambitious like this, I had to make important methodological choices, leaving room for improvements and extensions to this study. At the demand side, preferences were held homogeneous when it is well established that tastes vary between consumers and across occasion. Future studies may let these parameters free, allowing for more realistic and interesting cross-substitutability patterns. Also, different market contexts may be studied with more variability added to the attributes, including prices. There is a multitude of market characterizations that can be explored, in terms of the nature of the products and the competitive structure. At the firm level, there are also interesting opportunities. I have studied a static Nash Equilibrium for a Cournot game of complete information, but the assumptions of this kind of game may be relaxed to become more realistic. It is reasonable to consider that complete information is not available, i.e. in real markets information is usually incomplete and often asymmetric what motivates the study of Bayesian Nash Equilibrium. In some kind of markets, even the Cournot game may be replaced by a Bertrand game.

But what motivated this study was to build a compelling "call to action" argument to the choice modeler community and to the users of DCMs to assimilate choice process heterogeneity. In stated preferences studies, where even choice tasks of 12 alternatives are rarely used, there are no apparent reasons to avoid facing the challenge. In revealed preference, the issue is more complex but the consequence of ignoring can be enormous.

The complexity that we face is the outcome of the kind of decision-makers we are, pursuing multiples goals, smartly adapting to the environment, learning and forgetting, reasoning and simplifying as a way to make the most out of our biological machinery. In essence, the *homo economicus* is just a tiny serving of the amazing species, the *homo aptabilis*.

5 MULTIPLE META-GOALS BASED CHOICE: BALANCING REASONING AND EMOTION IN THE *HOMO APTABILIS* BEHAVIOR

5.1 Introduction

This project follows the proposition that stochasticity in choice behavior arises from individuals pursuing multiple and contradictory goals (Marley & Swait, 2017; Swait & Marley, 2013). It describes an econometric discrete choice model that accounts for behavioral rules heterogeneity as a way of balancing multiple meta-goals during the choice process, allowing the consumer to use a mixture of decision strategies when approaching a choice scenario. Moreover, the model assimilates the concept of archetypes (Li, 2013; Swait et al., 2016) allowing the individual to be consistent or adaptive in the way that meta-goals are pursued. The consistent consumer adopts a constant mixture of goals across choice occasions while the adaptive adjusts the mixture across occasions, responding to variations in the decision-making environment. Differently from the previous models adopting the archetypical perspective of the decision maker, in which consistency meant the adoption of a pure strategy as implied in the classical latent class choice model (Kamakura & Russell, 1993), this is a more general case in which consistency means a constant mixture of goals. The econometric model is tested using a Monte Carlo simulation that confirms that it is able to properly recover the true parameters if the proposed behaviors occur in the focal population.

Meta-goals, also known as process goals, are goals in which the desired end state relates to the choice process itself, instead of to its outcomes (Austin & Vancouver, 1996; Bettman et al., 1998; Dellaert et al., 2017; van Osselaer et al., 2005). To maximize choice accuracy and to minimize negative emotion resulting from the choice are two of the most important meta-goals, as already identified in the initial sections of this thesis, and they are going to be used to develop this section. Besides, they express a duality described in the goal literature that classifies this cognitive structure as having positive or negative desired consequences, causing approach (maximize accuracy) or avoidance (minimize regret) behaviors (Austin & Vancouver, 1996).

The pure strategy of the *homo economicus* is choice accuracy, described by the expected utility maximization decision-making. In the proposed model, the expected utility maximization is still present but as part of a mixed decision strategy, that may be constant or may be adapted while the *homo aptabilis* deploys her abilities to navigate the environment in search of the best approximation to her multiple goals. Notice that the *homo economicus* may still be present as a special case of the *homo aptabilis*, i.e., the case in which the consumer places all the weight, divided among goals, in expected utility maximization.

Negative emotion minimization is an outstanding meta-goal that has been typified in behavioral decision theories and in choice models as regret anticipation, which is defined from the choiceless utilities of the alternatives faced by individuals and from the comparisons made between chosen and non-chosen alternatives, such that, if the non-chosen alternative happens to be better, or if the individual imagines that they could have been better, the consumer feels regret and if the chosen alternative happens to be better she feels rejoice (Loomes & Sugden, 1982; Pieters & Zeelenberg, 2007; Zeelenberg, 1999; Zeelenberg & Pieters, 2007).

Finally, a third behavioral rule is a dual stage choice process; initially, the consumer simplifies the choice task and then she chooses an alternative. Dual stage choice process is a common description of consumer decision-making, both in behavioral decision theories (Chakravarti & Janiszewski, 2003; Hauser, 2014; Kardes et al., 1993) as well in discrete choice modeling (Adamowicz & Swait, 2012; Manski, 1977; Swait, 2001; Swait & Ben-Akiva, 1987). In the proposed dual stage choice process, the first step aims to reduce the alternatives in the choice scenario to form a subjective choice set that minimizes regret (Mussel, Göritz, & Hewig, 2014; Pieters & Zeelenberg, 2007), in which expected utility is maximized. This concept is consistent and inspired by the multiples goals choice based model proposed in Marley & Swait (2017)

Against this background, this section will be developed as follows: after this introduction, I will present a sub-section (second) exploring the literature to pursue a more precise definition of regret and a brief review of its assimilation in the discrete choice modeling literature. In the third sub-section, I will describe the model and its theoretical foundations. In the fourth sub-section, I will detail the design and the results of the Monte Carlo experiment executed to test the aptitude of the model to recover the proposed behaviors. Finally, in the fifth sub-section, I will discuss the results.

5.2 Literature review

Counterfactual thinking, as a pervasive cognitive activity in human life, is defined as the reasoning of what could have been and it involves fact-based comparisons, like social or past-temporal, or simulated based comparison, as in future temporal and counterfactual; this comparative thinking supports self-improvement and the affect-regulation function, with the focus on an alternative better than the current one motivating goal pursuit (Summerville & Roese, 2008). This comparative thinking influences decision-making focusing the individual in the appropriate subset of the existing information, triggering information about a judgmental standard and using this knowledge as a representation of target knowledge that is absent (Mussel et al., 2014). All in all, counterfactual thinking is an essential skill for the individual to "forge connections to create coherent solutions to problems" (Kray et al., 2010 p. 107).

Regret is one negative emotions triggered by the counterfactual thinking (fact or simulated based) that the situation would have been better had a different choice been made, i.e., its emergence requires the occurrence of a cognitive appraisal and a negative feeling. Additionally, to be discriminated from other negative emotions the individual have to accept self-responsibility for a bad or unjustifiable decision (Buchanan, Summerville, Lehmann, & Reb, 2016; Pieters & Zeelenberg, 2007; Zeelenberg, 1999).

The emotion has different roles depending on the domain. Specifically, it can arise from immutable outcomes of past decisions, and it will have a learning role to avoid a mistaken repetition, or it can arise while feasible corrective behaviors may serve to goal attainability. How regret intensity varies, as a function of this situational characteristics, is still unresolved (see Beike, Markman, & Karadogan, 2009 and Summerville, 2011 for different perspectives). It can also have effects on post-choice consumer behaviors (Bui, Krishen, & Bates, 2011; Shani & Zeelenberg, 2007), but at this moment I am interested in the regret elicited, or anticipated, during the choice process, i.e. while it is functional for goal pursuing.

Regret is an important concept to goal based choice, since it works as a feeling for doing emotion, signaling favorably or unfavorably to choice alternatives, enduring while opportunities to goal pursuit are open and inducing the engagement in improving behaviors to avoid the unpleasant emotion of a likely bad choice (Beike et al., 2009; Summerville, 2011;

Zeelenberg & Pieters, 2007). In the consumer choice domains regret is relevant and it is anticipated when the consumer faces complex or relevant choices, when she expects to quickly learn the outcomes of the chosen and nonchosen alternatives, when the choice involves preference uncertainty, and it can be felt regarding the choice process or the choice outcome. (Inman, 2007; Zeelenberg & Pieters, 2007).

To conclude, regret minimization is consistent with the multiple goal framework proposed in section I of this thesis, and since its regulatory strategies connect the decision situation, the multiple goals space, and the multi-attribute space. A few of these strategies are goal level adaptation; to increase decision quality, justifiability or to avoid decision; to restrict or enlarge the choice set, to guarantee choice reversibility or to avoid feedback about the nonchosen alternatives; and to anticipate regret. (Pieters & Zeelenberg, 2007).

5.2.1 Anticipatory regret in econometric choice models

One of the pioneering applications, if not the first, of regret theory based choice models to the marketing domain, studied the effects of coupon expiration date on consumer behavior using revealed preference data (Inman & McAlister, 1994). The dominant perspective was that redemption was an exponentially decreasing function in time past from the producer dropping the coupon. However, to reduce liabilities, producers changed this kind of marketing initiative model and started issuing coupons with expiration dates. The authors defended that the response was also a change in consumer behavior and developed an econometric model allowing for a bimodal temporal distribution of coupon redemption. The initial activity, right after the dropping of the coupon, is still exponentially decreasing in time, however as the expiration becomes proximal the model predicted that redemption would peak again, i.e., it would be exponentially increasing as the expiration date becomes closer. The concept of the model is that consumers anticipate regret and act, prior to expiration date, to avoid giving away the coupon's benefit. The regret model outperformed the traditional econometric models, based on utility maximization, both in fit and in out of sample prediction.

