UNIVERSITÀ DEGLI STUDI DI BOLOGNA
ALMA MATER STUDIORUM

DIPARTIMENTO DI SCIENZE AZIENDALI

DOTTORATO DI RICERCA IN DIREZIONE AZIENDALE
XX CICLO
**S.S.D. SECS-P/08**

TESI DI DOTTORATO

# CUSTOMER EVOLUTION IN SALES CHANNEL MIGRATION

Presentata da

**SARA VALENTINI**

Coordinatore dottorato
Prof. **FEDERICO MUNARI**

Chair
Prof. **ELISA MONTAGUTI**

Comitato
Prof. **GIAN LUCA MARZOCCHI**
Prof. **ELISA MONTAGUTI**
Prof. **SCOTT A. NESLIN**

Esame finale anno 2008

# Contents

# List of Figures



5.1 Average Amount Spent Over Time Break Down by Channel Usage .......... 94

# List of Tables

# Aknoledgements

# *Chapter 1*

# INTRODUCTION

## 1.1 Research Problem

In recent years, research on multichannel shopping behavior has increasingly expanded (see Neslin et al., 2006, and Rangaswamy and Van Bruggen, 2005 for a summary of the extant literature). Rangaswamy and Van Bruggen (2005) define customers who use more than one channel to interact with firms as *multichannel customers*.

Several managerial, (DoubleClick, 2004; Wall Street Journal, 2004; Myers, Van Metre and Pickersgill, 2004) as well as academic (Kumar and Venkatesan, 2005; Rangaswamy and Van Bruggen, 2005; Thomas and Sullivan, 2005) studies agree in considering *multichannel customers* as a great opportunity for firms.

An emerging generalization is that *multichannel customers* purchase higher volumes, exhibit higher loyalty, and may be exposed to more marketing as a consequence of being multichannel (Neslin et al., 2006a; Kumar and Venkatesan, 2005; Myers, Van Metre, and Pickersgill, 2004; Kushwaha and Shankar, 2005; and Ansari et al., 2007). Therefore, there is an increasing interest in understanding customers channel behavior in a multichannel environment.

In the past, consumers typically obtained all their channel services from a single integrated channel at all stages of their decision process, but in the last decade, the use of different channels at different stages of the decision process has become popular. It is straightforward that customers have rapidly expanded their channel experiences and preferences beyond traditional channels (such as stores) and they expect the company with which they do business to have a presence on all these channels (Blattberg, Kim and Neslin, 2008 p. 636). Doubleclick (2004), for example, reported a rise in multichannel shopping: 65% of shoppers used more than one channel of the same retailer to purchase items.

Firms have recognized the importance of increasing their channel variety in order to better address these diverse consumer needs (Deleersnyder, Geyskens, Gielens and Dekimpe, 2002). According to eMarketer (2006), 39% of Internet retailers operate three channels and 42% manage two channels. Furthermore, companies such as L. L. Bean and Dell have invested heavily in coordinating channel communications to the customer (Rangaswamy and Bruggen, 2005).

A deeper understanding of how customers' channel choices evolve over time is useful in order to achieve coordinated marketing actions in line with customers' propensity to buy using one or more channels.

Customers' behavior presents important dynamic components which might explain their channel choices. Customers gradually become aware of channel options and, they learn which channel best suits their needs in response to firm's direct marketing (Thomas and Sullivan, 2005). Actually, it is very difficult to envisage a static definition of a multichannel customer; it could be that customers become multichannel over time as a consequence of a progressive learning process. Little is know about this evolution, especially when customers are new to the firm.

Several exploratory studies have recognized the importance of dynamic elements in the study of customer channel choice. For example, Schoenbachler and Gordon (2002), Balasubramaniam et al. (2005), and Kumar and Venkatesan (2005) theoretically acknowledged the key role of past behavior and experience in their channel choice studies.

Formal approaches which model customers channel choices over time are particularly important to managers who wish to route customers to different channels over time, or learn how to create multichannel customers.

Researchers have begun to formally model the customer channel “migration” process. Migration can be thought of simply as channel choice, but this expression is used to convey that there is a particularly interest on how this choice process takes places over time (Blattberg, Kim and Neslin, 2008, p.647).

Specifically, in literature there are four works which have explored channel migration processes: Thomas and Sullivan (2005), Ansari, Neslin and Mela (2008), Venkatesan and Kumar (2007), and the Knox’s doctoral dissertation (2005).

This research we aims to add insights in the context of channel migration literature. In particular, we aim to study how the channel decision process of newly acquired customer evolves over time.

Although recently the dynamic of customers' channel choices has been explored a limited effort has been made to investigate and formally model the learning process, i.e. how customers’ decision process changes over time as consumers learn their preferences and become familiar with the firm’s marketing activities. The only exception is the work of Knox who models the process whereby consumers evolve to form several channel usage segments. He includes an initial or “learning” segment, and he found that customers from the learning segment migrate towards the online segment.

Despite this first contribution about the role of learning on channel choice evolution, many important questions still remain:

1) Why do some people switch channel decision processes while others don’t?
2) How the channel decision process changes over time among the people who switch?

3) When and where marketers can exert leverage on the channel choice process?

4) What are the main channel migration's patterns?

To investigate these substantive questions we develop and estimate a model of customer channel migration. Theoretically, we ground our work on the customers decision-making theory which suggests us that several patterns of decision-making strategies are possible. An extended problem solving combined with learning might induce customers to change their decision process. However, the decision-making theory suggests us that new customers are not necessarily more likely to be learning prone, in other words an extended problem solving for newly acquired customers is not "guaranteed". By contrast, customers may starts their relationship with the firm using simplified decision making rules, subsequently some events (e.g. negative experiences, the firm marketing stimuli, etc.) might trigger customers' motivation and commitment with the choice task and induce them to change decision process or customers may simply keep on a straightforward decision process over time.

In particular, we draw our modeling approach on the Aaker's (1971) new-trier model. He argues that during the *trial* period the customer is essentially sampling purchase options and learning in the process. At some point, the customer transitions to his or her equilibrium decision process. Along this line, if customers are learning we distinguish between two stages in their channel choices evolution. An initial or "*trial*" stage when the customer is acquiring experience with the company channel offer, and a second or "*post-trial*" stage representing the decision process the customer evolves to in the long term.

Our contribution is fourfold: i) we propose a set of key phenomena that are related to the customer's propensity to change channel decision process, ii) we show how these phenomena can be modeled and estimated, iii) we show how choice decision patterns might evolve over time, iv) we contribute to the understanding of the marketing role in different channel decision-making situations.

Among the key findings we outline several factors which significantly increase or decrease the probability to move towards a *post-trial* channel decision strategy. Furthermore, we find that customers' responsiveness to marketing strongly differs between the *trial* and *post-trial* stages. In addition, different migration patterns are delineated which present distinctive characteristics, e.g. different *trial* length and different marketing responsiveness. Finally we demonstrate that there is large evidence of people who remain always with their initial channel decision making strategy.

## 1.2 Structure of the Study

This dissertation is composed of 6 chapters, organized as follows. The present Introduction opens the work. We presented the structure of the problem we want to analyze and we have outlined its main contribution.

In Chapter 2 we present a literature overview to place our research problem in a theoretical perspective. We present the recent literature on multichannel customer behavior with the purpose to show the state of knowledge on this emergent topic. Specifically, we focus on the customers channel choice models and on its determinant. Moreover, since our interest is on the dynamics of channel choice, we present the modeling contribution about the evolution of customer channel choice over time. In particular we review the customer channel migration models and their main results. Finally we end reviewing the literature on channel choice decision making process.

In Chapter 3 we outline the research objectives and the conceptual framework. We present the original features of this work, the hypothesis and the upfront theory.

Chapter 4 concerns the model development. This chapter starts with an introduction on possible different modeling approaches. Then, we present developed model. Finally, we end giving information on the estimation approach.

In Chapter 5 we present the data, the model selection and the results. The chapter starts with a description of the data set which underlines its attractiveness to the present study. In addition, we show detailed descriptive statistics about this data set. Then, we present different tested versions of the outlined model and we indicate the best model. Finally, we present the results: the learning model estimates, the factor influencing the probability to move to a *post-trial* model, the overall parameter estimates with distinct parameters for the *trial* and *post-trial* stages, and we end showing different types of migration patterns.

Chapter 6 is the conclusive section where we outline our findings starting from the theoretical framework applied to the specific results and contribution of this work. We discuss our findings; illustrate limits and future directions of this research. Finally, we outline some managerial implications and make some concluding remarks.

# Chapter 2

# THEORETICAL BACKGROUND

This dissertation focuses on multichannel customer behavior. Consumers display complex shopping behaviors in the emerging multichannel environment (Alba et al., 1997; Peterson, Balasubramanian, and Bronnenberg, 1997). This evidence originated a definite research stream which aims to understand customers' behavior in multichannel setting. In reviewing the literature, I will follow this order. First, a general overview about the emergence of the multichannel phenomenon is provided. Second, I review the determinants of customers' channel choice. Third, I present research efforts made on the evolution of customers' channel choices over time and on customer channel migration models. Finally, I end this review presenting the first attempts made to build a multichannel theory. Therefore I discuss the channel choice decision making process, underling what is know, what it has been empirically tested and what it is still under-researched.

## 2.1The Multichannel Phenomenon

For decades, the increasingly diverse needs of an ever more fragmented market have compelled firms to increase their product variety as strategy. Recently, firms are also turning to a second strategy to better address these diverse consumer needs: they increase their channel variety (Deleersnyder, Geyskens, Gielens  and Dekimpe, 2002). In the past consumers typically obtained all their channel services from a single integrated channel at all stages of their decision process, but in the last decade, the use of different channels at different stages of their decision and shopping cycles has become more and more popular.

(Rangaswamy and Van Bruggen 2005). This trend is well-documented and it continues to increase (Coelho et al., 2003; Dutta et al., 1995; Easingwood e Storey, 1996; Frazier, 1999; Moriarty e Moran, 1990). As a consequence, customers, who rapidly expand their channel experiences and preferences beyond traditional channels (such as stores), expect the company with which they do business to have a presence on all these channels (Blattberg, Kim & Neslin, 2008). These channels include the Internet, call centers, sales forces, catalogs, retail stores, and in the near-future, interactive television (Blattberg, Kim and Neslin 2008).

Blattberg, Kim and Neslin (2008), in their analysis of multichannel literature, describe several factors that might have "pushed" and "pulled" the emergence of this phenomenon. They argue that the push toward multichannel has been driven by three main events: (1) by companies which have expanded their channel offer, (2) by customers who rapidly expanded their multiple channel usage, and (3) by competitive forces (e.g. if company A initiates a multichannel strategy, company B has to follow it). By contrast, the pull toward multichannel has been driven by potential advantages which have encouraged companies to adopt a multichannel management strategy, e.g. improvements in loyalty, in sales growth, and in efficiency.

Several managerial, (DoubleClick, 2004; Wall Street Journal, 2004; Myers, Van Metre & Pickersgill, 2004) as well as academic (Kumar & Venkatesan, 2005; Rangaswamy & Van Bruggen, 2005; Thomas & Sullivan, 2005) studies agree in considering multichannel customers as a great opportunity for firms.

Rangaswamy and Van Bruggen (2005) refer to customers who use more than one channel to interact with firms as *multichannel customers*[1]. A customer could be defined as multichannel in all the following situations: i) if he purchases a new bike at the closest store

[1] They also refer to marketing strategies to reach such customers as *multichannel marketing*.

but the bike equipment using the “Wal-Mart” web site, ii) if he searches for information on a new car online but than purchase it from the car dealer, iii) if he purchase books indifferently using the Barnes & Nobles website or the closest Barnes & Nobles store. From this example it emerges that the definition of multichannel customer ranges from channels as means to obtain information to channels as means to make purchases, from manufactures’ channels to retailers’ channels, from the same brand channels to competing brands channels.

It is important to clarify that customers may use distribution channels for at least two different purposes. On the one hand channels might be used as means to obtain information; on the other hand as means to purchase items or services (Van Baal e Dach, 2005; McLean e Blackie, 2004; Hdsapple e Sing, 2000).

Neslin et al. (2006) define “channel” a customer contact point, or a medium through which the firm and the customer interact. Interestingly, they do not include one-way communications, such as television advertising because these “channel” forms do not require an interaction between customer and firm, though they do include home shopping television networks and direct response advertising in mass media. Channels present different attributes and characteristics, for example they can differ in term of assortment, layout, location, price, services offered, delivery, etc. (Berman, Evans 1995).

Firms adopting a multichannel marketing strategy, i.e. synchronized distribution channels mix, have different tools to differentiate their offer and their positioning. A specific mix of distribution channels, for example, could attract variety seeking customers or new targets and it could even help firms to maintain the current customer base. Rangaswamy and Van Bruggen (2005) define multichannel customer management as the design, deployment, coordination, and evaluation of channels in order to enhance customer value through effective customer acquisition, retention, and development. Specifically, they argue that multichannel marketing enables firms to build long-term customer relationships by simultaneously offering their customers and prospects information, products, services, and support (or any combination of these) through two or more synchronized channels. Thus, for example, a firm might deploy multichannel marketing strategies and tactics to help customers to browse for product information at a Web site, then purchase at a store, and later obtain technical support over the telephone. By carefully synchronizing its channels, a firm creates superior channel service outputs and gives its customers fewer reasons or opportunities to switch to competitors because of inconvenient channel access, or loss of control in interacting with the firm. Also, by tracking customer behavior across channels, firms can improve their understanding of their customers' decision making and develop a basis for creating strong relationships with customers and improving retention.

It should be noticed that this stream of research centers on the consumers. This implies that multichannel customer management is a customer-centric marketing function, and it should not be confused with the traditional sales channel research, which focuses on the firm and distributors relationship. Neslin et al. note that marketers have always considered channel management to be a fundamental component of the marketing mix (e.g., Stern and El-Ansary 1972; Webster 1991). However, while traditional channel management has taken the perspective of the firm, multichannel customer management centers on the customer, on the creation of customer value as a means to increase firm value (Payne and Frow, 2005; Boulding, Staelin, Ehret, and Johnston, 2005; Rangaswamy and Van Bruggen, 2005).