It took more than one decade until regret theory started to be incorporated in discrete choice models, to evolve from a first specific formulation and to spread across disciplines like transportation, applied economics, and marketing. In the original formulation, the model compared each foregone alternative in a choice task with the chosen one and added to the regret function the difference in every attribute in which the chosen alternative performed worse than the unchosen ones, allowing for a semi-compensatory choice process (Chorus, Arentze, & Timmermans, 2008).

Two limitations of this model were overcome in a new formulation, first the fact that only the comparison between the chosen and every nonchosen was being considered, and second, the function discontinuity caused problems to derive the policy measures. The novel formulation, known as generalized random regret model (Chorus, 2010), is described by the following choice evaluative function, after integrating out the errors components (the derivation of the model is presented in the original paper):

(5-1)

$$\widetilde{R_i} = \sum_{j \neq i} \sum_{k=1,\ldots,K} ln\left(1 + exp\left(\beta_k\left[x_{jk} - x_{ik}\right]\right)\right)$$

Where every alternative j is compared to the focal alternatives i in every attribute from $1,\ldots,K$ and the difference is weighted by a regret parameter. The logarithm of the sum of these binary regrets smooths the function allowing for differentiation in its whole domain. Additionally, when incorporating to the regret measure the attributes in which the focal alternative performs worse and those in which it performs better compared to the competitor, the model reconciles itself with the earlier theoretical propositions, accounting for regret and rejoice. And given the form of the regret function, the measure is more sensitive to the attributes in which the focal firm perform worse than to the ones in which it performs better, accounting for loss aversion (Kahneman, 2003; Kahneman & Tversky, 1979; Tversky & Kahneman, 1981). Given that the individual goal is to minimize random regret, which is equivalent to maximize its negative, for the estimation the choice model is given by:

(5-2)

$$P_i = \frac{e^{-\widetilde{R_i}}}{\sum_{j \in J} e^{-\widetilde{R_j}}}$$

5.3 The econometric model

The presentation of the econometric model will rest on some definitions introduced in section 3 of this thesis, and they will be detailed and formalized to allow the model derivation supported by Marley & Swait, (2017). I will avoid repetitive citations, but the reader can find further details and identify some of the ideas that I will develop in the original paper. Besides this main reference, I will cite the other specific ones as eventually needed.

This is a multiple meta-goals choice based model, and meta-goals are knowledge structures, describing desired end states related to the choice process. These structures are associated with other concepts in memory, meaning that once a meta-goal is activated the related concepts in memory, including the means for goal striving, will become salient (van Osselaer et al., 2005). So, I start defining the meta-goals and the goal choice strategy.

5.3.1 Meta-goals and goal choice strategy

In Figure 5, section 3, the choice process is driven by the multiple goals space and it initiates with the goal choice strategy informing the activated meta-goals and their achievement criteria. Let $\mathcal{G} = \{accuracy, regret\}$ be the set of currently activated meta-goals, and let $\Gamma(\mathcal{G})$ be the set of nonempty subsets of $\mathcal{G}$, such that $\mathcal{A} \in \Gamma(\mathcal{G})$ is an unobserved meta-goal set. Notice that $|\Gamma(\mathcal{G})| = 2^{\mathcal{G}} - 1 = 3$, i.e., to attain accuracy only, regret only or both.

To describe the achievement criteria, I will define the multi-attribute space as the usual discrete choice model having $\mathcal{M}$ as a finite set of discrete alternatives available to the consumer, i.e., the choice scenario or choice task. Here $\Gamma(\mathcal{M})$ is the set of nonempty subsets of $\mathcal{M}$, and $\mathcal{C} \in \Gamma(\mathcal{M})$ is the unobserved choice set driving the consumer choice.

Each alternative $j \in \mathcal{M}$ is described by a vector x describing its attributes, such that, $x = \{1, \dots, \mathcal{K}\}$. Now, each j $\in \mathcal{M} = \{x_{j1}, \dots, x_{jjk}\}$, i.e., every alternative is a set of attributes measures, each one, as nominal, ordinal discrete, or continuous. To complete, every attribute k has one own vector of attribute levels, such that, $l_k = \{1, \dots, \mathcal{L}\}$. Attributes can have equal or different cardinalities and the levels are coded to allow the analyst to measure the desired statistical effects.

Now, let a goal attainment program to describe the achievement of every activated $g \in \mathcal{A}$, i.e., the goal attainment programs are the means associated with the activated goals in memory, and once the goals are triggered, the means become salient. The program associated with accuracy is to maximize expected utility, which for any j $\in \mathcal{M}$ is:

$$U_{nj} = \sum_{K} \beta_k x_{njk} + \varepsilon_{nj} \tag{5-3}$$

Where U_{nj} is the utility of j to consumer n, β_k is the taste parameter for an attribute, x_{njk} is the level of attribute k seen by consumer n in alternative j, and ε_{nj} is the stochastic term arising from the pursuit of conflicting objectives (Swait & Marley, 2013). Notice that in this formulation the tastes are homogeneous, and the utility varies across consumers because they can see different attribute levels in the choice task. The model is easily expanded to accommodate taste heterogeneity, but given that this is a regular feature of usual discrete choice models I am simplifying the simulation to emphasize the novel elements to be proposed. Finally, the maximization of the utility completes the goal attainment program for accuracy, formally defined as:

$$\max_{j \in \mathcal{C}} U_j \tag{5-4}$$

I will anticipate the formula describing the choice probability of every alternative as the multinomial logit model (McFadden, 1974), since it will be needed ahead to integrate the goals:

$$P_i^{accuracy} = \frac{exp(U_i)}{\sum_{j \in \mathcal{C}} exp(U_j)} \tag{5-5}$$

Notice that to choose $\max_{j \in \mathcal{C}} U_j$ or the alternative with the highest probability results in the same outcome.

Likewise, the goal attainment program for regret is the generalized random regret given in equation (5-1), although the model can well assimilate other formulations. Notice that the stochastic term is already integrated out from $\widetilde{R}_j$, but stochasticity is implicit in the model and received the same interpretation as in the expected utility. Thus, the goal attainment program for regret corresponds to minimize random regret, or:

$$\min_{j \in \mathcal{C}} \widetilde{R}_j \tag{5-6}$$

The next equation repeats, from the generalized random regret model, the choice probabilities under the regret goal attainment program:

$$P_i^{regret} = \frac{e^{-\widetilde{R}_i}}{\sum_{j \in J} e^{-\widetilde{R}_j}} \tag{5-7}$$

Now I will develop the goal attainment program for $\mathcal{A} = \{accuracy \wedge regret\}$, which demands a behavioral rule that strives to the simultaneous achievement of the two meta-goals. Given that the goals may conflict, i.e., in any choice task, the chosen alternative that maximizes expected utility is not necessarily the one that minimizes random regret, a goal attainment strategy is a procedure that reconciles the conflicting goals. The proposed strategy involves a two-stage choice process, the first reducing the number alternatives in the choice task and the second evaluating the remaining alternatives and choosing among them. Anticipatory regret is an emotion resulting from counterfactual thinking, emerging when choice task cause conflict, due to much information in the choice scenario or to attributes negatively correlated or when preferences are uncertain (Inman, 2007; Zeelenberg & Pieters, 2007). Among the regulatory strategies, to increase efforts searching for a better outcome, to focus comparisons on relevant alternatives or to use the painful feeling to learn about preferences are identified in the literature (Mussel et al., 2014; Pieters & Zeelenberg, 2007). Thus, the goal attainment strategy is to reduce the original alternatives to a choice set that minimizes regret, followed by expected utility maximization in the second stage.

It is necessary to formalize these ideas as the goal attainment strategy in the choice model, and I start with the choice set formation. Let $\mathcal{A} \in \Gamma(\mathcal{G})$ be the set of meta-goals activated in the first stage, regret in this case, and let $\Phi_{\mathcal{M}}(\mathcal{A})$ be a goal attainment strategy, that executes a goal attainment program ensuing in the choice set $\mathcal{C} \in \Gamma(M)$. Finally, $a_{\mathcal{M}}(\mathcal{A})$ is the probability of the set $\mathcal{A} \in \Gamma(\mathcal{G})$ being activated and $q_{\Gamma(\mathcal{M})}(\mathcal{C} \mid \Phi_{\mathcal{M}}(\mathcal{A}))$ is the probability of the choice set $\mathcal{C} \in \Gamma(M)$ being chosen when goal strategy $\Phi_{\mathcal{M}}(\mathcal{A})$ is executed.