In addition, multichannel marketing should not be confused with the traditional multiple-channel marketing, in which a firm interacts with different segments of the customer base through different channels, for example, using personal selling for large customers and using retailers for small customers. In multichannel marketing, customers can use alternative channels to reach the departments within the firm at their discretion, and they may choose different channels at different times (Rangaswamy and van Bruggen 2005).

In summary, customers are the core of a multichannel strategy because customers choose the channel through which to interact with a firm at a given time. That channel not only provides the information and services to meet the customers' needs, but also facilitates further interactions with other channels and divisions of the firm, if needed, to take care of the customer.

## 2.2 Channel Choice

### 2.2.1 Store and Retail Format Choice

Previous stream of literature contributes to an understanding of store choice (see Wrigley 1988 for a review) and on multi-format choice (e.g. Messinger and Narasimhan 1997; Fox et al. 2004; Kumar 2004; Inman Shankar and Ferraro 2004).

The modeling and theoretical contributions concerning store choice can be distinguish into two general classes.  The first encompasses Ehrenberg's wish whereby store choice is modeled via NBD and Dirichlet distributions. This work aims to examine store choice using data on penetration, purchase frequency and market share (Kau and Ehrenberg, 1984; Wrigley and Dunn, 1984; Uncles and Ehrenberg, 1986). The second stream of research encompasses discrete choice models in which store choice depend on factors concerning store attributes and customers characteristics (Barnard and Hensher 1992, Bell, Has and Tang, 1998; Bell and Lattin, 1998; Ho, Tang and Bell, 1998). This work shows that consumers' decision to shop at a store depends primarily on location, assortment and quality of the merchandise, service, price image, promotions on assortment, basket attractiveness, distance, fixed and variable costs, and experience (Barnard e Hensher 1992; Bell, Ho e Tang 1998; Bell e Lattin 1998; Ho, Tang e Bell 1998; Chan, Ma, Narasimhan e Singh 2005). More specifically, Bell and

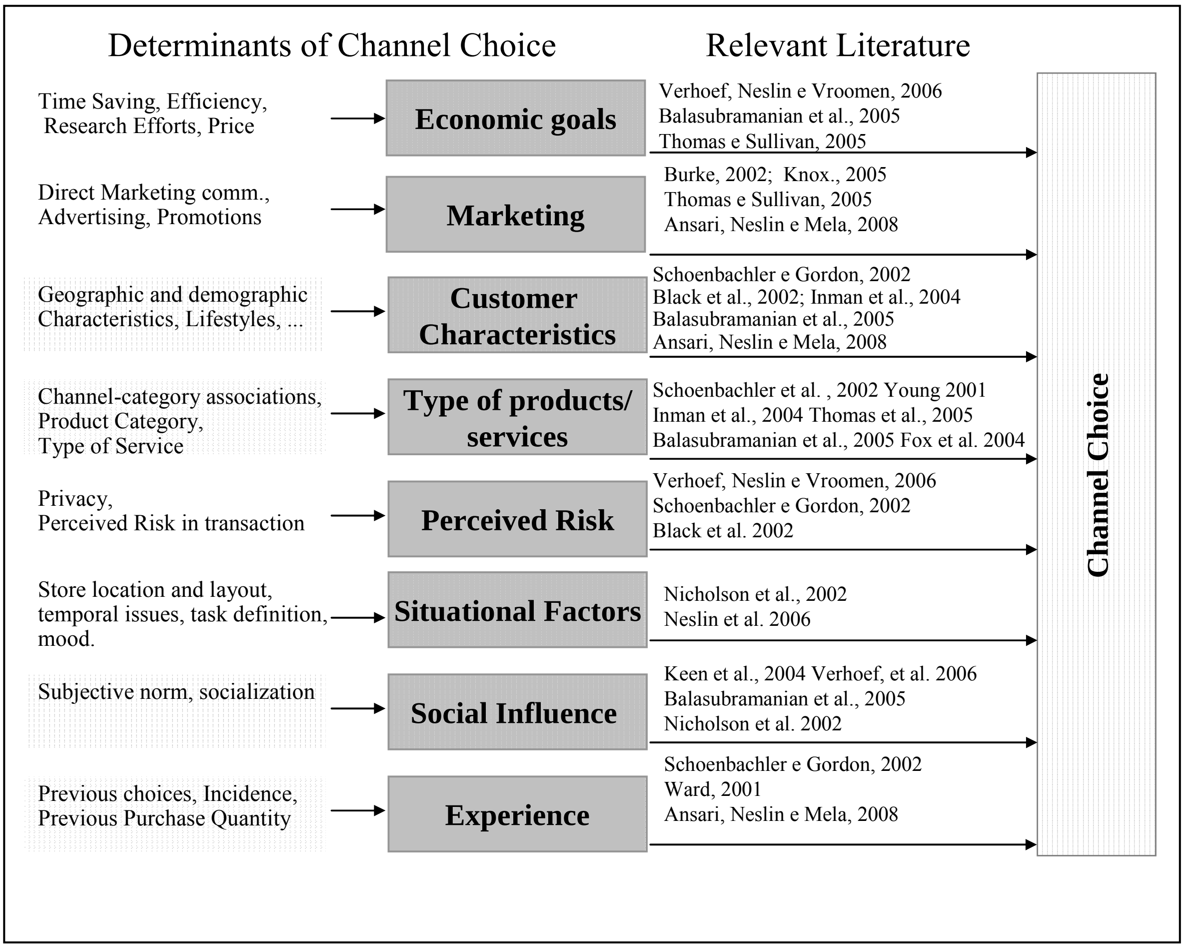
Lattin (1998) recognize that the customer' shopping behaviour is complex and that shoppers often visit multiple stores (Bell and Lattin, 1998). These results, however, were mainly based on grocery on date and no effort was made to examine different formats.

Naturally, the literature on store choice has been expanded to account for different retail formats with the purpose to understand what motivate customers to make use of retail formats (Messinger and Narasimhan 1997; Fox et al. 2004; Kumar 2004; Inman Shankar and Ferraro 2004). Fox et al. (2004), for example, in an empirical study on household behavior across retail formats (i.e. grocery stores, mass merchandisers, and drug stores), assess how competition differs across formats and they explore how retailers' assortment, pricing, and promotional policies, as well as household demographics, affect customer decision to use different retail formats. They find that consumer expenditures respond more to varying levels of assortment (in particular at grocery stores) and promotion than price. They also find that households that shop more at mass merchandisers also shop more in all other formats, suggesting that visits to mass merchandisers do not substitute for trips to the grocery store.

### 2.2.2 Channel Choice and its Determinants

The most heavily researched area of multichannel customer management concerns the antecedents of customer channel choice (see Neslin et al., 2006a, and Blattberg et al., 2008, p.641 for a review). Much work has been conducted on the determinants of channel choice. Figure 2.1 summarizes the most important determinants that have been studied. In the next paragraphs these determninants are discussed in detail. I end this section showing the first evidences about the role of experience in channel choice. This discussion opens the paragraph 2.3 which concerns the role of dynamic components in multichannel models.

*Figure 2.1: Determinants of Channel Choice*




### 2.2.2.1 Economic Goals

Channel choice research has identified customers' price expectations (e.g., Brynjolfsson and Smith 2000), and convenience (Forster 2004) as factors that may lead to a specific choice among channels or stores (Fox, Montgomery, and Lodish 2004).

Thomas and Sullivan (2005) include price (i.e. dollar amount paid for the product) as a predictor in their model of channel choice. Balasubramanian, Raghunathan, and Mahajan (2005) assert that the goals (e.g., economic) a consumer tries to achieve during his or her shopping experience affect channel choice. These studies suggest that customer may desire to achieve specific economic goals when they select channels. For example, for some customers it might be important to save money, or save time therefore they select channels which allow them to minimize money or time spent.

Balasubramanian et al. (2005) develop a conceptual framework with the purpose to clarify the utilities that consumers using a channel derive from both the purchase process and the purchased products. Their framework examines how consumers' pursuit of efficiency and utility can influence channel choices and consider pure economic goals. When customers goals are purely economic, consumers focus on maximizing net utility, defined as the utility they derive from the good less the total costs of obtaining it, which, apart from price, may include the real costs of travel, the opportunity cost of time, and the implicit cost of inconvenience (Balasubramanian, 1998). They argue that a consumer pursuing purely economic goals would base channel choice on a careful trade-off of the costs and benefits of using specific channels at the different stages of the purchase process.

Verhoef et al. (2007) aim at understanding the causes of research shopping phenomenon, i.e. customer tendency to use one channel for search and another for purchase. They argue that searching in one channel and purchasing in another channel may provide economic benefits. For instance, searching on the Internet may provide consumers with information on price, which allows them to have a better deal in the store through negotiation or better informed choices. In addition, they argue the customers search process as well as the customers purchase process depends on benefits and costs such as negotiation possibilities (i.e. whether consumers are able to negotiate on price and other aspects of the products), purchase effort (i.e. the difficulty and time costs consumers experience when purchasing a product using a specific channel), price level (i.e. consumers perceptions of prices in a specific channel), and search convenience (i.e. the easiness and speed at which consumers can gather information on products in the specific channel).

#### 2.2.2.2 Marketing

Recent work shows that marketing effort can drive channel choice. Thomas and Sullivan (2005a) found that direct marketing influences store, catalog, or internet choice and that this influence differ among groups of customers. Specifically, they found two segments. For one segment, direct marketing expenditures foster migration from the store to Internet. For the other segment, direct marketing expenditures favor customers' migration to the store rather than the Internet. Furthermore, they explored nonlinearities in the relationship between direct marketing communication and channel choice. They found that the nonlinear effects exist and, specifically, an increasing in marketing expenditures motivate segment one customers to make the first repeat purchase from the catalog and segment two customers from the store. The main finding is that marketing affects migration, and different segments may respond differently to such stimuli.

These results were found even in a more recent works by Ansari, Mela and Neslin (2008). These authors model incidence, channel choice and order size and they used emails and catalogs sent as a measure of direct marketing communications, instead of the amount of dollar spent. They also include interactions between communications. Interestingly, they found that marketing variables all have a positive direct effect in the incidence model and that the interactions between communications are negative, implying cannibalization and decreasing return effects. They found that emails were strongly associated with choice of the Internet. This is plausible considering that Emails and the Internet are basically the same technology, and the availability of a click-through URL in an email would encourage movement to the Internet (Blattberg, Kim and Neslin 2008). They also found diminishing marginal return to emails in terms of driving persons to the Internet implying that a pulsing strategy might be more effective.

Knox in his doctoral work, as Ansari et al (2008), analyze the effects of direct marketing communications (Email and catalogs) on incidence, order size and channel choice, and as Thomas and Sullivan (2005) he assumes customers to be in an unobserved segment. The possible segments include learning, online, offline and multichannel. He found support of newly acquired customers evolving behavior among these segments. As Ansari et al. (2008) he found that both emails and catalogs have a greater effect on purchase incidence, whereas emails have a greater effect on channel selection. In the learning segment all marketing communications (email, catalog, and both emails and catalogs) at the same time lead to increase choice of the online channel. The online segment was highly responsive to emails and this appeared to guide them toward the Internet. In the offline segment catalogs drive offline purchase, whereas sending both drive online purchase.

Finally, Venkatesan et al. (2007) studied the time it took customers to adopt and additional new channel, given they had bought earlier from another channel. They measure marketing communications (direct mail or Email) as the ratio of the sum of the number of marketing communications sent by the firm between two consecutive customer transactions to the total number of transactions the customer made. These authors found that marketing communications (direct mail and email) had an inverse U-shaped relationship with the timing of new channel adoption. Up to a point, increasing communications would shorten the time till adoption, however, after that threshold, increasing communications would actually lengthen the time till adoption.

In summary, marketing efforts influence channel choice, and the influence is heterogeneous across customers. The influence of marketing communications on customer behavior is nonlinear in nature an increase in the level of marketing communications might motivate customers to use specific channels (Thomas and Sullivan 2005), but after a certain threshold it can have dysfunctional consequences because customers might begin to perceive the firm as not understanding their needs and simply pushing its products (Venkatesan et al., 2007). Ansari et al. (2007), and Knox (2005), have also found that marketing influences purchase incidence. As a result, marketing influences sales volume as well as the channel that produces that volume.

#### 2.2.2.3 Customer Characteristics

Several individual difference variables are associated with channel choice, including age, gender, education, income, family size, and region (Neslin et al., 2006; Ansari et al. 2007; Gupta, Su, and Walter 2004; Inman, Shankar, and Ferraro 2004; Kushwaha and Shankar 2005; Verhoef et al. 2007). Income may play an important role in channel choice, particularly the choice of the online channel. Age is another likely determinant of channel choice. For example, there is a strong association between the age of US population and the likelihood of Internet usage (Kushwaha and Shankar 2005). Inman, Shankar and Ferraro (2004) found that different socio-economic classes have different predispositions to buy different product categories from different types of channels and consumer demographics play an important role in determining the share of volume of a channel. Therefore, demographics are predictors of the internet use, online shipping and catalog shopping and are likely to influence single and multiple channel shopping behavior (Schoenbachler and Gordon, 2002; Black et al. 2002; Kushwaha and Shankar 2005).

Black et al. (2002) include among factors which appear to influence channel choice the consumer confidence, lifestyle factors, motivations and emotional responses. Similarly, Schoenbachler and Gordon (2002) argue that customers' lifestyle factors (e.g. need for convenience, views on shopping for entertainment) have some influence on motivation to buy through a specific channel. Inman, Shankar, and Ferraro (2004) show that their "geodemographics" factor impacts on channel choice. They explain that people in different social classes differ not only in terms of the products they buy but also in terms of the type of store they frequent to buy products. This is in line with previous marketing literature which demonstrates that shopping sites tend to take on fixed class identities, i.e. each store, even if it is a grocery store, acquires status identification. (Miller et al., 1998; Martineau, 1958).

Interestingly, Thomas and Sullivan (2005b) find that also the stage in the customer lifecycle determines channel choice.

#### 2.2.2.4 Type of Products/Services

Another factor proposed to affect whether a consumer is a single or multichannel customer and which channel is used is the product/service category. Certain product categories are more often purchased via certain channels and others are more often purchased via traditional retail means (Schoenbachler and Gordon, 2002).