Once a choice set minimizing random regret endures the first stage, let $\mathcal{B} \in \Gamma(\mathcal{G})$ be the set of meta-goals activated in the second stage, accuracy in this case, and let $\Psi_{\mathcal{C}} = (\mathcal{B})$ be the goal attainment strategy implementing a goal attainment program resulting in choice. Here $b_{\mathcal{C}}(\mathcal{B})$ is the probability of the goal set $\mathcal{B} \in \Gamma(\mathcal{G})$ being started, and $r_{\mathcal{C}}\big(j \mid \Psi_{\mathcal{C}} = (\mathcal{B})\big)$ is the probability of alternative $j \in \mathcal{C}$ being chosen, given $\Psi_{\mathcal{C}} = (\mathcal{B})$, or utility maximization in this case.

I am ready to derive the choice model for this behavioral rule, and following Marley & Swait (2017) the overall probability for the alternative i is:

(5-8)

$$P_{\Gamma(\mathcal{G})\mathcal{M}}(i) = \sum_{\mathcal{C}\in\Gamma(\mathcal{M})} \left(\sum_{\mathcal{A}\in\Gamma(\mathcal{G})} a_{\mathcal{M}}(\mathcal{A}) q_{\Gamma(\mathcal{M})}\big(\mathcal{C} \mid \Phi_{\mathcal{M}}(\mathcal{A})\big) \right) \left(\sum_{\mathcal{B}\in\Gamma(\mathcal{M})} b_{\mathcal{C}}(\mathcal{B}) r_{\mathcal{C}}\big(j \mid \Psi_{\mathcal{C}} = (\mathcal{B})\big) \right)$$

To streamline the notation, from equation (5-8) we define the probability of a choice set in stage one:

(5-9)

$$Q_{\Gamma(\mathcal{G})\Gamma(\mathcal{M})}(\mathcal{C}) = \left(\sum_{\mathcal{A}\in\Gamma(\mathcal{G})} a_{\mathcal{M}}(\mathcal{A}) q_{\Gamma(\mathcal{M})}\big(\mathcal{C} \mid \Phi_{\mathcal{M}}(\mathcal{A})\big) \right)$$

The equation (5-9) needs a specific formulation to explicitly simplify the choice scenario and this is given by the independent availability logit model - IAL - (Swait & Ben-Akiva, 1987). The full model is derived in the original paper, but the important message is that in the first stage there is a threshold τ that drives the availability of every alternative in the subjective choice set, and the probability of occurrence of each choice set depends on the probability of each expected alternative in the choice task to overcome the threshold and also on the probability of the unexpected alternatives in the choice task to not overcome the threshold.

And, now, I define the probability of any alternative j_i given a choice set:

(5-10)

$$R_{\Gamma(\mathcal{G})\mathcal{C}}(i) = \left(\sum_{\mathcal{B}\in\Gamma(\mathcal{M})} b_{\mathcal{C}}(\mathcal{B}) r_{\mathcal{C}}\big(j \mid \Psi_{\mathcal{C}} = (\mathcal{B})\big) \right)$$

Finnaly, equation (5-8) is expressed as:

(5-11)

$$P_{\Gamma(\mathcal{G})\mathcal{M}}(i) = \sum_{\mathcal{C}\in\Gamma(\mathcal{M})} Q_{\Gamma(\mathcal{G})\Gamma(\mathcal{M})}(\mathcal{C})\, \mathrm{R}_{\Gamma(\mathcal{G})\mathcal{C}}(\mathrm{j})$$

And the goal attainment program for this two stage process is:

(5-12)

$$\max_{j\in\mathcal{C}} P_{\Gamma(\mathcal{G})\mathcal{M}}(j)$$

Now it is time to develop the choice model, that is specified in the next sub-section

5.3.2 The Multiple Meta-goal Based Choice Model

To develop the choice model, it is necessary to integrate the goal choice strategy, independently specified for the two meta-goals and for goal attainment strategies and programs, into a common framework. The objective is an econometric model that identifies a mixture of meta-goals and two archetypes behaviorally different in the use of these mixtures. The consistent archetype uses a constant mixture of the meta-goals and the adaptive archetype varies the mixture across occasions. It means that the consumer chooses an initial mixture and then chooses to keep or to adapt it in a different choice occasion. This is a similar process of the anchoring and adjustment described in Li (2013) and in Swait et al (2016), but those authors adopted a pure strategist as the consistent archetype. The first author used a traditional latent class model (Kamakura & Russell, 1993) to identify the consistent archetype pursuing the functional goals. The last authors used the same approach to identify the consistent full information processor in their information archetype mixture model. In practice, it means that once the consumer chooses the archetype, she will also pursue one goal or be full information processor, and heterogeneity in goal pursuit or information usage arises among adaptive consumers. The consistent archetype in the model being proposed weights meta-goals desirability at each choice, but the weights are constant across occasions and there is heterogeneity in the weights selected by different consumers using a mixed strategy. There may be consumers that put all the weight in one meta-goal, and this individual becomes a pure strategist. In this sense, the consistent pure strategist is a special case of the consistent mixed strategist.

The first step to define any mixture is to combine different goal attainment programs, and I start following the lead from Dawes (1964, p.108):

> Combining procedures may be viewed mathematically as follows. Each separate procedure maps the vector representing an individual into a point on a line—the decision line. If more than one procedure is employed, each vector is mapped into a number of such points. These points, in turn, define the coordinates of a new vector, and the method of combining selection procedures maps this new vector into a final decision line. If selection is based on a combination of procedures, the order in which they are considered does not affect the outcome -- unless, of course, the individuals change during the selection procedure itself.

Given that we observe the choice scenarios, the individuals' characteristics and the outcomes; and that meta-goals, behavioral rules and choice sets are unobserved, the inference of latent variables must be supported by the observed ones. To identify the initial mixture of meta-goals that combines the outcomes of the different behavioral rules, a scalar optimization (Marley & Swait, 2017; Swait & Marley, 2013) leads to the following solution:

$$\mathcal{W}_{n\mathcal{g}}^{\mathbb{C}} = \frac{exp(\alpha + \zeta Z_n)}{\sum_{\mathcal{G}} exp(\alpha + \zeta Z_n)} \tag{5-13}$$

Where $\mathcal{W}_{\mathrm{n}\mathcal{g}}^{\mathbb{C}}$ is the weight the consistent consumer n assigns to goal $\mathcal{g}$, α is a constant capturing the average weight attributed to goal $\mathcal{g}$ in the population while $\mathrm{Z_n}$ is a vector of variables describing the individual or the situation in which the choice is performed. For instance, the necessity to justify the choice to others is a task characteristic likely to elicit anticipatory regret, and, if present, may increase the weight of the meta-goal regret. Likewise, individual characteristics are known to be related to meta-goals, e.g., consumers identified as maximizers in the maximize-satisfice personality scale (Dar-Nimrod et al., 2009) are more prone to activate the accuracy goal.

Let, from equations (5-5), (5-7) and (5-11), $P_j^{\mathcal{G}}$ be the matrix with choice probabilities under every $\mathcal{g}$ and $\mathcal{W}$ be a vector line with weights for every $\mathcal{g}$, thus the final choice probabilities for the consistent archetype are given by:

$$\boldsymbol{P} = P_j^{\mathcal{G}} . \mathcal{W}^{\mathbb{C}\prime} \tag{5-14}$$

And to complete the goal choice strategy for the consistent mixture archetype, execute:

(5-15)

$$\max_{j \in \mathcal{C}} P_j^{\mathcal{G}} . \mathcal{W}^{\mathbb{c}\prime}$$

At this point, the adaptive mixture archetype needs support to adjust her behavior to the context, and this is done providing appropriate information for meta-goal weighting adjustment, such that:

(5-16)

$$\mathcal{W}_{n\mathscr{g}}^{\mathbb{a}} = \frac{exp(\alpha + \zeta Z_n + \gamma M_t)}{\sum_{\mathcal{G}} exp(\alpha + \zeta Z_n + \gamma M_t)}$$

In equation (5-16) $\mathrm{M_t}$ is a variable, or a vector of variables, describing the decision environment in occasion t. It can be a context variable, like the inclusive value (Ben-Akiva & Lerman, 1985) that represents the expectancy of the attainment of each meta-goal in the choice set. A measure, like the inclusive value, offers support for weight adaptation as a function of the context, since the attainability of the meta-goals may be affected by the multi-attribute space. It can also be a situational characteristic that varies across occasions like fatigue or learning stage. The fact that $\mathrm{M_t}$ adds variability to the choice occasion offer the support for the adaptive mixture archetype and completes the information for implementing the last piece of the goal choice strategy for the adaptive mixture archetype. The choice probabilities are:

(5-17)

$$\boldsymbol{P} = P_j^{\mathcal{G}} . \mathcal{W}^{\mathbb{a}\prime}$$

And the goal choice strategy for the adaptive mixture archetype is completed executing the command:

(5-18)

$$\max_{j \in \mathcal{C}} P_j^{\mathcal{G}} . \mathcal{W}^{\mathbb{a}\prime}$$

To conclude the model, it is necessary to bring the archetypes together, and this is also done through a scalar optimization, but now operating over the likelihood of the choice sequence from a consumer. The process begins analogous to the studies using the archetypes (Li, 2013; Swait et al., 2016), such that the probabilities of a consumer to be consistent or adaptive are given by the logistic model:

(5-19)

$$Q^{consistent} = \left(1 + exp(\delta)\right)^{-1}$$

And the probability of the consumer to be adaptive is:

(5-20)

$$Q^{adaptive} = 1 - Q^{consistent}$$

Once again, the archetype is unobserved, and some information is required to estimate δ, that should be enduring information about the individual or the choice environment. For instance, need for cognition (NFC) is a personality trait (Cacioppo & Petty, 1982; Cacioppo et al., 1984) that is known for influencing the choice process. One hypothesis to be explored is whether the probability of being an adaptive mixture of meta-goal is increasing in the scale measuring (NFC), since as the score increase so does the pleasure of the individual to engage in cognitively demanding tasks. If the adaptive archetype needs to explore the variations in choice occasions, it is possible for this archetype to being attractive to higher NFC consumers.

Now, it is possible to enable this consumer with a complete choice model, and the unconditional probability of observing a sequence $T_n$for consumer *n*, is:

(5-21)

$$P_n(T_n) = P_n(T_n \mid Consistent) * Q_n^{consistent} + P_n(T_n \mid Adaptive) * Q_n^{adaptive}$$

And the last step is to specify the likelihood function for this model, which is the product of the unconditional choice probabilities of the chosen alternatives, over N consumers:

(5-22)

$$L = \prod_n P_n(T_n)$$

$$= \prod_n \left(P_n(T_n \mid Consistent) * Q_n^{consistent} + P_n(T_n \mid Adaptive) * Q_n^{adaptive}\right)$$

As already mentioned, there is a difference in the consistent mixture archetype proposed here, and the non-adaptive in previous studies and the modification emerges from the likelihood function. In the previous models, the consistent archetype is a traditional latent class, implying that the mixture between the archetypes happens at end of the choice sequence. The consistent archetype is a pure strategist and the adaptive is responsive to the environment. Now, the consistent archetype is a mixed strategist, who still does not respond to changes in the environment but brings more *ex-ante* flexibility.

Figure 19 is the visual summary of the model formalized throughout this section. It includes the adaptive mechanism operating at the archetypical level, the multiple meta-goals pursued by the consumer at the decision rules heterogeneity level and the goal evaluation strategy that maps alternatives into the goal space through the preference components.

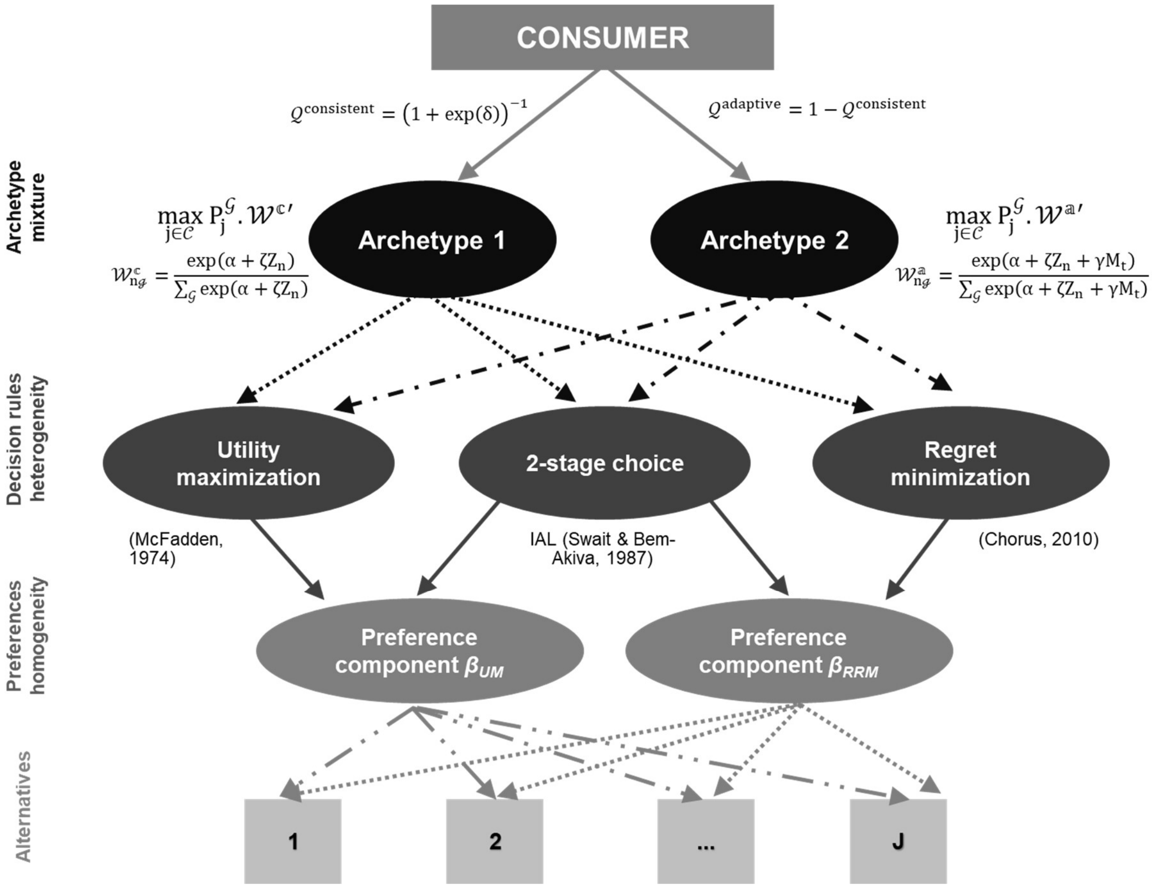

Figure 19 - Multiple meta-goals based choice model

To test the ability of this model to recover the behaviors described, if they happen to be present in the data generation process, I have conducted Monte Carlo experiment that is described in the next sub-section.

5.4 Monte Carlo experiment to simulate Multiple Meta-goals Choice Based Model

The true data generation process originated the data described by the behavioral model presented in the previous sub-section, and I, now, use the same notation to detail how data and parameters were created for this simulation.

The dataset was composed by 1.000 respondents, each one responding to eight choice tasks, i.e., $t = \{1, \dots, 8\}$; each choice task had four or eight alternatives. For each respondent half of the tasks was composed with four alternatives and the other half with eight alternatives, so $|\mathcal{M}| = 4\ or\ |\mathcal{M}| = 8$. The position of tasks as a function of its cardinality, in the individual sequence, was random. Thus, 8.000 thousand choice scenarios were randomly generated, without any special concern about design, except to not have two identical alternatives in the same choice task. Each alternative, in every choice task, is described by four nominal generic attributes, or $|x| = 4$, and all the attribute have five levels, $\mathcal{L} = 5$ with the specific values drawn from a uniform distribution. For modeling purposes, the attributes were effect coded with the last level of each omitted for identification. Hence, given the alternatives configuration there were sixteen parameters to be estimated for the utility function, and other sixteen for the regret function. The true utility and regret parameters are detailed the first line of Table 32. Notice that this setting implied in preference homogeneity. This data set supports all the replications in the Monte Carlo simulation.

After generating the data set, parameters and data defining mixtures for archetypes and meta-goals were generated, as follows.

The archetypes mixtures were fixed across replications, and the true proportions were 50% consistent archetype and 50% adaptive archetype. This proportion is chosen to create the situation of maximum entropy and, therefore, the most challenging to the model. The attribution of every respondent to each archetype was based on a constant-only model implemented through a random variable $\phi_n \sim U(0,1)_N$, such that if $\emptyset_n \leq 0.5$ consumer n is consistent, otherwise she is adaptive. This definition was held constant across replications and the actual draw resulted in $\overline{\emptyset} \cong 0.497$, which is equivalent to a true $\delta \cong 0.012$ under the logistic model expressed in equation (5-19).

After assigning every respondent to one of the available archetypes, the models to create the proper $\mathcal{W}^{\mathbb{c}}$ or $\mathcal{W}^{\mathbb{a}}$ were created. These models are drawn at every replication of the experiment.

Starting with the consistent archetype, equation (5-13) implies the availability of α, ζ and Z_n to estimate $\mathcal{W}^{\mathbb{c}}$. Given the three meta-goals described in section 5.3.1 the two degrees of freedom were used to draw the model parameters as $\alpha \sim U(-1{,}1)_{|\mathcal{A}-1|}$ and $\zeta \sim U(-1{,}1)_{|\mathcal{A}-1|}$. Now, the estimation of ζ demanded support of exogenous data generated as $Z_n \sim U(-1{,}1)_{|N,1|}$. Notice that beyond the substantive interpretation discussed in section 5.3.2, statically speaking α informed the mean meta-goals weights for the sample and $\zeta . Z_n$ added between-subjects behavioral heterogeneity.