Channel choice research has identified the product group to be purchased (Young 2001), as factors that may lead to a specific choice among channels or stores (Black et al. 2002, Fox, Montgomery, and Lodish 2004, Thomas and Sullivan 2005). Consumers may be pleased with specific products at specific stores that appear to suit their needs or to be good deals (Balasubramanian et al., 2005).

Inman, Shankar, and Ferraro (2004), extending the literature on brand associations to the context of channels, introduce the concept of “Channel-Category Associations” as predictors of channel patronage, i.e. the categories of consumer goods that are most closely associated with particular channels. They argue that when the consumer is considering purchasing a given set of goods, the likelihood of a particular channel coming to the fore should be a function of the sum of its associations with each product category being considered and the specific features/ benefits offered by each channel. They found strong evidence of channel- category associations. For example, the grocery channel is associated with food products, the drug channel with medications and health-related products, and the mass merchandiser channel with household items. In contrast, cleaning supplies, automotive, gifts, beauty care, miscellaneous household items, and paper goods are most closely related to the mass merchandiser channel; and tobacco, alcohol, candy, magazines, and soaps are perceived as closest to the drug channel. Their findings suggest that the channel-category associations influence channel share of volume both directly and indirectly.

#### 2.2.2.5 Perceived Risk

Purchase risk concerns perceived uncertainty in buying products through a specific channel. It refers to the individual's personal assessment of the risk associated with the purchase. The perceived risk can be financial, social or physical or some combination (Schoenbachler e Gordon, 2002). Schoenbachler and Gordon (2002) argue that perceived risk depends on from several factors including customers' familiarity with the channel or he brand name, with the specific company and with the price of the product/service.

Interestingly, researchers have pointed to the importance of trust for online shopping, as consumers cannot physically check the quality of a product or monitor the security of sending sensitive personal information (privacy) or financial information and payments (Hoffman, Novak and Peralta 1999; McKnight, Choudhury and Kacmar 2002, Black et al. 2002; Verhoef et al. 2007). In the same way, Verhoef et al. (2007) found that the store channel is particularly strong on risk, and privacy. He also found that privacy is not grouped into the risk factor, but appears to be a separate factor.

#### 2.2.2.6 Situational Factors

Nicholson, Clarke, and Blakemore (2002) seek to explore how and why consumers select particular modes of shopping in specific situations. Specifically, they apply a Belkian analysis of situational variables present in the consumer setting with the aim to establish which Belkian environmental dimensions dominate when a particular channel becomes the preferred shopping mode. For example, consumer A shops in a store on a specific occasion because the purchase is urgent and, for instance, making remote ordering is impossible. According to Belk, situational variables are all those factors particular to a time and place of observation (Belk 1974). Such attributes are classifiable according to five distinct dimensions of situational influence: (1) physical setting (e.g., weather), (2) social setting, (3) temporal issues (time of day, urgency of the purchase), (4) task definition (e.g., type of product), and (5) antecedent state (e.g., mood). These authors found that temporal factors appear to exert a powerful positive influence in the selection of the Internet shopping option for instance, but have little impact upon selection of the store-based option. Similarly, Neslin et al. (2006) cite the five situational factors suggested by Nicholson, Clarke, and Blakemore (2002) and they note that particular attention has been devoted to task definition, hypothesizing for example that experience goods are more likely to be purchased at a store, while search goods are more likely to be bought on the Internet (Mathwick, Malhotra, and Rigdon 2002), and customizable products are more likely to be purchased on the Internet (Mahajan, Srinivasan, and Wind 2002).

#### 2.2.2.7 Social Influence

Several studies demonstrate that channel usage depend also on the influence of the opinions of “significant” others and on socialization experiences.

The earlier study of Nicholson et al. (2002) considered a “social setting” variable in their conceptual framework which aim to explain channel choice. This variable focuses on the presence or absence of others, together with their social roles, role attributes and opportunities for interaction. It is therefore a dimension which encapsulates everything from awareness of security staff in the store and opportunities for interaction with in-store sales staff. Similarly, Balasubramanian et al. (2005) include in their conceptual framework the quest for socialization, i.e. how consumers' need to be part of social milieus or of stimulating environments can influence channel choices. They argue that the opportunities for social interaction that it affords the presence of others during shopping may increase utility.

Keen, Wetzls, de Ruyter, and Feinberg (2004), used the theory of reasoned action framework and they included "subjective norm" as a driver of channel choice. These authors found that subjective norm is an important driver of channel choice, although channel format and price are, in their analysis, turns out most important. In a similar way, Verhoef, et al. (2007) based on the well-known theory of reasoned action their study and they modeled customers' channel preferences as a function of several attributes, including among these "clientele", i.e. the perceived use of this channel for either search or purchase by relatives and acquaintances (reference groups). These authors found that clientele particularly influenced customers' choice of the Internet. They explain this result arguing that Internet is a new channel and one would expect consumers to model their behavior after peers when they personally have less experience. It also could relate to the "community" concept behind the Internet.

#### 2.2.2.8 Experience

Another important individual-difference variable is channel experience. Ward (2001) provides a theory for the role of channel experience. He proposes that customers make human capital investments in learning to use particular channels. If the skills gained through these investments "spillover" to other channels, customers can become multichannel users. For example, a skill in using a catalog is the ability to determine the best product without actually touching it. This skill spills over to using the Internet, and as a result, the catalog and Internet become substitutes. It is very interesting the way this author test these assumptions empirically. To measure spillover, Ward obtains data on customer purchases, by channel, in several product categories. The author estimates an equation of the customer's propensity to purchase a category in each channel as a function of channel-specific and category-specific dummies. The residuals from these regressions represent effects that cause deviations from which channel we would expect the customer to use on average. By correlating these residuals between channels, the author estimates spillover. Results suggest that the spillover effects are largest between online and direct marketing. His works is interesting because it supports the idea that customers learn when they are using channels and that this learning contributes to create a channel related experience which might influence channel future channel choices. This idea is also supported by Schoenbachler and Gordon (2002) in their theoretical framework which examines the role of experience on multichannel behavior. They argue that customers purchasing from a channel are more likely to purchase from the same channel in the future, implying a strong positive relationship between past experience and channel choice. These authors do not test the proposed model empirically but they recommend the inclusion of experience recognizing that in channel choice past behavior may be a strong predictor of future behavior. Along this line, Montoya-Weiss et al. (2003), Inman et al. (2004), and Ansari et al. (2008) find that experience in using a particular channel makes it more likely the customer will use that channel in the future. However, they do not explore whether this result is due to mindless inertia or to cognitive learning (Blattberg, Kim and Neslin 2008).

The role of experience on channel choice is complex because this variable is naturally dynamic and it implicitly includes different components which might be difficult to isolate. Actually, there are several reasons that might generate channel choice dependence over time and previous experience represents only one possible explanation. In the next paragraphs I will discuss deeply this issue, presenting the earlier dynamic channel choice models.

## 2.3 Dynamic Components in Multichannel Models

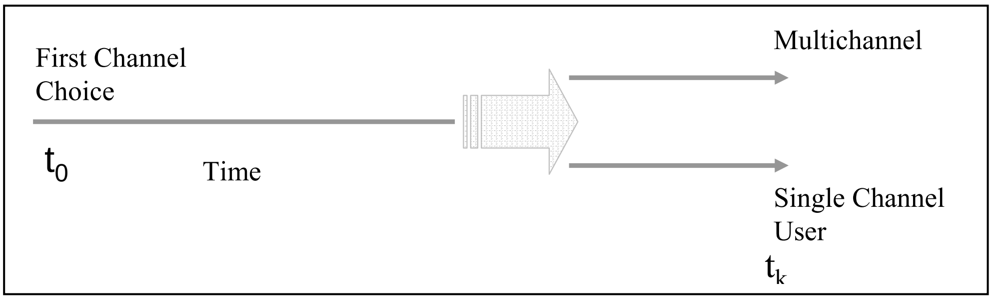
A deeper understanding of customers' habits concerning channel choice is useful in order to achieve coordinated marketing actions in line with customers’ propensity to buy using one or more channels.

Although recently the dynamic of customers' channel choices has been explored (Ansari, Mela & Neslin, 2008; Thomas & Sullivan, 2005; Knox, 2005; Venketesan & Kumar, 2007) a limited effort has been made to investigate and formally model the learning process, i.e., how customers’ decision process changes over time as consumers learn their preferences and become familiar with the firm’s marketing activities. A deeper understanding of this evolution and the process by which customers become loyal to certain channels is important.

Static choice models are based on a questionable assumption: choice at time t does not depend upon previous choices. Actually, customers’ channel behavior can be dynamic as customers become aware of channel options and gradually learn which channel best suits their needs in response to firm’s direct marketing (Thomas and Sullivan 2005). Furthermore, it is very difficult to envisage a static definition of a multichannel customer. It is more likely that customers become multichannel over time as a consequence of a progressive learning process (see figure 2.2). Thus, it is through analyzing the whole sequence of repeat purchases with a firm that one can effectively understand if a customer 1) purchases through only one channel

2) purchases through more than one channel 3) purchases through only one channel, but as a result of a migration process.

*Figure 2.2: Multichannel Customer Dynamic*




### 2.3.1 Dynamic on Choice Behavior

Customer choice in consumption experiences often depends on past behavior. For this reason assuming a static choice behavior may produce a distorted representation of market structure and of choice determinants effects.

There are two important reasons for incorporating dynamics into empirical choice models. Substantively, the dynamics may be more "realistic" and, hence, may provide a better description of behavior. More importantly, there may be patterns in the data that are simply not captured by a static model. Hence, ignoring the dynamics could potentially "throw away" valuable information and, worse, could generate misleading conclusions about behavior (Dubè et al. 2005).

In a variety of contexts it is often noted that individuals who have experienced an event in the past are more likely to experience it again in the future, whereas individuals who have not experienced the event are not as likely to experience the it in the future (Hsiao 1999).

In other words, the conditional probability that an individual will experience the event in the future is a function of past experience.

Erdem and Keane (1996) in the context of brand choice demonstrate that ignoring consumer choice dynamics in market structure modeling may lead to misleading implications and unrealistic conclusions for product management. A vast literature in brand choice deals with choice dynamics by primarily focusing on the impact of past purchases on current choices.

There are two diametrically opposite explanations for often observed empirical regularity in choice behavior with which individuals who made a choice in the past are more likely to experience that choice in the future: habit persistence and individual heterogeneity. In the former case (i.e. habit persistence) as a consequence of experiencing an event, preferences relevant to future choices are altered. Past experience has a genuine behavioral effect in the sense that an otherwise identical individual who has not experienced a product (or a brand, channel, etc.) will behave differently in the future than an individual who has experienced it. It occurs because tastes and customers' preferences may systematically change as a consequence of habit persistence (e.g., Kuehn 1962, Jeuland 1978, Chamberlain 1978; Pollak 1970, 1976; Spinnweyn 1981; Heckrnan 1981), that is reinforcement of tastes and preferences over time or of an explicit desire among individuals to seek variety in choice (McAlister and Pessemier 1982, McAlister 1982) or both (Kahn, Kalwani and Morrison 1986). In the latter case (i.e. individual heterogeneity) individuals may differ in certain unmeasured variables that influence their choice probability but that are not influenced by the experience in their choice behavior. Previous experience appears to be determinant of future experience exclusively because it as a proxy for temporally persistent unobservable factors that determine choices (e.g. risk aversion). It is well known that heterogeneous preferences can lead to a spurious appearance of time dependence (Heckman 1981).

Finally, it should be noted that external information obtained from the market or from firm communications may generate time depenence. The customer in several consumption situations may update his or her beliefs about the specific object of the choice basing on retrieved information. These information may influence choice. For this reason marketing variables (e.g. advertising, promotions, direct communications, etc.) might generate temporal dependence on choice.

### 2.3.2 First Evidences of Dynamic Elements on Channel Choice

In the context of multichannel literature several exploratory studies have included dynamic elements in consumer channel choices. For example, Schoenbachler and Gordon (2002), Balasubramaniam et al. (2005), and Kumar and Venkatesan (2005) do not explicitly estimate dynamic channel choice models, but they theoretically acknowledged the key role of past behavior and experience in channel choice.

Schoenbachler and Gordon (2002) include in their theoretical framework of "multichannel buying" past direct marketing experience (see paragraph 2.2.2.8). They argue that customers who have purchased from a catalog are more likely to purchase from a catalog in the future and customers who have purchased online are more likely to purchase online. Similarly, Balasubramaniam et al. (2005) include experience and consumers' reliance on schemas and script for shopping which can be thought as intrinsically dynamic variables.

One of the earlier model which accounts empirically dynamic variables is the work of Kumar and Venkatesan (2005). These authos in the B2B context use an ordered logistic regression where the dependent variable is the number of channels a customer used. They use several covariates which depends on customers past behavior, e.g. returns (number of returned items in the customer lifetime), tenure (number of years between the customer's first purchase), past customer value (historic cumulative profits obtained form a customer). Once more, this work is not a model dynamic because they mase cross-sectional analysis but the authors recognize the importance of past events in choosing and operazionalizing their covariates of multichannel behavior to explain multichannel decisions.

### 2.3.3 Customer Channel Migrations

Formal approaches which model customers channel choices over time are particularly important to managers who wish to route customers to different channels over time, or learn how to create multichannel customers.

Researchers have begun to model the customer channel "migration" process. Migration can be thought of simply as channel choice, but this expression is used to convey that there is a particularly interest on how this choice process takes places over time (Blattberg, Kim and Neslin, 2008, p.647).

Specifically, in literature there are four works which have explored channel migration process: Thomas and Sullivan (2005), Ansari, Neslin and Mela (2008), Venkatesan and Kumar (2007), and the Knox's doctoral dissertation (2005). Although these works explore the dynamic of customers' channel choices over time several challenges in this area still remain. However, their results are very interesting and trigger further research on this area. In the following paragraphs I present in detail each work, highlighting main results and modeling features.

Among the earlier models of customers' channel migration there is the Thomas and Sullivan (2005) work. These authors model the choice among the store, the catalog, and Internet. They use a multinomial logit model to predict channel choice as a function of: the dollar amount of direct communications that the firm spends, the price (i.e. dollar amount paid for the product), the product category (they use data on eleven basic product categories), the distance (number of miles the customer lives from the closest store), a time-varying measure that equals the current purchase occasion number of the customer and the prior channel (i.e. dummy variables that indicate the prior channel from which the customer made a purchase). This last variable incorporates channel choice dynamics, i.e. the effects of past channel experience on current choice. Their predictions help them segment customers based on channel choice and develop targeting strategies based on these segments. The authors find that marketing can affect various channel migrations differently (e.g. it can affect the choice of catalog vs. internet differently than it affects the choice of catalog versus store). One limitation is the use of the dollar amount of direct communications that the firm spends which does not vary at individual level and it does not reflect the type of marketing.

Ansari, Mela and Neslin (2008) jointly model three decisions: channel choice, whether to buy, and how much to spend and they examine the antecedents and consequences of channel migration from offline to online. They use a framework in which marketing communications determine customer behavior, in the form of channel selection (choice), purchase frequency, and order size. These behaviors are related contemporaneously and reinforced over time through "experience effects." The authors use a type-2 tobit model of purchase frequency and order size, and they integrate it with a binary probit model of choice between catalog and Internet. Specifically, they observe in each time period (i.e. month) whether the customer purchases from the firm in period *t*, if so, how much is spent and which channel is chosen. They model latent utilities that drive the observed data as functions of customer characteristics, previous behavior or experience, marketing, and seasonality/trend. They use individual random intercepts to distinguish the effect of individual persistence in preferences from the effect deriving from experience. Experience effects include variables such as expenditures in the previous period, channel choice in the previous period, etc. The authors also include cumulative web usage–since the Internet was new at the time, they wanted to investigate permanent learning that might occur. Marketing includes catalogs and emails, modeled as stock variables and interactions. Time effects include seasonality and trend. The authors found that catalog choice exhibited strong inertia, i.e., spending a lot of money on a catalog in the previous month increased the likelihood the catalog will be chosen if a purchase were made this month. They also found that Emails were associated with choosing the Internet, although with decreasing returns to scale and that Catalogs did not influence catalog choice at low levels, although did so at high levels. Finally, they observe negative association between cumulative use of the Internet and purchase incidence, suggesting that Internet purchasing may undermine customer retention.

Similarly, Knox develops a model to capture purchase frequency, order size, and channel choice. He uses a nested logit approach to capture incidence and channel choice. The main focus of Knox's work, however, is on modeling the process whereby consumers evolve to form three channel usage segments: online oriented, offline oriented, and multichannel. The channel utilities include an intercept, marketing variables with response parameters that are allowed to vary over states, in addition to a time trend and last amount of money spent in the particular channel. He includes the same marketing variables in the purchase and choice decision, this specification allows for a rich array of marketing responses. Knox models the process by which customers become loyal to certain channels, or choose to adopt certain channels (i.e. channel adoption) as a hidden Markov process. The customer is assumed to be in an (unobserved) segment at any point in time. The possible segments include the initial or "learning" segment, as well as the ultimate online, offline, and multichannel segments. Knox assumes that offline, online, and multichannel are absorbing states once customers evolve to one of those segments, they stay there. The probabilities of migrating from the learning segment depend on marketing. This is modeled using a Dirichlet distribution for each marketing instrument combination (received catalog, received email, received both, and received neither). This captures heterogeneity in transition likelihood across customers depending on what if any marketing communications are received. Knox finds strong evidence in support of evolving customer behavior. Models that do not capture this non-stationarity do not capture well the growth of the online channel over time. Knox finds that customers migrate to one of the three segments, that marketing influences the migration probabilities, and that the multichannel segment accounts for the highest sales volume. Knox findings confirm the results of Ansari et al. (2008) which found that Emails and catalogs influence channel choice and that both and catalogs have a greater effect on purchase incidence. He finds two segments in the population: a migration segment that starts offline and gradually moves online, and a hardcore offline loyal segment. Thus there does appear to be a large channel migration in effect. Marketing instruments play different roles in the channel selection process. He interestingly found that marketing has an effect on slowing down evolving customer behavior. For example, catalogs effectively slow down the transition from the migration state to the offline state. Emails have a moderate effect on slowing down customer transitions.

Finally, Venketesan et al. (2007) model the time until the adoption of a second channel, given the customer is initially using one channel, and then the time until the adoption of a third channel, given the customer is using two channels. The authors consider a retailer using a full-priced store, discount-priced store, and the Internet. The authors define $t_{ij}$, where $j$ equals either 2 or 3, as the time taken to adopt the second or third channel. The authors formulate a hazard model of $t_{ij}$ as follows. They include customer characteristics and marketing variables at customer level which occur between the adoption of the j-1th and jth channel. They test how these variables impact on the hazard probability that customer i will adopt his or her jth channel channel at time tij since the adoption of the j-1th channel (j = 2, 3). The authors found that marketing encourages faster channel adoption, although with decreasing returns. Cross-buying and purchase frequency shortens the time of channel adoption, presumably because these customers need more channels. Interestingly, the number of returns especially increases the length of time to adopt the third channel. Once customers are using two channels, they have to be shopping at least at one type of store, so their returns needs are satisfied. Finally, they found that customers adopt their second channel less quickly than their third. This suggests that once the customer learns to adopt a second channel, they have acquired the skill to shift channel.

## 2.4 Channel Choice Decision Making Process

When consumers realize that they want to make a purchase they go through a series of steps in order to make it. These steps can be described as (1) problem recognition, (2) information search, (3) evaluation of alternatives, (4) choice and, (5) post-choice evaluations. These steps represent the standard stages depicting the consumer decision-making theory (see for example Solomon, Bamossy and Askegaard, 2002). This theory can be used for studying different types of customer's decisions and, among these, of course channel choice.

In literature, search behavior (e.g. Ratchford et al., 2003; Wendel and Dellaert 2005), purchase behavior (Ansari et al., 2008; Alba et al., 1997; Fox et al., 2004; Inman et al., 2004; Venkatesan et al., 2007; Thomas and Sullivan, 2005; Knox, 2005) or both (Verhoef et al., 2007; Balasubramanian et al., 2005) have been considered in channel choice models or in channel decision theoretical frameworks.

Early analysis about customers channel choice decision making process have extensively discussed consumers decisions in the online environment (e.g. Ariely, 2000;

Bakos, 1991 and 1997; Brynjolfsson and Smith, 2000; Hoffman and Novak, 1996) and in the offline environments (see paragraph 2.2.1). However research efforts which focus specifically on consumers' use of multiple channels considering the whole channel decision making process are relatively sparse (Balasubramanian et al. 2005).

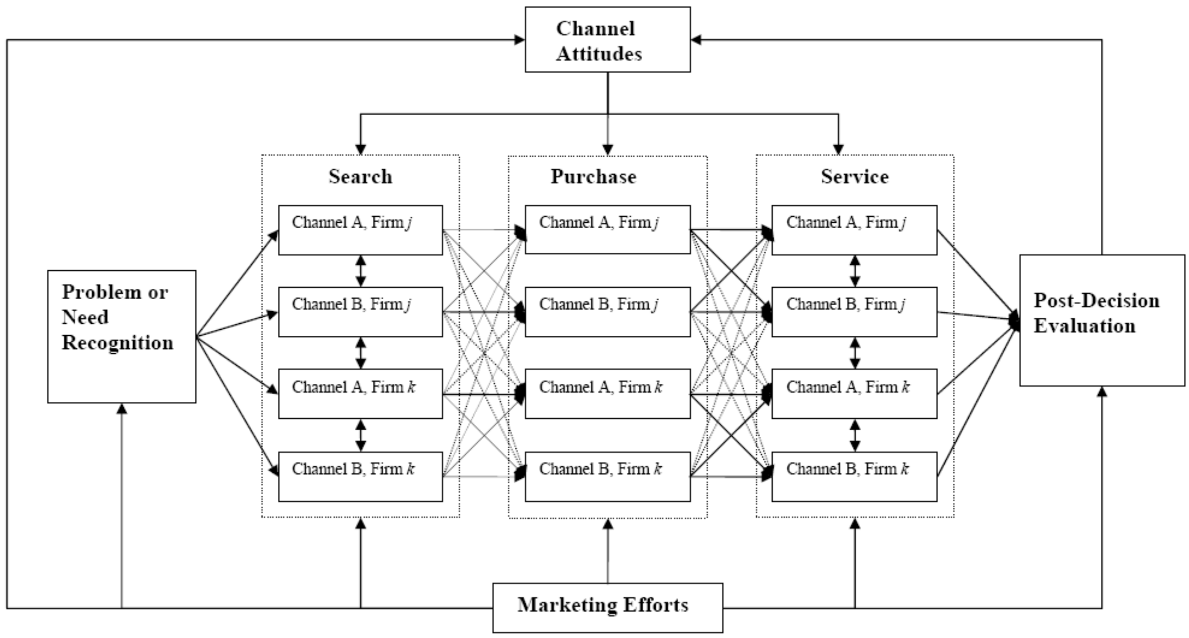
For the sake of our knowledge we can list only two works which takes into account jointly search and purchase decision in a multiple channels framework. The first contribution is the work of Balasubramanian, Raghunathan and Mahajan (2005) which presents a theoretical conceptual framework drawn on focus groups interviews with customers. Specifically, Balasubramanian et al. (2005) argue that consumer goals at various stages of the decision process are in accord with the characteristics of various channels. Hence, they emphasize the importance of acquiring knowledge on how consumers construct their goals and choose different channels at various stages of the decision process, in order to help managers to influence consumers' choice of channels and sellers. Consequently, they strongly recommend the developing of a single theory which captures consumer behavior in the multichannel environment. The second contribution, the work of Verhoef, Neslin and Vroomer (2007), represents a first effort in this direction. They empirically test a model for understanding the causes of research shopping phenomenon, i.e. the tendency of customers to use one channel for search and another for purchase. They identify three fundamental mechanisms causing research shopping: (1) attribute-driven decision making (this mechanism is based on consumer perception that one channel excels on attributes that determine search, while the other channel excels on attitudes that drive purchase), (2) lack of channel lock-in (i.e. higher attitudes toward searching on channel A translate into higher attitudes toward purchasing on channel A) and (3) cross-channel synergy (i.e. searching on channel A enhances the experience of purchasing on channel B). They found that Intenet→Store research shopping is the most common form of research shopping. This is reasonable thinking to the strong search attribute advantage of internet compared to the store, coupled with strong purchase attribute advantage of the store. Interestingly, they also found a lack of statistically significant lock-in for the internet but a very strong lock-in for store and catalog.

Neslin et al. (2006b), in their review on multichannel research efforts, underline three crucial aspects concerning the channel choice decision making process: (1) customer perceptions and preferences drive channel choices (e.g., the customer may prefer the Internet for search because it is easy to use but the store for purchase), (2) the customer learns from and evaluates his or her experiences, which feedback into the perceptions and preferences that guide his or her next shopping task (e.g., the customer may learn that the Internet search did not answer all the important questions), (3) the customer chooses both channels and firms, so from the customer perspective, this represents a two-dimensional choice.

Blattberg, Kim and Neslin (2008, p.637), taking into consideration these first literature evidences, adapted the traditional decision-making framework to the channel choice context (see figure 2.3) showing that the customer can access various channels at various companies in different stages. The process is guided by the customer's attitudes toward the various channels, by the firms' marketing efforts, and by the outcomes of previous stages in the process. Finally, the customer evaluates the experience and updates his/her attitudes. The main purpose of this framework is to describe the complexity of the channel choice process, the choice between channel and firm, and the dynamics which might occur among search, purchase and after-sales evaluations.

In addition, Blattberg et al. (2008) clarify that the search stage is divided into "learning" and "shopping." Learning is gathering information about general product attributes. Shopping is specifying exactly what product is wanted at what price. The decision making process described in figure 2.3 and the process described in the majority of the empirical channel choice studies implies that the different stages in decision-making exists and that marketing managers should carefully take into consideration these stages in developing their marketing strategies.

*Figure 2.3: A General Model of Customer Channel Choice (from Blattberg, Kim and Neslin, 2008, p. 637)*




These considerations implicitly suppose that customers go through an elaborate sequence of stages for each channel selection. In order words, these contributions are describing a channel choice decision process which requires a medium or high involvement on the part of consumers (see Balasubramanian et al., 2005; Black et al. 2002).

However, we should not forget to remark that often consumers simply do not go through this elaborate sequence for every decision. In the most common consumer behavior textbooks different types of consumer decision processes are described. When the process is very complex it is called extended problem solving (EPS), when it is characterized by a low degree of complexity it is called limited problem solving (LPS). Consumer researcher have found convenient to think in terms of a continuum where at one extreme we have habitual decision making and at the extreme EPS (Solomon et al. 2002).

Research efforts on channel choice decision making typically do not consider these different types of decision making processes. Actually, some authors recognize that often channel choice might be a behavior which is not satisfactorily described by a complex decision making process. For example, Balasubramanian et al. (2005) highlight that channel choice can be a matter of routine and that when consumers follow established shopping schemas and scripts in employing a channel they are rarely involved into complex decision rules. In recognizing this evidence, they underline that this issue remain under-researched. Similarly, Blattberg et al. (2008) emphasize that experience in using a particular channel makes it more likely the customer will use that channel in the future but, whether this is due to mindless inertia or to cognitive learning has not been explored. We will discuss more deeply this aspect in chapter 3. Here we merely aim to observe that the channel choice models and modeling strategies discussed in paragraph 2.3.2. are implicitly built upon the above described channel choice decision making process, which supposes the existence, at least at the beginning, of an extensive search phase and post-purchase evaluations, i.e. the existence of learning prone customers.

However, it should be remarked that while steps in decision-making are followed by consumers for purchases, such a process in not an accurate portrayal of many purchase decisions (Olshavsky and Granbois 1979). Often consumers simply do not go through this elaborate sequence for every decision. Hence, research literature on decision making has characterized different types of decision making processes distinguishing them in term of the amount of effort that goes into the decision each time it must be made (Solomon, Bamossy and Askegaard 2002). Sometimes consumers undertake a complex decision process but more common are rather simplistic processed in which relatively little time and effort are devoted to the decision (Olshavsky and Granbois 1979, Blackwell, Miniard and Engel 2002). Consumer researchers have found convenient to think in terms of a continuum. At one end we have habitual decision-making and at the other extreme extended problem solving.