To estimate $\mathcal{W}^{\mathbb{a}}$ for the adaptive archetype, the support of $\gamma \mathrm{M_t}$ was required. The parameter is a random variable drawn as $\gamma \sim U(-1{,}1)_{|\mathcal{A}-1|}$ and the exogenous data needed for its estimation was given by $\mathcal{M}_t \sim U(-1{,}1)_{|N*T,1|}$. Besides the substantive interpretation of this term, it added within-subject behavioral heterogeneity.

Specifically for the two-stage goal attainment program, the independent availability model used to operate the choice set formation in the first stage required the definition of a threshold, specified over the regret function. Notice in equation (5-1) that random regret is increasing in the number of alternatives in the choice scenario and, as defined in the data generation process, the number of alternatives per choice task varied in the data set. Thus this definition needed to account for this variation. With this motivation the threshold was:

(5-23)

$$\tau_{nt} = 3 * \frac{\sum_{j \in \mathcal{M}_{nt}} \tilde{R}_{jnt}}{|\mathcal{M}_{nt}|}$$

In words, equation (5-23) means that the threshold was defined by choice scenario and it was three times the average random regret of the alternatives in the choice task, i.e., if a given alternative generates a random regret that was three times the average of the regret generated by the sample, it will not be available for the second stage of the choice process.

At this point, the only missing information to complete the data generation process was the vector of observed true choices, given the previous assignment of consumers to archetypes. For every choice task of each respondent the matrix $P_j^{\mathcal{G}}$, with the choice probabilities, was computed using equations (5-5), (5-7) and (5-11) and multiplied by the weighting vector, given the individual's archetype, to obtain the final vector of choice probabilities $\boldsymbol{P}$ defined in equations (5-14) or (5-17).The final choice in each choice set resulted from a random draw with the probabilities respecting $\boldsymbol{P}$. To realize the choice, for every choice task a random variable $i \sim U(0,1)_{|J,1|}$ was generated and the chosen alternative was the one resulting from the indicator function $\mathrm{I}[\min CDF(\boldsymbol{P_{nt}}) \geq i_{nt}]$, as it was described in Guevara, Chorus, & Ben-Akiva (2016)

Given the model definitions and the specified data sets, I have written a code for the software Gauss 16.0 to recover the vector of parameters $\theta(\beta_{UM}, \beta_{RR}, \delta, \alpha, \zeta, \gamma)$ using maximum likelihood estimation.

The results are presented in Table 32 and the expected *z* statistic to reject differences between true and estimated parameters is less than 1.96, taking 95% as the confidence level. Although there were still statistically significant differences between the true and the estimated parameters, the absolute and the relative differences were quite small. There were five βs used in the utility function with *z* statistics larger the 1.96, however, the largest absolute difference was 0.012 meaning that the estimated parameter, in this case, was 1.2% larger that the true one. There were also three βs used in the random regret function with *z* statistics larger than 1.96, but the largest absolute difference between the true parameter and the estimated one was 0.003. The *z* statistic was also larger than 1.96 for δ and the absolute difference 0.18, which means that under the logistic model the estimated probability for the consistent archetype was around 47.8% against the 49.7% in the true model. Lastly, the *z*-statistic of α_2 was also larger than 1.96 but the absolute difference was only 0.006.

All in all, the model is working well in recovering preferences and mixture parameters and although there were some undesirably large *z*-statistics, the absolute difference were very small and it seems that with enough additional replications the estimated parameters tend to asymptotically converge to the true ones.

Table 32 - Results for Monte Carlo experiment

	UM_1	UM_2	UM_3	UM_4	UM_5	UM_6	UM_7	UM_8	UM_9	UM_{10}	UM_{11}	UM_{12}	UM_{13}	UM_{14}	UM_{15}	UM_{16}
θ	1.00	0.50	0.00	-0.50	0.00	-0.50	-1.00	-0.50	1.00	0.50	0.00	-0.50	-1.00	-0.50	0.00	0.50
$\hat{\theta}$	1.00	0.50	0.00	-0.50	0.00	-0.50	-1.01	-0.50	1.01	0.50	0.00	-0.51	-1.01	-0.50	0.00	0.50
df	798	798	798	798	798	798	798	798	798	798	798	798	798	798	798	798
z	0.68	1.55	1.16	0.03	0.32	2.36	3.27	1.43	5.29	1.84	0.01	3.42	3.45	1.91	0.13	1.66
dif	0.001	0.003	0.002	0.000	0.001	-0.005	-0.008	-0.003	0.012	0.004	0.000	-0.007	-0.008	-0.004	0.000	0.003

	RR_1	RR_2	RR_3	RR_4	RR_5	RR_6	RR_7	RR_8	RR_9	RR_{10}	RR_{11}	RR_{12}	RR_{13}	RR_{14}	RR_{15}	RR_{16}
θ	0.50	0.00	-0.50	-1.00	-0.50	0.00	0.50	1.00	-1.00	-0.50	0.00	0.50	0.00	0.50	1.00	0.50
$\hat{\theta}$	0.50	0.00	-0.50	-1.00	-0.50	0.00	0.50	1.00	-1.00	-0.50	0.00	0.50	0.00	0.50	1.00	0.50
df	798	798	798	798	798	798	798	798	798	798	798	798	798	798	798	798
z	2.44	0.96	0.41	0.97	1.43	2.00	1.18	0.67	1.01	0.19	0.53	1.28	0.54	0.25	2.09	0.02
dif	0.003	-0.001	-0.001	-0.002	-0.002	0.003	-0.002	0.001	-0.001	0.000	-0.001	0.002	0.001	0.000	0.003	0.000

	δ	$\alpha 1$	$\alpha 2$	$\zeta 1$	$\zeta 2$	$\gamma 1$	$\gamma 2$
θ	0.01	-0.01	0.04	0.01	0.03	0.01	0.00
$\hat{\theta}$	0.19	-0.01	0.05	0.01	0.03	0.01	0.01
dif	798	798	798	798	798	798	798
z	5.15	0.99	3.48	1.20	0.17	0.04	0.97
dif	0.180	-0.002	0.006	-0.003	0.000	0.000	0.009

df = degrees of freedom or the number of replications (R) minus 1.

$$z = abs\left(\frac{\theta_R - \hat{\theta}_R}{\sqrt{\frac{s_r}{df} - \frac{\hat{s}_r}{df}}}\right)$$

$$dif = mean(\theta_R - \hat{\theta}_R)$$

5.5 Final considerations

I have presented a flexible meta-goal based choice process, based on three components accounting for the multiple-goal pursuit and the possibility of behavioral adaptation to the context or to the task characteristics. The model was inspired by some recent papers using the concept of mixtures, archetypes, and alternatives to the full information utility maximization paradigm.

The model innovated by the assimilation of different behavioral rules that were used as mixed strategies, through consistent mixtures, or adaptive strategies, through adaptive mixtures.

The conceptual foundation of the model was well grounded in the behavioral decision theory field, that emphasizes multiple-goal pursuit as the psychological concept bringing rationality to the consumer choice process as an alternative to the traditional *homo economicus* paradigm. It is also well grounded in discrete choice modeling literature, since that the development of models accounting for heterogeneity in the choice process, beyond tastes, are receiving a lot of attention from this community. A novel approach to explain stochasticity in the choice process is to explain it as an outcome of multiple goals pursuit, coupling the choice modeling stream with the behavioral view of decision-making. The model presented adopts this perspective of multiple goals choice based process and introduces heterogeneity at the behavioral rule level as the mean to achieve the meta-goal pursuit. This issue has been studied for a long time in behavioral decision theories and there are choice models already exploring heterogeneity at this level. But, to the best of my knowledge, this is the first to introduce the decision strategy heterogeneity as a mixed strategy, and to allow adaptation within-individual, across choice occasion.

These ideas are extremely aligned with the concept of the *homo aptabilis* that served as the north for developing this thesis.

There is still a long way to validate this meta-goal choice based model, and the main limitation of this section is that it still misses an empirical application. This is the only way to know if the behavioral propositions find support in reality and if the model is able to recover.

6 FINAL CONSIDERATIONS

The objective of this thesis was to explore consumer response heterogeneity, specifically in the choice process. I have explored the literature in economics, behavioral decision theories, and choice modeling.

In the first section, I have organized the knowledge based on how BDT see the choice process, reminding that this school of thought has been advocating heterogeneity in the choice process for a long time (Bettman et al., 1998; Payne, 1982). From the idea of an adaptive choice process, I have reviewed the concepts driving human decision-making and identified the parallel concepts and models already developed in the choice modeling literature.