Olshavsky and Granbois (1979) even sustain that a significant proportion of purchases may not be preceded by an extended decision process. They argue that this conclusion does not simply restate the familiar observation that purchase behavior rapidly becomes habitual, with little or no pre-purchase processes occurring after the first few purchases. They sustain that for many purchases a decision process might never occur, not even on the first purchase. Specifically, their review on shopping and suggests that extended search and evaluation typically does not precede store patronage (Granbois 1977).

Similarly, many years later Balasubramanian et al. (2005) underlined that shopping can be a matter of routine and that when consumers follow established shopping schemas and scripts in employing a channel, they rarely use alternative channels to compare costs and benefits. Often researchers assume that the consumer choice task involves some comparison of alternatives. This is appropriate when consumers do choose between competing offerings. When consumers patronize a channel because it figures in some schema, this standard decision scenario may not apply. In other words, Balasubramanian et al. (2005) show that when consumers are guided by a schema or by a shopping script, they are unlikely to employ distinct channels at the various stages of the shopping processes.

The purchase process, at least for products that require medium to high involvement on the part of consumers, consists of distinct stages (Lilien, Kotler, & Moorthy, 1992).

# Chapter 3

# CONCEPTUAL FRAMEWORK AND RESEARCH OBJECTIVES

## 3.1 Overview

This research aims to study customer channel migration process; specifically we aim to obtain insights into how the channel decision process of newly acquired customer evolves over purchase occasions. Our main purpose is to understand if customers change their channel decision process over time. We argue that when changing channel decision consumers first develop a decision strategy (*trial* stage) then they update and switch from it after a certain number of purchase occasions (*post-trial* stage). Time to switch might vary among customers.

We also contend that customers might remain their initial channel decision strategy, but when switching they change in term of customers' channel preferences, dependence upon previous channel choices, and responsiveness to the marketing with respect to the *trial* phase. Along these lines, we can delineate different types of decision making strategies and we can track the migrations that might occur among them. This leads us to take an overall view on the channel choice migration process and to ground our work within a general decision-making framework which aims to describe the evolution of channel decision behavior of new customers to the firm.

The present chapter is organized in two main parts. Each part aims to answer specific question:

5) Why do some people switch decision processes while others don't?

6) How the decision process changes over time among the people who switch?

In the first part (paragraph 3.2) we answer to the first question discussing three main issues. First, we present the Aaker (1971) new trier logic which motivates this study. Second, we discuss the role of learning in switching. Third, we present some factors which might trigger the customer's learning proneness.

In the second part (paragraph 3.3) we answer to the second question. First, we debate the existence in the channel choice context of different decision-making patterns Second, we outline four different types of channel choice decision patterns which might takes place over time: (1) always inertial channel decision strategy, (2) always preference-based channel decision strategy, (3) inertial channel decision strategy which turns out preference-based, and (4) preference-based channel decision strategy which turns out inertial. Third, using this framework we focus on the customers' marketing responsiveness and we track the evolving patterns depending on the marketing role as well.

## 3.2 Part 1: The Customer's Propensity to Switch Decision Processes

We aim to obtain insights into how the channel decision process evolves over purchase occasions. We draw on the Aaker's (1971) new-trier model, which was applied to the purchase of new products. Aaker (1971) argues that during the *trial* period the customer is essentially sampling purchase options and learning in the process. At some point, the customer transitions to his or her equilibrium decision process. Aaker's notion is that the evolution from the *trial* to *post-trial* phase takes place if customers learn something from their initial experiences that changes their decision process. For example, they may try a relatively new product and learn that they highly prefer that product. Their *post-trial* decision process would therefore reflect a high preference for that product. Of course, it may also occur that they find they dislike the new product, and their *post-trial* decision process therefore reflects a low preference for that product.

While Aaker assumes that all new consumers eventually change their initial decision process to a final decision process, it is quite plausible that not all new consumers make this transition, i.e. they may stick with their original decision process. In particular, it can be argued that when customers do not exhibit a high leaning proneness, maybe because they are not committed with the specific decision task, actually they do not learn from their initial experience; hence they do not change their initial decision process.

At this point is important to clarify that the depicted framework is not grounded on the Bayesian learning literature (e.g. Erdem and Keane 1996), in other words we do not assume customers as rational and fully-informed Bayesian updaters. For example, Erdem and Keane (1996) model both usage experience and advertising as sources of information regarding uncertain brand attributes in a structural model in which customers are suppose to continually update their brand beliefs as new information sources are available. Their work demonstrates that under uncertainty (specifically in turbulent markets) customers learn in a Bayesian fashion, in this way they give a strong structural explanation on the "why" current choices depends on past choices, hence on the why customers learn. Specifically, they demonstrate that customers are forward looking and they learn about brand attributes with usage experience and advertising. The evidence that learning induces customer to update their beliefs about choice alternative attributes supports our idea that customers might gradually change their decision process over time. However, we do not aim to estimate a structural model with perfectly rational and forward looking customers. Even if this might be an interesting issue, in this context of analysis we are only marginally interested in the way customers update their channel attributes beliefs. On the contrary, our purpose is to estimate the probability, at individual level, that two distinct stages describing customer's channel decision process over time exist, i.e. that customers might switch decision processes over time. Then, if the switching takes place we aim to estimate two channel choice models: one for the *trial* phase and another for the *post-trial* in order to assess if different parameters estimates characterized these models.

### 3.2.1 Why People Switch Decision Processes?

As suggested from the above, the reason for switching decision processes is learning. During the *trial* phase customers learn the benefits and costs of different channels, or the usefulness of various marketing communications. As a result, customers are likely to move to the *post-trial* phase when their decision process is different. It might involve more established preferences for certain channels, or paying more (or less) attention to various marketing communications.

Learning is the process by which experience leads to changes in knowledge and behavior (Blackwell et al. 2001). Such changes might be relatively permanent (Solomon et al. 2002). For this reason if a learning process occurs we can observe main changes into the customers' channel choice pattern over time.

Customers might or might not initiate a learning process. Hoch and Deighton (1989) point out that learning from experience is not a simple process of discovering objective truth. Customer learning from experience is open to influences that could be internal or external (Hoch and Deighton, 1989). These influences could facilitate learning or impede it. For example, strong inertia or strong channel preferences along with channel choice satisfaction could impede learning and therefore inhibit switching (i.e. customers do not revise their initial channel beliefs and stay with their initial decision process).

#### 3.2.1.1 Factors Influencing Learning

Literature on learning by doing, motivation to search and process information have identified several factors that might influence learning. Among these factors we can list:

- *Customers' familiarity with the domain* (Brucks 1985, Hoch and Deighton, 1989, Miyake and Norman 1979, Alba and Hutchinson 1987). Following Alba and Hutchinson (1987) we can refer to familiarity as the number of channel related experiences accumulated by the customer. In general customers with greater familiarity have a richer store of prior knowledge and more clear beliefs and expectations about experience. Additionally, relationship between familiarity and search for information may be inverted-U-shaped with both high and low familiarity leading to less external search for information (Johnson and Russo 1984).
- *Customers' motivation to learn* (Bettman 1979, Hoch and Deighton, 1989). Highly motivated customers enter more actively into the search for information and encoding more extensively than customers less motivated to learn. This motivation to learn could be related to the customers risk profile (Cox 1967, Evans et al. 1996); ability to acquire relevant information with which purchase uncertainty can be addressed (Murray, 1991, MacInnis & Jaworski, 1989) and to the desire to reduce information search costs (Shugan 1980, Blackwell et al. 2002).
- *Customers' characteristics and demographics* (Blackwell at al. 2001, Moorthy, Ratchford and Talukdar 1997).
- *Situational factors and external information.* (Hock and Deighton, 1989, Ha Hoch 1988, Blackwell et al. 2001). Customers in the market are exposed to different forms of external information (e.g. advertising, firm marketing strategies, word of mouth, etc.) or events out of their control that might have an impact on the learning process.

- *Customer satisfaction and negative experiences* (Blackwell et al. 2001). Satisfaction and dissatisfaction represent *post*-choice evaluations. If customers are satisfied with the results of their initial decision process, e.g., the channel they are currently using, they will continue to use this process. But if dissatisfying experiences happen, customers might desire to search for additional information and they could activate a learning process in order to correct the problem, and eventually change their decision process (Tschirgi 1980, Weiner 1985)

These factors can influence the learning process. Therefore, they might affect how the length of the learning phase (Hoch and Deighton, 1989) and the probability to observe a change in the customer's decision process. Basing on this, we segment the market in two groups of customers: customers with a low probability to leave their initial channel choice process (i.e. low probability to switch) and customers with a high probability of changing their decision process (i.e. high probability to switch). We call the first group "*stayers*" and the second group "*switchers*".

The factors listed above affecting learning can increase or decrease the chance of being a *stayer* or *switcher*, i.e. they may trigger the learning process or not. We do not have data that directly measures these. However, we have potential proxies as follows:

*Customers' familiarity*

- *Channel Choice Consistency*: Channel choice consistency combined with a high number of channel related experiences denotes strong channel familiarity. Customers with a strong early channel familiarity are more likely to be *stayers* (H1).

*Customers' motivation to learn*

- *Use of the Internet*: The Internet, as the newest channel, would be the channel where most customers would have limited experience. Hence we would expect initial users of the Internet were experimenting and therefore had more to learn (see Ansari, Mela, and Neslin 2008). Furthermore, there is consensus in both the academic and business press about the ability of electronic channel to convey information to consumers at lower costs than other channels (Alba et al. 1997, Bakos 1997 Abeer and Lohse 1999, Lynch and Ariely 2000). Thus, internet usage may increase customers' motivation to search for information reducing the cost of searching. Therefore we hypothesize that H2a: Early use of the Internet would be associated with switching decision processes, and H2b: Customers whose first channel used was the Internet would be more likely to switch decision processes.

- *Providing email address*: Customers providing an email address are indicating they are interested in suggestions that might come from emails. This indicates proclivity to experiment and therefore we hypothesize H3: Customer who provide emails are more likely to switch decision processes.

*Customers' characteristics and demographics*

- *Demographics*: We have information on age and gender. We are not aware of an association between gender and "venturesome" behavior. However, younger people should be more likely to experiment, yielding H4: Younger customers are more likely to be *switchers*.

*Negative Experiences*

- *Returns*: Customers who make returns early in their usage may be dissatisfied with the results of their channel decision process and therefore switch to a different decision process. We hypothesize: H5: Consumers who return products are more likely to switch decision processes.
- *Customer acquisition method*: Customers are acquired by various means. Sometimes promises are made during this process that the service does not live up

to. We therefore hypothesize: H6: Customer acquisition method should be associated with switching decision process.

*Table 3.1: Hypotheses Summary*

| HYPOTHESES | |
|---|---|
| **Why people should stay or switch?** | |
| $H_1$ | Customers with a early strong channel familiarity are more likely to be *stayers* |
| $H_{2a}$ | Early use of the Internet would be associated with switching decision processes |
| $H_{2b}$ | Customers whose first channel used was the Internet would be more likely to switch decision processes |
| $H_3$ | Customer who provide emails are more likely to switch decision processes |
| $H_4$ | Younger customers are more likely to be *switchers*. |
| $H_5$ | Consumers who return products are more likely to switch decision processes. |
| $H_6$ | Customer acquisition method should be associated with switching decision process. |

## 3.3 Part 2: How Choice Decision Patterns Might Evolve

Previous works studied customers channel choice process (see paragraph 2.4 in chapter 2). The customer recognizes a need, searches for information for a product that addresses the need, purchases the product, and then seeks after-sales service. Along the way, the customer can access various channels at various companies. The process is guided by the customer's attitudes toward the various channels, the firms' marketing efforts, and the outcomes of previous stages in the process. Finally, the customer evaluates the experience and updates his/her attitudes (Blattberg, Kim and Neslin 2008 p. 637).

This theoretical framework represents a situation of extended problem solving (EPS) which perfectly suits the channel choice behavior of customers who are learning prone and who are willing to spend energy and time in order to search for information about channel alternatives. However, Olshavsky and Granbois (1979) remarked that while steps in decision-making are followed by consumers for purchases, such a process in not an accurate portrayal of many purchase decisions (Olshavsky and Granbois 1979). Often consumers simply do not go through this elaborate sequence for every decision.

The type of decision process used in different choice situations depends on customer's motivation to process information. Previous research in marketing conceptualized motivation to process information in terms of consumer's involvement and commitment with the informational stimuli (Bloch and Richins 1983, Burnkrant and Sawyer 1983, Cohen 1983, Greenwald and Leavitt 1984, Houston and Rothschild 1978, Lastovika and Gardner 1979, Mitchell 1981, Petty and Cacioppo 1981, Wright 1974, Zaichkowsky 1985). Involvement is defined as the level of perceived personal importance and interest evoked by a stimulus within a specific situation (Blackwell et al. 2001). Therefore, the more important the product or service to a customer, the more motivated he or she is to search and be involved in the decision.

Hence, literature on decision making has characterized different types of decision making processes distinguishing them in term of the amount of effort that goes into the decision each time it must be made (Solomon, Bamossy and Askegaard, 2002; Olshavsky and Granbois, 1979; Blackwell, Miniard and Engel 2002). We can think of a continuum from one extreme where there are, limited problem solvers (LPS), and the other extreme extended problem solvers ones (EPS). Actually, Solomon et al. (2002 second edition, p.237) distinguish between routine and habitual decision makers and limited problem solvers and they put at one extreme routinized behaviors, then LPS and finally EPS. They argue that both extended and limited problem solving involves some degree of information search and learning proneness, and that routinized decisions are made with little or no conscious effort. The same concept is expressed by Alba and Hutchinson (1998) and Kujala and Johnson (1993) who explain that choices characterized by automaticity are performed with minimal effort and without conscious control. Olshavsky and Granbois (1979) even remark that for many purchases this type of decision behavior might occurs from the first purchase. They argue that this conclusion does not simply restate the familiar observation that purchase behavior rapidly becomes habitual, with little or no pre-purchase processes occurring after the first few purchases, by contrast this mean that "unconscious" choices might occurs since the first purchase ever made. Their argument is rather "strong" and it might be questionable, however literature demonstrates that in many situations consumers use simplified heuristics to make their choices and sometimes simplified decision processes might be observed since the first purchase. The notion of peripheral route processing (Petty and Cacioppo, 1981) represents another theory in the literature which supports the extensive use of simplified heuristics by customers. Several works (e.g. Inman, McAlister and Hoyer, 1990) demonstrate that it is important to be aware that customers in many situations use simplified, and even unconscious, decision processes in order to delineate effective marketing strategies. For example, the advertising or promotion effectiveness, strongly depend on the type of decision process which characterized customer behavior.