The main idea is that choice is an outcome of the interaction among context, task and individuals' characteristics. This interaction is what challenges the normative rationality implied by the *homo economicus*, preventing the behavioral invariance implicit in the expected utility theory. Thus, characteristics of the choice set, like the amount and the structure of information available to the consumer, as well as tasks characteristics, like time pressure and necessity of choice justification, trigger different cognitive processes that lead to different outcomes when compared to the expect utility maximization framework. In the extreme opposition to the normative rationality is the concept of constructive choice process, proposing that preferences fully emerge during decision-making, leaving a marginal role to preferences as a driver of the choice process. However, task characteristics, like consumer involvement or accumulated experience, and individual characteristics, like personality traits, give room to preferences playing varying roles in the choice process.

The variability in the cognitive processes involved in human decision-making rests in the limits of the storage and processing power of the human brain. The adaptive responses are: actively managing the amount of information to consider in the choice process; relying on perceptual judgments; and using heuristics that produces outcomes, which are ecologically as good as the ones produced by the normative rationality.

The psychological construct that orders the choice process is the individual's goals, defined as cognitive structures associated with other concepts in memory, supporting the adaptive process, and working as the reference to the consumer approach over the choice scenario. These goals are hierarchical structures encompassing since the goals proximal to the multi-attribute space that defines the choice task up to long-term goals related to the choice process itself or the more subjective goals shaping lifestyles and values. Multiple goals based choice brings rationality to the human decision-making process, preferences may have a role that in contingent on the whole choice setting, and constant adaptation is the skill that allows the individual to navigate the environment in search of the goals. Now, the normative rationality may give room to the procedural rationality.

Following my original objective, I have reviewed the choice modeling literature to illustrate how BDT has been used in the development of econometric models of choice, and how this is an increasing concern among these scholars. I have demonstrated the benefit of the study of choice supported by the three schools of thought that I have chosen to explore and how the integration is already happening. Goal based choice is a common denominator to overcome the limits of normative rationality, or the *homo economicus*, and goal-based choice models that bring a flexible and adaptive rationality to the choice process, with stochasticity in the choice process being explained by multiple goals pursuit. I have named the individual that behave in such a flexible way as *homo aptabilis*.

A common effect of this flexible and adaptive choice process is the simplification in the choice scenario leading to choice set formation, i.e., a process of selection or elimination of information causing the choice to be based on unobserved choice sets. The idea of multiple meta-goals-pursuit, choice set formation, and flexibility in the choice process have driven my empirical investigation.

The first empirical project has explored the risks of ignoring choice process heterogeneity, which includes channeling the phenomenon all the way down to the preferences. I have started using Monte Carlo experiments to develop demand representations for two common marketing contexts: a fast movable consumer goods and a service context. These representations rest in choice set heterogeneity and taste homogeneity, depicting a relatively simple view of the *homo apatabilis*. In the modeling side of this investigation, I relied on the (yet) dominant practice of accounting for taste heterogeneity and choice process homogeneity, i.e., the assumption of the *homo economicus* behavior. The result was that all the parameters were biased and the firm-specific constants, although being the most seriously affected variables, were not able to capture the totality of biases. The firm-specific attributes (prices in my empirical application) were also importantly affected by the model misspecification, with an interplay such that the larger the bias in firm-specific constant the smaller it was in the firm-specific attributes, and vice-versa. Additionally, since one of the firm-specific constants needed to be omitted for identification, the effect over this firm rested entirely on the firm-specific attribute. At the end, even the generic attributes ended up biased as an outcome of the wrong model. To conclude, the biases not only spread all over the parameter space, as its distribution is affected by the analyst's decisions of how to code the attributes' matrix.

The model parameters do not enter the firm decision-making process, which takes the policy measures into account. Through the Monte Carlo experiment, it was possible to demonstrate that the parameters' biases were large enough to cause severe deviation in the choice probabilities and in the choice elasticities, meaning that information used as input for strategic decisions at the firm level were also wrong.

Finally, to understand the impact of bad information entering the firm decision-making process I used a game theoretical approach to evaluate the effects of misattribution consumers' choice process heterogeneity into tastes. The result confirmed that firms were driven away from payoff maximization, with profits of the focal firm and producer surplus being significantly reduced. My expectation is that this result serves as a compelling argument to choice modelers to adopt choice process heterogeneity as a standard feature in their models.

This empirical application can be expanded in several directions at the demand and at the supply side. At the demand, it is possible to introduce preference heterogeneity in the data generation process and to study other conditions, in terms of substitutability patterns. It is also possible to study contexts with the consumers' experiences varying more intensely, which can include prices varying across occasions, for instance. It also possible to develop more advanced theoretical games, introducing the possibility of the firm adopting mixed strategies, i.e., adding uncertainty about the others players behaviors. Also, private intelligence can be included as a feature of the game, creating information asymmetry. Finally, a dynamic game may study the possibility of corrective behaviors by the firm and the time that would take to the market achieve the optimum equilibrium, if it ever happens.

After this enterprise, I have developed a multiple meta-goal based choice model, incorporating the strive for accuracy, for reducing negative emotion, and a combination of both based on theoretical knowledge. The model has also given room for adaptation across choice occasions at the individual level. It was an effortful development, assimilating choice process heterogeneity, sitting on the best knowledge offered by both schools of thought.

The model allows the pursuit of multiple meta-goals, combined through a linear optimization process. The weights of the goals may vary as a function of context, task and individual characteristics and consumer may choose to be consistent in the way they weight the meta-goals or they can adjust such weights across occasions. Moreover, the model also opens the possibility of choice set heterogeneity allowing for a two-stage choice process.

One important, and obvious, opportunity is an empirical test of this model to understand if it recovers parameters for the proposed choice processes in stated or revealed preference data. Only in the empirical application the performance may be compared to competitive models, both in internal and external validity. Moreover, the econometric model is also flexible enough to accommodate other meta-goals, like reason-based choice, cost/benefit considerations in information usage among others. All in all, it is a model with the personality of the *homo aptabilis*.

Finally, the integration between behavioral decision theories and choice modeling is still a challenge. My perspective is that one improves the weaknesses of the other. Many of BDT theories have strong internal validity but suffer from the lack of external validity. The laboratory conditions do not hold in reality, and there it is difficult to isolate the effects to validate the theories and to understand their effects in real contexts.

Discrete modeling operates closer to the "real world", sometimes in the laboratory, sometimes in reality. And there are tools to fuse the data from different sources. However, econometric models are "as if" models, meaning that they are compatible with some psychological processes but, usually, do not establish the occurrence of such processes. The choice process is unobserved, and the farther the analyst is from the observed choice, the more latent is the nature of the phenomenon.

Manipulation of context variables is a common feature in discrete choice experiments. The tool box can be enriched by manipulation of task variables and self-reported information, structured or not, that help the understanding if psychological processes assumed in the econometric formulation of the choice models are really occurring. To incorporate theories and methods to develop and test behavioral decision theories using choice models is a great opportunity already being explored.

To conclude, the accumulated behavioral knowledge, in three scientific streams explored in this thesis, works as a Pandora box that once opened cannot be ignored. Add to this knowledge the enormous technology advancement that defines our time, and the call to widespread the space of the *homo aptabilis* into choice models is roaring.

7 REFERENCES

Adamowicz, W., Bunch, D., Cameron, T. A., Dellaert, B. G. C., Hanneman, M., Keane, M., … Swait, J. (2008). Behavioral frontiers in choice modeling. *Marketing Letters*, *19*(3–4), 215–228.

Adamowicz, W., & Swait, J. (2012). Are food choices really habitual? Integrating habits, variety-seeking, and compensatory choice in a utility-maximizing framework. *American Journal of Agricultural*.

Allenby, G. M., & Rossi, P. E. (1998). Marketing models of consumer heterogeneity. *Journal of Econometrics*, *89*(1), 57–78.

Andrews, R. L., Ansari, A., & Currim, I. S. (2002). Hierarchical Bayes versus Finite Mixture Conjoint Analysis Models: A Comparison of Fit, Prediction, and Partworth Recovery. *Journal of Marketing Research Journal of Marketing Research This*, *39*(1), 87–98.

Andrews, R. L., & Currim, I. S. (2005). An experimental investigation of scanner data preparation strategies for consumer choice models. *International Journal of Research in Marketing*, *22*(3), 319–331. https://doi.org/10.1016/j.ijresmar.2004.12.002

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin*, *120*(3), 338–375.

Baron, J. (2004). Normative Models of Judgment and Decision Making. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 19–36). Oxford: Blackwell Publishing.

Baumeister, R. F., Sparks, E. A., Stillman, T. F., & Vohs, K. D. (2008). Free will in consumer behavior: Self-control, ego depletion, and choice. *Journal of Consumer Psychology*, *18*(1), 4–13.

Bavetta, S. (2004). Measuring freedom of choice: An alternative view of a recent literature. *Social Choice and Welfare*, *22*(1), 29–48.

Beike, D. R., Markman, K. D., & Karadogan, F. (2009). What We Regret Most Are Lost Opportunities: A Theory of Regret Intensity. *Personality and Social Psychology Bulletin*, *35*(3), 385–397.