Accordingly, we can envisage channel choice situations in which customers use simplified heuristics in order to select channels. We find support for this idea in the channel choice literature. Granbois (1977) suggests that extended search and evaluation typically does not precede store patronage. Balasubramanian et al. (2005) underlined that shopping can be a matter of routine and that when consumers follow established shopping schemas and scripts in employing a channel, they rarely use alternative channels to compare costs and benefits. In addition, these authors argue that when consumers patronize a channel because it figures in some schema, the steps which characterize the standard decision-making scenario may not occur.

Despite this evidence almost all channel choice studies consider situations in which learning takes place, implicitly describing Extended Problem Solving (EPS) decision making. For example, Balasubramanian et al. (2005) recognize that the purchase process which they describe requires minimum of high involvement on the part of consumers. Similarly, Black et al. (2002) briefly reflect on the involvement towards the decision process in their work. Similarly, channel migration models implicitly consider EPS a necessary precondition for their channel choice models. For example, Knox (2005) contends that if a customer is new to the firm the customer channel choice decision process starts with a learning phase, i.e. he argues that a learning phase takes place at the beginning of the relationship with the firms and he asserts that customers who are new to the firm are likely to be learning. This is, of course, plausible, but we wish to underline that new customers are not necessarily more likely to be learning prone, in other words an EPS decision making for newly acquired customers is not "guaranteed". By contrast, customers may starts their relationship with the firm using simplified decision making rules, subsequently some events (e.g. negative experiences, the firm marketing stimuli, etc.) might trigger customers' motivation and commitment with the choice task and induce them to change decision process or customers may simply keep on a straightforward decision process over time.

These considerations lead us to take a larger view on the channel choice migration process and to ground our work taking into consideration a more general decision-making framework describing the evolution of channel choice behavior of new customers to the firm. Therefore we suppose that different types of decision processes might take place.

### 3.3.1 How Do Switchers Change?

In the previous sections we aimed to stress an important concept: in the study of the evolution of customers' channel migration process we should take into consideration that customer might use different decision making strategies over time and that these strategies can range from habitual decision making to EPS. We also highlighted that channel migration models literature is implicitly based on EPS, therefore authors have outlined managerial implications mainly based on committed, involved and learning prone customers without formally taking into account that different decision processes might describe channel customers' behavior. This, of course, has implication in the evolution of customer channel choice patterns over time.

Here, we argue that we can think to, at least, two types of customers: those committed in channel choice task who consciously choose the channel that they prefer, and those relying their channel decisions on the previous channel chosen, and do not exhibit a strong commitment in this choice task. This evidence is well-supported in the choice modeling literature where researchers have made efforts in order to distinguish different sources of choice persistence over time. For example, Keane (1997) argues that customers' exhibit persistence in brand choice in two diametrically opposed patterns of consumer behavior. The former takes places simply because customers have different preferences over choice alternatives (heterogeneity in customers' preferences). The latter is due to positive state dependence. State dependence is defined as the dependence of the current customers' choice on previous choices made (Heckman 1981), in particular an high and positive state dependence is defined as inertia (Seetharaman et al. 1999). Inertia depicts a situation where, for example, a brand is bought merely because less effort is required, but if another alternative is introduced which for some reason is easier to buy (e.g. it is cheaper, or the usually chosen alternative is not available) the customer will not hesitate to choose it since there is little or no underlying commitment with the choice alternative (Solomon et al. 2002, p. 259). Keane (1997) stressed that it is of fundamental importance distinguishing between choice persistence driven by inertia or by customers' "conscious" preferences because this has important managerial implications.

Basing on these evidences we argue that in channel choice situations, as well, we can envisage, at least, two different types of decision making processes going on. One is driven by inertia which means less commitment, involvement and so on. The other is driven by preferences, which simply means "conscious" choices and of course commitment (e.g. I choose the channel that I prefer, and I know that I prefer it because I have experiment it in the past or I have a strong aversion for other channels, e.g. catalog and internet which do not have contact personnel). These considerations lead us to the definition of four possible patterns describing the channel choice process (see table 3.2)

*Table 3.2: Channel Decision Patterns*

| | **Final Channel Decision Process preferences>inertia** | **FinalChannel Decision Process preferences<inertia** |
|---|---|---|
| **Initial Channel Decision Process preferences >inertia** | *Conscious Stayers:* Customer are committed with the choice task, they exhibit strong conscious preferences since the beginning, and stated that way | *Switchers b (preference-based$\rightarrow$inertial):* For those customers we can distinguish two stages: *trial* and *post-trial*. Customer started with conscious channel preferences but over time they revise their decision process and became inertial. |
| **Initial Channel Decision Process preferences <inertia** | *Switchers a (inertial$\rightarrow$preference-based):* For those customers we can distinguish two stages: *trial* and *post-trial*. Customer started low involved with the choice task, but over time they revise their decision process, and they exhibit conscious channel preferences. | *Inertial Stayers:* Customer act as not interested in the channel choice task, they choose relying on the previous choice and stated that way |

*"Conscious" Stayers* are committed with the channel choice task and exhibit conscious channel preferences since the beginning of their relationship with the firm. This might happen for several reasons (e.g. strong aversion for some channel alternatives, for example the Internet, channel expertise developed with other companies, and so on) Their conscious preferences leads them to perform a persistent and habitual channel choice pattern over time. For those customers we do not observe a switching behavior, probably because they are satisfied with the channel/channels chosen and they are not motivated to undertake an EPS.

*Switchers (pattern a)* start inertial with the channel choice task but as time passed some factors might trigger their inner propensity to search and process information, thus they get involved and motivated with the channel choice task and, after a EPS or learning phase, they develop channel preferences and they exhibit conscious channel preferences in the *post-trial* stage.

*Switchers (pattern b)* exhibit at the beginning quite developed initial channel preferences and commitment during a *trial* stage; they exhibit since the beginning a limited or even extended problem solving about channel alternatives. However, in the *post trial* they do end up show strong channel preferences and they might even be inertial in channel selection. This could happen because they do not achieve the development of "conscious" and well established channel preferences, on the contrary they might become bored or over-taxed with having to think so much about channel decisions, and so resort to an easy heuristic – choose what I chose last time, i.e., they become routinized, more inertial.

*Inertial Stayers* are extremely uninvolved with the choice task. They show strong inertia in their channel decision process. Habits are built on inertia. These customers seem to have no reasons to switch decision process, i.e. to switch towards a preference-based decision process. They select a channel basing on their previous choice without developing channel preferences.

### 3.3.2 The Role of Marketing

In the previous paragraph we highlighted that: i) customers switch when they learn, ii) several factors might foster customer learning, and induce customers to undertake an EPS, iii) two types of decision processes might describe channel choice: preference-based or inertial.

We stressed that new-to-the-firm customers do not necessarily exhibit a high learning propensity from the beginning and that an EPS might even never occur, hence we envisaged a taxonomy which depicts different channel choice decision patterns basing on switching behavior and the preference-based versus inertial decision processes. Another important issue that has not been discussed relates to the role of marketing on channel decision process. Implicitly we ask: when and where marketers can exert leverage on the channel choice process?

In order to answer this question we develop a general framework which takes into account that marketing indifferently might be effective or not in preference-based or inertial situations.

Marketing literature has demonstrates that marketing communications can be processed as information or they can serve just as cue to reinforce choice (Vakratsas & Ambler 1999, Hoch and Deighton 1989). For this reason, high marketing could explains indifferently the behavior of an involved customers who is learning prone and uses marketing as a source of external information or the behavior of an inertial decision maker who chooses a simple decision making process (Vakratsas & Ambler 1999).

Low marketing responsiveness can mean either the customer is inertial or he has well-developed channel preferences. In the former case customer pays not attention to the channel choice and he/she does not process marketing information. In the latter case he or she has strong, well thought preferences, and he/she is less influenced by marketing. Nevertheless, it is important to explore the role of marketing on the choice process either when customers exhibit low or high responsiveness.

Basing on our considerations, we depict four possible conditions which discriminate four different customers' decision styles:

1) pattern 1: preference-based decision making – high marketing responsiveness,
2) pattern 2: preference-based decision making – low marketing responsiveness
3) pattern 3: inertial decision making – high marketing responsiveness,
4) pattern 4: inertial decision making – low marketing responsiveness.

The four decision styles describing the channel selection process present the following characteristics:

- *Customers in pattern one*. These customers are committed with the choice task and they exhibit channel preferences. Their behavior is not routinized and the marketing has a positive effect. Marketing may serve as a source of information. These customers might need marketing information to reinforce their channel preferences or to further develop them (Smith and Swinyard 1988, etc.).
- *Customers in pattern two*. These customers are committed in the channel choice decision, as well as customers in pattern one. However, marketing does not seem to influence their channel choice. This result is plausible because the effect of marketing is low when preferences are already formed and customers exhibit strong familiarity with the choice (Hoch and Deighton 1989).
- *Customers in pattern three*. These customers are uninvolved with the channel choice task. Inertia plays a significant role in explaining uninvolved customer choices, but marketing has a positive impact on channel choice. For example, we can think of a consumption situation in which the customer does not pay attention to the channel choice, he or she may desire to simplify this decision task. The positive effect of marketing in this situation is reasonable because it may serves as cue to reinforce channel choice.
- *Customers in pattern four*. These customers are uncommitted with the channel choice task. They act as strongly uninterested in channel choice, therefore they do not pay

attention to marketing information. Their channel choices are completely guided by channel state dependence, i.e. by previous channel choices.

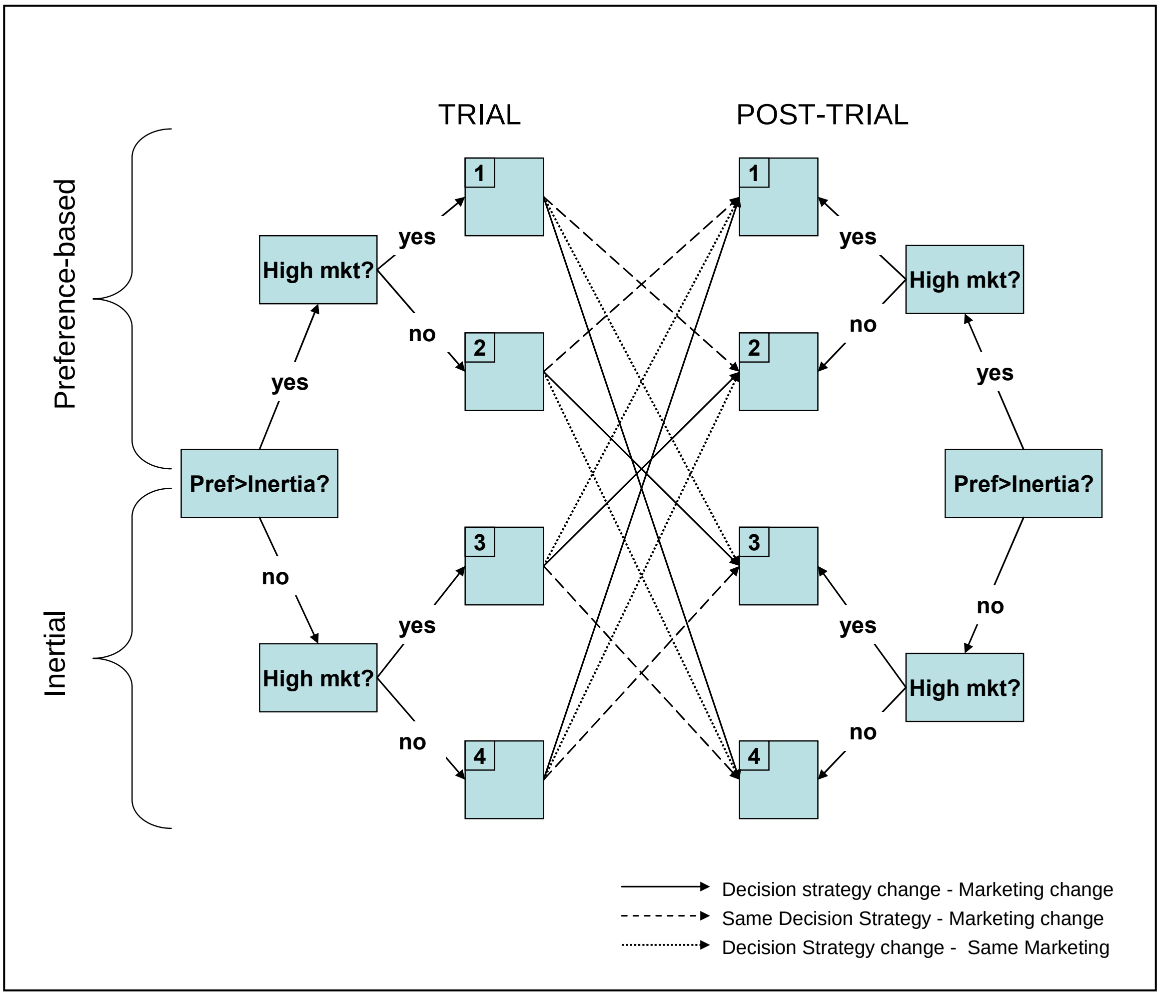
We argue that these four patters and their possible evolutions can describe *stayers* and *switchers*' channel choice migration process. S*tayers* behavior might be describe indifferently by one of these four patterns. Essentially *stayers* do not change their decision pattern over time, thus they might use one of the four patterns and they will continue stay with it over time. By contrast for *switchers* we can observe evolving patters between *trial* and *post-trial* stages. In figure 3.1 we map all the possible evolving patterns which might take place for *switchers*[2].

Specifically, we map twelve possible *trial* / *post-trial* combinations (see arrows in figure 3.1). *Switchers* can go from preference-based to inertial based decision strategy, preference-based to inertial decision strategy, or stay in their current decision strategy but just change their use of marketing.

The preference>inertia node distinguishes between preference-based and inertial decision strategy. Specifically, our purpose is to estimate a channel choice model which might give us information on channel preferences and inertial behavior, in other word we aim to estimates individual level parameters which allow us to asses the probability that the customers behavior is driven by inertia or by channel preferences (see model development in chapter 4 for details). In this way we can compare the probability that the customer is using an inertial decision strategy or a preference-based decision strategy. If the probability that he or she is inertial is greater than the probability that he or she is choosing basing on "conscious" channel preferences, we classify this customer in pattern three or four (see figure 3.1). Following the same logic, i.e. basing on the estimation of the marketing responsiveness we will classifies customers as high or low responsive to marketing stimuli.

[2] We do not map combinations 1=>1, 2=>2, 3=>3 and 4=>4 because these represent *stayers* combinations and for *stayers* we can not distinguish two stages (*trial* and *post-trial*).

*Figure 3.1: Migrations from Trail to Post-Trail*




As Keane (1997) point out a cost/benefit analysis of the marketing strategies will depend critically on the assumed forms of heterogeneity and positive state dependence, i.e. on the different types of decision strategies used in the considered population. Similarly, we believe that accounting for the role of marketing and distinguishing its effect among different types of *stayers* and *switchers* customers, i.e. among different types of decision making strategies and the evolution of these strategies over time is extremely important in developing specific channel marketing strategies. We believe that the general framework represented in figure 3.1 which takes into account that marketing indifferently might be effective or not in inertial or preference-based situations could help manager in targeting effective marketing strategies.

*Chapter 4*

# MODEL DEVELOPMENT

## 4.1 Overview

Our modeling logic is centered on the customers channel migration process. In channel choice context migration can be thought of simply as channel choice, but this expression is used to convey that there is a particularly interest on how this choice process takes places over time (Blattberg, Kim and Neslin, 2008, p. 647).

A deeper understanding of the channel migration process can help managers design marketing programs that evolve over time and evaluate the profitability of different customers in terms of channel choice behavior.

In the literature review section (chapter 2) we highlighted that a limited effort has been directed to investigate and formally model how customers' decision process changes over time as they learn their preferences and become familiar with the firm's marketing activities. Therefore, our purpose is to formally model the adoption and the channel migration process of a cohort of new customers. Specifically, we aim to obtain insights into how this channel choice process evolves over purchase occasions.

We draw on the Aaker's (1971) new-trier model. Aaker (1971) argues that during the *trial* period the customer is essentially sampling purchase options and learning in the process. At some point, the customer transitions to his or her equilibrium decision process. Aaker's notion is that the evolution from the *trial* to *post-trial* phase takes place if customers learn something from their initial experiences that changes their decision process, i.e. the new-trier purchase precipitated a learning experience (Aaker, 1971 p. 441). For example, customers may try a relatively new channel and learn their preference for it over time. Their *post-trial* condition will, therefore, reflect a high preference for that channel. Of course, they might find they dislike the new channel, as a result their post-channel decision process will reflect a low preference for it. While Aaker (1971) assumes that all new consumers eventually move from their initial decision process into a new decision process, it is quite plausible that not all new consumers make this transition and they stick to their original decision process. This might occur if customers do not learn anything additional from their initial purchases. This is in line with the decision making literature stating that under some choice conditions customers do not revise their initial choice beliefs and remain into their initial decision (see, for example, Solomon et al. 2001 p.235). For example a strong inertial behavior or strong channel preferences along with channel choice satisfaction might not foster learning and therefore switching.

Basing on this literature, we hypothesize the existence of a “staying” and a “switching” behavior. The former assumes that customers are not learning prone, so they remain with their initial channel decision over time. The latter considers that customers are learning prone and for this reason they go through two distinct stages over their purchase occasion histories with the company: a *trial* stage and a *post-trial* stage.

If we observe a switching behavior it is reasonable to argue that not all the customers require the same time to learn. The literature on learning by doing supports this contentium identifying conditions under which learning might be fast, long or difficult to initiate (Hoch and Deighton 1989). Accordingly, we aim to understand whether the length of a *trial* phase is homogeneous among customers or not. In other words we try to investigate whether differences in the length of the *trial* among customers exists.

For this reason, we use a modeling approach which allows us: first to estimate the probability that the customer is learning prone (i.e. the probability the customer switches to the *post-trial* model) at individual level, secondly to estimate how many *trial* purchase occasions he or she needs to go thought before switching to the *post-trial* stage (i.e. the probability the customer is using the *post-trial* model in the $n^{th}$ purchase occasion).

We contend that the *trial* and *post-trial* stages are governed by a different set of parameters. Specifically, we seek to capture, at individual level, the impact of direct marketing communications (catalogs and emails), state dependence and intrinsic preferences on channel migration process.

Before discussing each component of the model in details I summarize below the main behavioral assumptions behind the modeling approach:

- New to the company customers show different learning propensities. Some customers initiate a learning process that leads them to an equilibrium channel choice pattern. For these customers we can distinguish two stages in their channel choice process: a *trial* and a *post-trial* stage. Other customers exhibit a non-learning propensity. For this reason their initial channel choice model will explain their future channel decisions.
- The length of the *trial* period is heterogeneous among customers who are learning prone.
- The *trial* stage and the *post-trial* stage (if present) are described by a distinct set of parameters which govern the channel choice probabilities.
- Customers respond differently to marketing stimuli. The influence of marketing variables might evolve over time and it varies among customers and between the *trial* and the post *trial* stages.
- Customers are heterogeneous in their channel preferences and propensities. Channel preferences might evolve over time therefore they might differ in the *trial* and the post *trial* stages.

- Customers might exhibit different degrees of persistence in the channel choice over time because of an inertial behavior. The degree of this persistence might evolve over time and it might be strongly different between the *trial* and the post *trial* stages.

## 4.2 Alternative Modeling Approaches

The observed channel choice of the individual is denoted by a discrete variable which can take values 1,2,….., J. Discrete choice models' selection depends on three main issues:

1) the object of the choice and sets of the alternatives available to decision makers,
2) the type of explanatory variables considered,
3) assumptions and axioms concerning the selection probabilities ("selection probability axioms" Louviere et al. 2000),

The first issue is about the type of decision makers (e.g. households) and the number and types of choice alternatives. To fit within a discrete choice framework the choice set needs to exhibit three characteristics (Train, 2002): i) the choice alternatives must be mutually exclusive, ii) the choice set must be exhaustive, iii) the number of alternatives must be finite. Therefore, before taking into consideration the different modeling alternatives, one should critically reflect on the type of decision maker and on the specific choices involved in the study. This has important implication in the selection of the most appropriate modeling approach. For example, in the context of brand choice literature researchers have made efforts in order to understand whether one should include the no-purchase option as another choice alternative in a consumption situation. The decision to model the incidence decision (i.e. buy or not buy) brings to quite different modeling strategies depending on the role the no-purchase has in the specific research situation.

By far the easiest and most widely used discrete choice model is logit. Its popularity is due to the fact that the formula for the choice probabilities takes a closed form and it is readily interpretable (Train 2002). With reference to the second issue different versions of logit model can be obtained depending on the type of explanatory variable included. In general, we can envisage three types of explanatory variables (Franses and Paap 2005):

1) variables that are different across individuals but are the same across categories (e.g. age),
2) variables that are different for each individual and are also different across categories,
3) variables that are the same for each individual but different across categories

Depending on the type of explanatory variables included we can use the classical multinomial logit or the conditional multinomial logit.

Finally, the third issue concerns assumptions made on selection probabilities. In order to address this issue we should think about the logic behind random utility models (Thurstone, 1927, Marschak, 1960 and Luce, 1959). Discrete choice models are usually derived under an assumption of utility-maximizing behavior. The household would obtain a certain level of utility from each channel alternative. This utility is known to the household but not by the researcher. The household *h* chooses the channel alternative that provides the greatest utility. The behavioral model is therefore: choose channel *j* if and only $U_{hj} > U_{hi}$ where $j \neq i$. The researcher does not observe the household's utility but he/she observes the stated choice, some attributes of the alternatives, and some attributes of the household. The analyst can specify a function that relates these observed factors to the household's utility. These factors describe the deterministic or representative part of the utility function. But there are aspects of utility that the researcher does not or cannot observe, thus utility includes terms which capture the factors that affect utility but are not included in the representative part. The error terms characteristics (i.e. distribution, assumptions, etc.) depend critically on the researcher's specification of the representative part of the utility (Train 2002). In summary, the idea behind random utility theory is that the customer might have a perfect discrimination capacity but the researcher having incomplete information, should account for uncertainty (Shugan, 2006). For these reasons, different choice models are obtained from different assumptions about the distribution of the unobserved portion of utility and from assumptions about the characteristics of choice probabilities, namely the independence from irrelevant alternatives (IIA)[3]. IIA assumption is related on the beliefs about the structure of the error terms which better describes the choice probabilities. Actually, the origin of the IIA propriety is the assumption that the error terms of the utilities equations are uncorrelated and that they have the same variance across alternatives. For these reasons we have different modeling options about an unordered multinomial dependent variable depending on the assumption concerning the error terms. Specifically, the multinomial logit model assumes that the errors terms are distributed as a Weibull (i.e. independent and uncorrelated errors). More flexible models relax some conditions about the error terms, e.g. they allow some sort of correlations among the errors terms, relaxing at the same time the IIA propriety[4]. For example, the multinomial probit model hypothesizes a multivariate normal distribution for the error terms which allows the

[3] IIA propriety states that if the ratio of probabilities to choose alternative *j* versus *i* does not depend on any alternatives other than *j* and *i*. the relative odds of choosing *j* over *i* are the same no matter what other alternatives are available or what the attributes of the other alternatives are. Since the ratio is independent from alternatives other than *j* and *i*, it is said to be independent from irrelevant alternatives (Train, 2002). Since probabilities sum to one over alternatives, an increase in the probability of one alternative necessarily means a decrease in probability for other alternatives. The pattern of substitution among alternatives has important implications in many situations.

[4] In the context of the random utility model the IIA assumption comes about because the errors terms are assumed to be independent (i.e. Weibull random variables).

definition of a variance covariance matrix, i.e. it takes into account the existence of possible relationships among the error terms. Other extensions of the pure logit model have been developed (Maddala 1983, Amemiya 1985, Ben-Akiva dn Lerman 1985) in order to cope with the IIA propriety. A popular extension is the nested logit which assumes that choice alternatives can be divided into clusters such that the variances of the error terms are the same within each cluster but different across clusters. This implies that IIA assumption holds within each cluster but not across clusters (Franses and Paap, 2005).

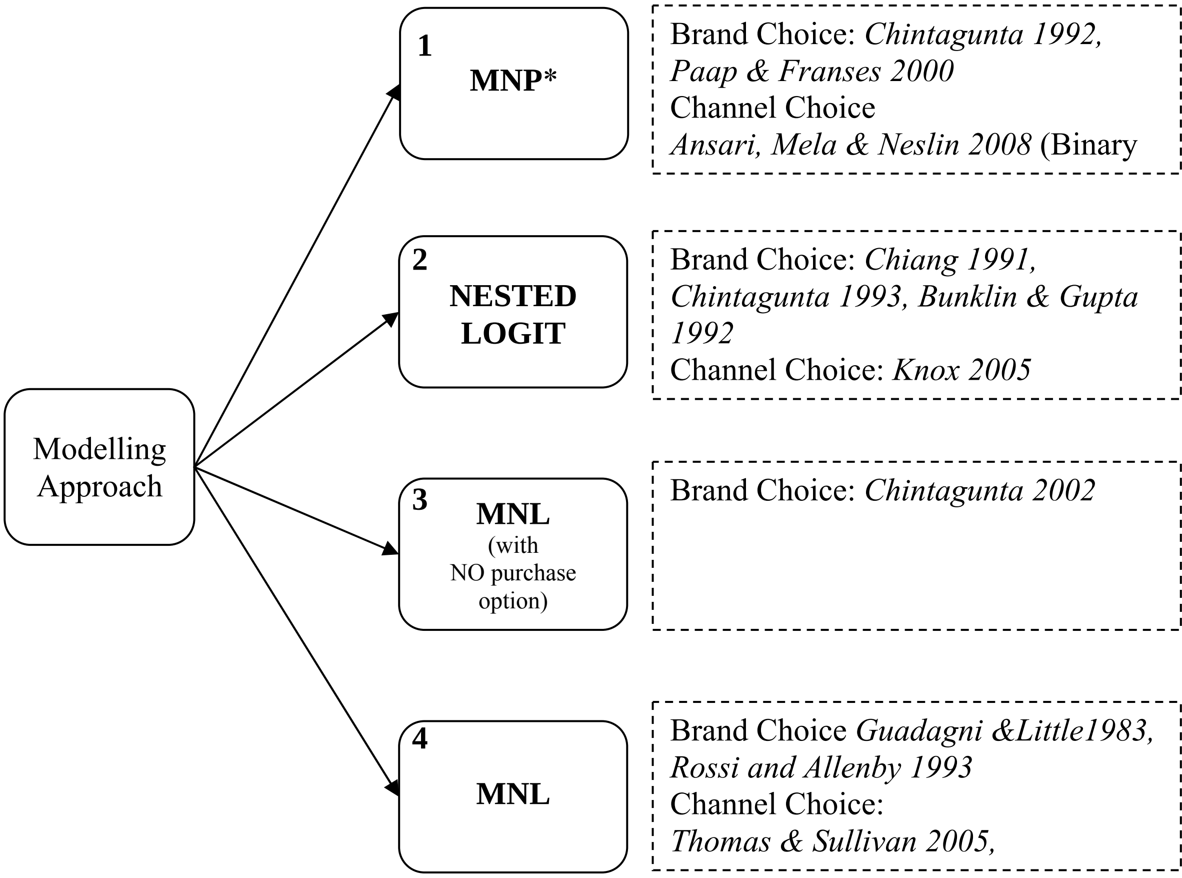
The channel choice literature borrows from the brand choice literature in modeling channel stated preferences. Often models developed for consumer scanner data are being adapted to study customer channel migration (Blattberg, Kim and Neslin 2008, p.647). Analogous to the brand choice / purchase incidence / purchase quantity (e.g., Bell, Chiang, and Padmanabhan 1999), we now have channel choice / purchase frequency / order size.