Belk, R. W. (1975). Situational variables and consumer behavior. *Journal of Consumer Research*, 157–164.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. Cambridge, MA: MIT Press.

Berlin, I. (1969). *Four Essays on Liberty*. Oxford: Oxford University Press.

Bettman, J. R., Luce, M. F., & Payne, J. W. (1998). Constructive consumer choice processes. *Journal of Consumer Research*, *25*(3), 187–217.

Botti, S., & Iyengar, S. S. (2006). The dark side of choice: When choice impairs social welfare. *Journal of Public Policy & Marketing*, 24–38.

Bown, N. J., Read, D., & Summers, B. (2003). The lure of choice. *Journal of Behavioral Decision Making*, *16*(4), 297–308.

Buchanan, J., Summerville, A., Lehmann, J., & Reb, J. (2016). The Regret Elements Scale: Distinguishing the affective and cognitive components of regret. *Judment and Decision Making*, *11*(3), 275–286.

Bui, M., Krishen, A. S., & Bates, K. (2011). Modeling regret effects on consumer post-purchase decisions. *European Journal of Marketing*, *45*(7/8), 1068–1090.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*(1), 116–131.

Cacioppo, J. T., Petty, R. E., & Feng Kao, C. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, *48*(3), 306–307.

Chakravarti, A., & Janiszewski, C. (2003). The Influence of Macro-Level Motives on Consideration Set Composition in Novel Purchase Situations. *Journal of Consumer*

Research, *30*(2), 244–258.

Chartrand, T. L., Huber, J., Shiv, B., & Tanner, R. J. (2008). Nonconscious Goals and Consumer Choice. *Journal of Consumer Research*, *35*(2), 189–201.

Chernev, A. (2003). When more is less and less is more: The role of ideal point availability and assortment in consumer choice. *Journal of Consumer Research*, *30*(2), 170–183.

Chorus, C. G. (2010). A new model of random regret minimization. *Ejtir*, *2*(10), 181–196.

Chorus, C. G. (2014). A generalized random regret minimization model. *Transportation Research Part B: Methodological*, *68*, 224–238.

Chorus, C. G., Arentze, T. A., & Timmermans, H. J. P. (2008). A Random Regret-Minimization model of travel choice. *Transportation Research Part B: Methodological*, *42*(1), 1–18.

Chorus, C. G., Rose, J. M., & Hensher, D. (2013). Regret minimization or utility maximization: it depends on the attribute. *Environment and Planning*, *40*(1), 154–169. Retrieved from http://epb.sagepub.com/content/40/1/154.short

Clarke, K., & Belk, R. W. (1979). The effects of product involvement and task definition on anticipated consumer effort. *Advances in Consumer Research*, *6*(1).

Cohen, A. R., Stotland, E., & Wolfe, D. M. (1955). An experimental investigation of need for cognition. *The Journal of Abnormal and Social Psychology*, *51*(2), 291.

Dana, J., & Davis-Stober, C. P. (2016). Rational Foundations of Fast and Frugal Heuristics: The Ecological Rationality of Strategy Selection via Improper Linear Models. *Minds and Machines*, *26*(1–2), 61–86.

Dar-Nimrod, I., Rawn, C. D., Lehman, D. R., & Schwartz, B. (2009). The maximization paradox: The costs of seeking alternatives. *Personality and Individual Differences*, *46*(5), 631–635.

Dawes, R. M. (1964). Social selection based on multidimensional criteria. *The Journal of Abnormal and Social Psychology*, *68*(1), 104.

Dellaert, B. G. C., Brazell, J. D., & Louviere, J. (1999). The effect of attribute variation on consumer choice consistency. *Marketing Letters*, *10*(2), 139–147.

Dellaert, B. G. C., Swait, J., Adamowicz, W., Arentze, T. A., Bruch, E. E., Cherchi, E., … Salisbury, L. C. (2017). Individuals' decisions in the presence of multiple goals. *Customer Needs and Solutions*, *71*(1), 1–14.

Desarbo, W., Ansari, A., Chintagunta, P., Himmelberg, C., Jedidi, K., Johnson, R., … Wedel, M. (1997). Representing Heterogeneity in Consumer Response Models. *Marketing Letters*, *8*(3), 335–348.

DeShazo, J. R., & Fermo, G. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management*, *44*(1), 123–143.

Dhar, R., & Nowlis, S. M. (1999). The effect of time pressure on consumer choice deferral. *Journal of Consumer Research*, *25*(4), 369–384.

Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of Marketing Research*, 146–160.

Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, *73*(3), 221–230.

Fishburn, P. C. (1968). Utility theory. *Management Science*, *14*(5), 335–378.

Folkman, S., & Lazarus, R. S. (1988). Coping as a mediator of emotion. *Journal of Personality and Social Psychology*, *54*(3), 466–475.

Gibbons, R. (1992). *A primer in game theory*. New York: Harvester Wheatsheaf.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. (Penguin, Ed.).

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, *62*, 451–482.

Gilbride, T. J., & Allenby, G. M. (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science*, *23*(3), 391–406.

Gilbride, T. J., & Allenby, G. M. (2006). Estimating Heterogeneous EBA and Economic Screening Rule Choice Models. *Marketing Science*, *25*(5), 494–509.

Guevara, C. A., Chorus, C. G., & Ben-Akiva, M. E. (2016). Sampling of Alternatives in Random Regret Minimization Models. *Transportation Science*, *50*(1), 306–321.

Hahn, M., Lawson, R., & Lee, Y. G. (1992). The effects of time pressure and information load on decision quality. *Psychology & Marketing*, *9*(5), 365–378.

Haugtvedt, C. P., & Petty, R. E. (1992). Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology*, *63*(2), 308.

Hauser, J. R. (1978). Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information-Theoretic Approach. *Operations Research*, *26*(3), 406–421.

Hauser, J. R. (2014). Consideration-set heuristics. *Journal of Business Research*, *67*(8), 1688–1699. https://doi.org/10.1016/j.jbusres.2014.02.015

Hauser, J. R., & Wernerfelt, B. (1990). An Evaluation Cost Model of Consideration Sets. *Journal of Consumer Research*, *16*(4), 393.

Haynes, G. A. (2009). Testing the boundaries of the choice overload phenomenon: The effect of number of options and time pressure on decision difficulty and satisfaction. *Psychology & Marketing*, *26*(3), 204–212.

Hensher, D. A. (2014). Process heuristics in choice analysis: An editorial. *Journal of Choice Modelling*, *11*(11), 1–3.

Hensher, D. A., & Greene, W. H. (2003). The Mixed Logit model: The state of practice. *Transportation*, *30*(2), 133–176.

Hensher, D. A., & Greene, W. H. (2010). Non-attendance and dual processing of commonmetric attributes in choice analysis: a latent class specification. *Empirical Economics*, *39*(2), 413–426.

Hensher, D. A., Greene, W. H., & Chorus, C. G. (2013). Random regret minimization or random utility maximization: an exploratory analysis in the context of automobile fuel choice. *Journal of Advanced Transportation*, *47*(7), 667–678.

Hoeffler, S., & Ariely, D. (1999). Constructing stable preferences: A look into dimensions of experience and their impact on preference stability. *Journal of Consumer Psychology*, *8*(2), 113–139.

Houthakker, H. S. (1952). Compensated changes in quantities and qualities consumed. *The Review of Economic Studies*, *19*(3), 155–164.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 90–98.

Huber, J., & Puto, C. (1983). Market boundaries and product choice: Illustrating attraction and substitution effects. *Journal of Consumer Research*, 31–44.

Huffman, C., Ratneshwar, S., & Mick, D. G. (2000). Consumer Goal Structures and Goal-Determination Processes: An integrative framework. In S. Ratneshwar, D. G. Mick, & C. Huffman (Eds.), *The Why of Consumption: Contemporary perspectives on consumer motives, goals, and desires* (pp. 9–35). London: Routledge.

Inman, J. J. (2007). Regret Regulation: Disentangling Self-Reproach From Learning. *Journal of Consumer Psychology*, *17*(1), 19–24.

Inman, J. J., & McAlister, L. (1994). Do Coupon Expiration Dates Affect Consumer Behavior? *Journal of Marketing Research*, *31*(3), 423.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006.

Jacoby, J., Chestnut, R. W., & Fisher, W. A. (1978). A behavioral process approach to information acquisition in nondurable purchasing. *Journal of Marketing Research*, 532–544.

Kahneman, D. (2003). A perspective on judgment and choice: mapping bounded rationality. *American Psychologist*, *58*(9), 697–720.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kamakura, W. A., Kim, B.-D., & Lee, J. (1996). Modeling Preference and Structural Heterogeneity in Consumer Choice. *Marketing Science*, *15*(2), 152–172.

Kamakura, W. A., & Russell, G. J. (1993). Measuring brand value with scanner data. *International Journal of Research in Marketing*, *10*(1), 9–22.

Kardes, F. R., Kalyanaram, G., Chandrashekaran, M., & Dornoff, R. J. (1993). Brand Retrieval, Consideration Set Composition, Consumer Choice, and the Pioneering Advantage. *Journal of Consumer Research*, *20*(1), 62.

Keeney, R. L. (1972). Utility functions for multiattributed consequences. *Management Science*, *18*(5-NaN-1), 276–287.

Koehler, E., Brown, E., & Haneuse, S. J.-P. A. (2009). On the Assessment of Monte Carlo Error in Simulation-Based Statistical Analyses. *The American Statistician*, *63*(2), 155–162.

Kray, L. J., George, L. G., Liljenquist, K. A., Galinsky, A. D., Tetlock, P. E., & Roese, N. J. (2010). From what might have been to what must have been: Counterfactual thinking creates meaning. *Journal of Personality and Social Psychology*, *98*(1), 106–118.

Lancaster, K. J. (1966). A new approach to consumer theory. *The Journal of Political Economy*, 132–157.

Lazarus, R. S. (1991). Cognition and motivation in emotion. *American Psychologist*, *46*(4), 352–367.

Lee, L., Amir, O., & Ariely, D. (2009). In search of homo economicus: cognitive noise and the role of emotion in preference consistency. *Journal of Consumer Research*, *36*(2), 173–187.

Li, L. (2013). *The Role of Antecedent Volition on Consumer Evaluative Processes and Choice Behavior*. University of Alberta.

Li, L., Adamowicz, W., & Swait, J. (2015). The effect of choice set misspecification on welfare measures in random utility models. *Resource and Energy Economics*.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*(368), 805–824.

Louviere, J., Hensher, D., & Swait, J. (2000). *Stated choice methods : analysis and applications*. Cambridge University Press.

Luce, M. F. (1998). Choosing to Avoid: Coping with Negatively Emotion-Laden Consumer Decisions. *Journal of Consumer Research*, *24*(4), 409–433.

Luce, M. F., Bettman, J. R., & Payne, J. W. (1997). Choice processing in emotionally difficult decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(2), 384–405.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical surveys*. Wiley New York, NY.

Lurie, N. H. (2004). Decision Making in Information-Rich Environments: The Role of Information Structure. *Journal of Consumer Research*, *30*(4), 473–486.

Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research*, 419–430.

Malhotra, N. K., Jain, A. K., & Lagakos, S. W. (1982). The information overload controversy: An alternative viewpoint. *The Journal of Marketing*, 27–37.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, *8*(3), 229–254.

Marley, A. A. J., & Swait, J. (2017). Goal-based models for discrete choice analysis. *Transportation Research Part B: Methodological*, *101*, 72–88.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York: Academic Press.

Miron, A. M., & Brehm, J. W. (2006). Reactance theory-40 years later. *Zeitschrift Für Sozialpsychologie*, *37*(1), 9–17.

Mogilner, C., Rudnick, T., & Iyengar, S. S. (2008). The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research*, *35*(2), 202–215.

Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin*, *32*(8), 1024–1036.

Mussel, P., Göritz, A. S., & Hewig, J. (2014). Which choice is the rational one? An investigation of need for cognition in the Ultimatum game. *Personality and Individual Differences*, *60*, S8.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, *19*(3), 113–126.

Nowlis, S. M. (1995). The effect of time pressure on the choice between brands that differ in quality, price, and product features. *Marketing Letters*, *6*(4), 287–295.

Parker, J. R., & Schrift, R. Y. (2011). Rejectable choice sets: How seemingly irrelevant no-choice options affect consumer decision processes. *Journal of Marketing Research*, *48*(5), 840–854.

Payne, J. W. (1982). Contingent decision behavior. *Psychological Bulletin*, *92*(2), 382–402.

Pieters, R., & Zeelenberg, M. (2007). A theory of regret regulation 1.1. *Journal of Consumer Psychology*, *17*(1), 29–35.

Pihlens, D. A. (2008). *The Multi-attribute Elimination-By-Aspects (MEBA) Model*. (doctoral dissertation) University of Technology of Sydney.

Pilli, L. E. (2012). *Julgamento e decisão: limites para a racionalidade do consumidor*. Retrieved from http://www.teses.usp.br/teses/disponiveis/12/12139/tde-17122012-161621/pt-br.php

Pilli, L. E., & Mazzon, J. A. (2016). Information overload, choice deferral, and moderating role of need for cognition: Empirical evidence. *Revista de Administração*, *51*(1), 36–55.

Pindyck, R. S., & Rubinfeld, D. L. (2013). *Microeconomics*. Pearson.

Raiffa, H. (2006). Preferences for multi-attributed alternatives. *Journal of Multi-Criteria Decision Analysis*, *14*(4–6), 115–157.

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

Russo, J. E., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 676–696.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Scammon, D. L. (1977). "Information load" and consumers. *Journal of Consumer Research*, 148–155.

Schoemaker, P. J. H. (1982). The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature*, 529–563.

Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisficing: happiness is a matter of choice. *Journal of Personality and Social Psychology*, *83*(5), 1178.

Shani, Y., & Zeelenberg, M. (2007). When and why do we want to know? How experienced regret promotes post-decision information search. *Journal of Behavioral Decision Making*, *20*(3), 207–222.

Shugan, S. M. (1980). The Cost of Thinking. *Journal of Consumer Research*, *7*(2), 99.

Simon, B. (2004). *Identity in modern society: A social psychological perspective*. Blackwell Publishing.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99–118.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*(1), 1–20.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research*, 158–174.

Simonson, I. (2008). Regarding inherent preferences. *Journal of Consumer Psychology*, *18*(3), 191–196.

Smith, S. M., & Levin, I. P. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making*, *9*(4), 283–290.

Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy*, *69*(3), 213–225.

Street, D. J., & Burgess, L. (2007). *The construction of optimal stated choice experiments: theory and methods* (3rd ed.). New Jersey: John Wiley & Sons.

Summerville, A. (2011). The Rush of Regret. *Social Psychological and Personality Science*, *2*(6), 627–634.

Summerville, A., & Roese, N. J. (2008). *Dare to compare: Fact-based versus simulation-based comparison in daily life*. *Journal of Experimental Social Psychology* (Vol. 44).

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

Swait, J., & Adamowicz, W. (2001). Choice environment, market complexity, and consumer behavior: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organizational Behavior and Human Decision Processes*, *86*(2), 141–167.

Swait, J., & Ben-Akiva, M. (1987). Incorporating random constraints in discrete models of choice set generation. *Transportation Research Part B: Methodological*, *21*(2), 91–102.

Swait, J., Brigden, N., & Johnson, R. D. (2014). Categories shape preferences: A model of taste heterogeneity arising from categorization of alternatives. *Journal of Choice Modelling*, *13*, 3–23.

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

Swait, J., & Marley, A. A. J. (2013). Probabilistic choice (models) as a result of balancing multiple goals. *Journal of Mathematical Psychology*, *57*(1), 1–14.

Swait, J., Popa, M., & Wang, L. (2016). Capturing Context-Sensitive Information Usage in Choice Models via Mixtures of Information Archetypes. *Journal of Marketing Research*, *53*(5), 646–664.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). New York: Cambridge University Press.

Tversky, A. (1972a). Choice by elimination. *Journal of Mathematical Psychology*, *9*(4), 341–367. https://doi.org/10.1016/0022-2496(72)90011-9

Tversky, A. (1972b). Elimination by aspects: A theory of choice. *Psychological Review*, *79*(4), 281–299.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Tversky, A., & Kahneman, D. (1981). The framing of decisions. *Science*, *211*, 453–458.

Tversky, A., & Shafir, E. (1992). Choice under Conflict: The Dynamics of Deferred Decision. *Psychological Science*, *3*(6), 358–361.

van Osselaer, S. M. J., & Janiszewski, C. (2012). A Goal-Based Model of Product Evaluation and Choice. *Journal of Consumer Research*, *39*(2), 260–292.

van Osselaer, S. M. J., Ramanathan, S., Campbell, M. C., Cohen, J. B., Dale, J. K., Herr, P. M., … Tavassoli, N. T. (2005). Choice Based on Goals. *Marketing Letters*, *16*(3/4), 335–346.

Verplanken, B., Hazenberg, P. T., & Palenewen, G. R. (1992). Need for cognition and external information search effort. *Journal of Research in Personality*, *26*(2), 128–136.

Von Neumann, J., & Morgenstern, O. (1947). *The theory of games and economic behavior* (2nd ed.). Princeton: Princeton University Press.

Weber, E., & Johnson, E. (2009). Mindful judgment and decision making. *Annual Review of Psychology*, *60*, 53–85.

Woodworth, R. S. (1938). *Experimental Psychology*. New York: Holt.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of Behavioral Decision Making*, *12*(2), 93–106.

Zeelenberg, M., & Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer Psychology*, *17*(1), 3–18.