In a channel choice situation the household $h$ ($h$=1,…., $H$) observed over $t$ ($t$=1,2,….$T_h$) purchase occasions, either purchases or does not purchase and to make the purchase the household selects one of $J$ channel of a firm. We observe two outcome variable $y_{ht}$* which takes 1 if the household h purchases at time $t$ and the variable $y_{ht}$ which takes the value $j$ ($j$=1,2,…,$J$) if the household select channel $j$ at time $t$.

In literature different modeling approaches have been proposed to handle brand choice, most of these approaches have been used to model channel choice as well. Figure 4.1 summarizes the most commonly used approaches, i.e. multinomial probit (MNP), nested logit, multinomial logit with no purchase option, multinomial logit (MNL).

Each modeling approach has its advantages and disadvantages. Models developed for brand choice are being adapted to study customer channel migration process. Figure 4.1 shows several references about empirical application in brand choice and channel migration literature distinguishing among different modeling approaches. The purpose of this graphical representation is to show that each approach certainly has strong support in the literature. The contribution of this work doesn't hinge on nested logit versus non-purchase option logit or probit. We're just looking for a model that has support in the literature. Therefore, to model channel choice we passed thorough all these different alternative approaches in order to find out the one which best suit our purposes.

*Figure 4.1: Alternative Modeling Approaches*




* *In order to handle the no purchase option several methods have been applied: MNP with no purchase option, a system of equations (channel choice and incidence) specifying the random effects to be correlated both within and across equations, etc.*

### 4.2.1 Multinomial Probit

We start considering the probit as a good potential candidate for our model. The probit model is certainly the most appealing modeling approach. The appeal of the multinomial probit relies on the relaxation of the IIA propriety. In addition, the multinomial probit handles the problem of correlated unobserved factors over time.

The only limitation of probit models is that they require normal distributions for all unobserved components of utility. This is not a disadvantage per se because it allows for correlations between the errors variables and for different variances for different alternatives[5], but sometimes this implies huge estimation efforts.

Taking into consideration these advantages and disadvantages we attempted to perform a multinomial probit model with no purchase option. Our model assumes customers decide each period whether to buy from channel $j$ ($j$=0,1,…$J$) or not buy ($j$=0)[6]. The decision is governed by a multinomial probit model (with no-purchase option). In addition, we assume that the customer $h$ ($h$= 1,…., $H$) starts off using one probit model, and then switches to another[7]. The first probit represents the decision process while the customer is learning. The second probit represents the *post-trial* decision process. The time it takes for the customer to switch from the first probit (the *trial* phase) to the second probit (the *post-trial* phase) is governed by a binomial probit model ($Z_{ht}$). $U_{hjt}$ represents the utility that household $h$ derives from choosing alternative $j$ in period $t$ ($t$=1,2,…$T$):

[5] Note that when the covariance matrix is an identity matrix, the IIA propriety will again hold (Franses and Paap 2005).

[6] see Sriram and Kalwani, Management Science 53(1) 2007 p. 48.

[7] Here we anticipate the switching modeling strategy used. A more detailed discussion about the *trial* and *post-trial* switching modeling strategy is in the paragraph 4.3

$$Z^*_{ht} = \delta_{h0} + \delta_{h1} t + \eta_{ht}$$

$$if \quad \begin{matrix} Z^*_{ht} > 0 & \mathrm{Z} = 1\,(\text{post-trial}) \\ Z^*_{ht} < 0 & \mathrm{Z} = 0 \quad (\text{trial}) \end{matrix} \tag{1}$$

$$\begin{aligned} & z = 0 \\ & U_{hjt} = \beta_{hjo} + \beta_{hj1} CS_{ht} + \beta_{hj2} ES_{ht} + \beta_{h3} LC_{ht} + \varepsilon_{hjt} \quad {}_{j=1,2,\ldots,J} \\ & U_{hjt} = \varepsilon_{hjt} \quad {}_{j=0} \\ & \\ & z = 1 \\ & U'_{hjt} = \beta'_{hjo} + \beta'_{hj1} CS_{ht} + \beta'_{hj2} ES_{ht} + \beta'_{h3} LC_{ht} + \varepsilon'_{hjt} \quad {}_{j=1,2,\ldots,J} \\ & U'_{hjt} = \varepsilon'_{hjt} \quad {}_{j=0} \end{aligned} \tag{2}$$

The errors will be assumed to follow a multivariate normal distribution:

$$\Omega = \begin{pmatrix} \varepsilon_1 \\ . \\ . \\ . \\ \eta \end{pmatrix} \sim MVN(0, \Sigma)$$

We made assumptions on the possible correlations structure which better describes the relationships among the error terms. Matrix Σ contains these assumptions. Specifically, we hypothesize that: i) the errors terms of the no purchase option alternative utility (i.e.$\varepsilon_0$) are correlated with the channel utilities equations errors, ii) the errors terms of the *trial* equations are correlated with the same error terms of the *post-trial* equations, iii) different variances (heteroschedasticity). These assumptions are summarized in the variance covariance matrix of the error terms below.

$$\Sigma = \begin{array}{cc} & \begin{array}{ccccccccc} \varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \varepsilon_0 & \varepsilon'_1 & \varepsilon'_2 & \varepsilon'_3 & \varepsilon'_0 & \eta \end{array} \\ \begin{array}{c} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_0 \\ \varepsilon'_1 \\ \varepsilon'_2 \\ \varepsilon'_3 \\ \varepsilon'_0 \\ \eta \end{array} & \left[\begin{array}{ccccccccc} \sigma_1^2 & & & & & & & & \\ 0 & \sigma_2^2 & & & & & & & \\ 0 & 0 & \sigma_3^2 & & & & & & \\ \rho_{01} & \rho_{02} & \rho_{03} & \sigma_4^2 & & & & & \\ \rho_1 & 0 & 0 & 0 & \sigma_5^2 & & & & \\ 0 & \rho_2 & 0 & 0 & 0 & \sigma_6^2 & & & \\ 0 & 0 & \rho_3 & 0 & 0 & 0 & \sigma_7^2 & & \\ \rho_{0'1} & \rho_{0'2} & \rho_{0'3} & \rho_0 & \rho_{0'1'} & \rho_{0'2'} & \rho_{0'2'} & \sigma_8^2 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_9^2 \end{array}\right] \end{array} \tag{3}$$

#### 4.2.1.1 Problems with this Specification

In some situations, normal distributions provide an adequate representation of the random components and it might be easy to estimate. However, in many situations if the number of alternative *J* or the number of observations (H) is large numerical integration is no longer feasible because the number of function evaluations becomes too large (Franses and Paap 2005). For this reason, in many brand choice models researchers avoid introducing correlations between the brand utilities because correlations are difficult to identify when J is large (Chib, Seetharaman, and Strijnev, 2004). Specifically, in this case the number of observation is large (*H* is large and *T* is large as well). Furthermore, we have individual level parameters. This means that we should estimate *H* variance covariance matrixes for each time period *t*. This compromises the estimation feasibility[8].

[8] We tried this solution with a subset of our sample and even a simplified version which did not consider different stages, but even with this reduced sample size and simplified form we had identification problems with the estimation of the variance covariance matrix (see appendix 1a for the syntax used).

### 4.2.2 Nested Logit

The nested logit assumes that choice alternatives can be divided into clusters. The IIA assumption is relaxed across clusters. Specifically, we create two clusters: the incidence cluster (decision to purchase) and channel choice cluster.

This model of channel choice behavior considers two households decisions:

1) Whether or not to buy an item (incidence model)

2) Which channel use (multinomial logit model)

We distinguish between two different incidence and multinomial logit models: one for the learning phase and another for the *post-trial* phase. We model the length of the learning phase using a geometric distribution (see paragraph 4.3 for details). Following, we distinguish between two different incidence models. One for the *trial* phase ($Z_0$) and another for the *post-trial* phase ($Z_1$).

$$Z_{0ht} = \alpha_{0h} + \lambda_{0h} IV_{0ht} \tag{4}$$

$$Z_{1ht} = \alpha_{1h} + \lambda_{1h} IV_{1ht} \tag{5}$$

We allow for heterogeneity with individual-level intercepts ($\alpha_0$ and $\alpha_1$). IV indicates the inclusive value. It can be seen as the expected maximum utility from the channel choice decision (Louviere, Hensher and Swait, 2000; Herriges and Kling, 1996; Ben-Akiva and Lerman, 1985). We allow for heterogeneity in the inclusive value parameters ($\lambda_0$ and $\lambda_1$) as well. After the incidence decision customer *h* will choose a channel *j* (*j*=1,2,…,J).

$U_{hjt}$ represents the utility that customer *h* derives from choosing channel *j* in period *t*. We distinguish between two different utilities and we model these utilities following the classical multinomial logit structure. We indicate with $U_{0hjt}$ the utility of choosing channel *j* in the *trial* period at "*trial* time" *t* and with $U_{1hjt}$ the utility of choosing channel *j* at "*post- trial* time" *t'*.

$$U_{0hjt} = \beta 0_{0hj} + \beta 1_{0hj} C_{ht} + \beta 2_{0hj} E_{ht} + \beta 3_{0h} LC_{hjt} \tag{6}$$

$$U_{1hjt'} = \beta 0_{1hj} + \beta 1_{1hj1} C_{ht'} + \beta 2_{1hj} E_{ht'} + \beta 3_{1h} LC_{hjt'} \tag{7}$$

See appendix 1a for detail on the syntax used for this modeling specification.

#### 4.2.2.1 Problems with this Specification[9]

The nested logit model in this particular choice context might cause several problems. The first problem concerns the estimation of individual level lambdas (i.e. IV parameters) which might be problematic in this context where large numbers of choices are not observed per individual; therefore the estimates at the individual level could be very noisy. But, the main problem arises because the formula for the IV, in this particular context of analysis, is inconsistent. We try to explain this problem with an example. Suppose we have three choice alternatives, i.e. three different utility functions. In this situation the true IV is:

$$IV = ln(exp(U_1)+exp(U_2)+exp(U_3)) \tag{8}$$

This can be written also as:

$$S_1: IV = U_3 + ln(exp(U_1-U_3)+exp(U_2-U_3)+1) \tag{9}$$

Or:

$$S_2: IV = U_2 + ln(exp(U_1-U_2)+exp(U_3-U_2)+1) \tag{10}$$

---

[9] We thank Professor Sanjon Misra for his useful considerations. We summarize in this section his main arguments which helped us to better understand the perils of the use of the nested logit approach in this specific context of analysis.

Note that the *IV* values under specifications $S_1$ and $S_2$ are identical (as long as utilities are known). Normalizing under different base channels causes some changes to the *IV* values but this is not a problem in typical choice/nested logit models since the differences in normalizations are reflected only in the channel intercepts (which will adjust accordingly). So, for example, if the intercept in $U_3$ is normalized to zero (in $S_1$) or the intercept in $U_2$ is normalized to zero (in $S_2$) we might get different IV values but this will not bias parameters or probabilities since the variation in the *IV*'s due to any changes in the brand specific *X*'s will remain the same.

The problem comes up when one has to specify a base channel in a pure MNL (not alternative specific covariates). This, unfortunately, is our specification. In such cases there is a severe problem since any changes in covariates (in our case Catalogs or Emails) change the relative intercepts themselves. This can cause weird results. The reason is that we cannot recover the true *IV*. This is easy to see by setting $U_3$=0 in $S_1$ and $U_2$=0 in $S_2$, we will not just get different values of *IV* but any changes in a covariate will impact the compensating differential differently too. This problem has never been address in the channel choice literature. To our knowledge, the only application of a nested logit approach to model channel choice is the work of Knox (2005). Anyway, Knox modeled a binomial channel choice. For this reason he has not to set a reference channel as base. In the brand choice literature, where we can find many applications of the nested logit approaches this problem does not come out because usually brand choice models includes alternative specific covariates (e.g. price, format, etc.).

#### 4.2.3 MNL with No Purchase Option

Using this formulation we use the classical multinomial logit model but we account for purchase incidence by including a “no purchase” alternative in the customer’s choice set.

Using this modeling approach the typical customer *h* (*h*=1,2,….,*H*) in period *t* (*t*=1,2,….,*T*) may select channel j (j=1,2,..J) or he can decide to not purchase (j=0). As usual, we distinguish between two different set of models (*trial* and *post-trial*):

$$\begin{aligned}
&Trial\\
&U_{hjt} = \beta_{hjo} + \beta_{hj1} CS_{ht} + \beta_{hj2} ES_{ht} + \beta_{h3} LC_{ht} + \varepsilon_{hjt} \qquad {\scriptstyle j=1,2,\dots,J}\\
&U_{hjt} = \alpha_{h1} Season + \alpha_{h2} Trend \qquad {\scriptstyle j=0}\\
&Post - Trial\\
&U'_{hjt} = \beta'_{hjo} + \beta'_{hj1} CS_{ht} + \beta'_{hj2} ES_{ht} + \beta'_{h3} LC_{ht} + \varepsilon'_{hjt} \qquad {\scriptstyle j=1,2,\dots,J}\\
&U'_{hjt} = \alpha'_{h1} Season + \alpha'_{h2} Trend \qquad {\scriptstyle j=0}
\end{aligned} \tag{11}$$

See appendix 1a for details on the syntax.

#### 4.2.3.1 Problems with this Specification

Many studies have found that the purchase incidence decision is theoretically distinct from the brand choice decision. Hence, it may be not appropriate to model the no-purchase decision as just another alternative in the choice set with the IIA restriction holding across brand and no-purchase (Chintagunta, 2001).

This could lead to unstable parameter estimates and to convergence problems[10]. Table 4.1 are summarizes the principal characteristics of the above described models. In order model channel migration behavior at individual level the above described modeling approaches represents possible modeling strategies. Each modeling approach described has been widely used in brand choice literature. The channel migration literature is still underdeveloped, however we can notice that some of these approaches have been used as

[10] This is what happened in our case. We didn't reach convergence after 1 million iterations.

well, but often with a binomial dependent variable (e.g. Knox 2005, Ansari et al. 2008). Our context of analysis presents some peculiarities. First, we have more than 2 channel options (i.e. a multinomial dependent variable). Second, we do not have alternative specific variables. Third, our purpose is to estimate individual level parameters. These features might create some estimation and conceptual problems which make the above mention modeling approaches non optimal solutions. We have summarized for each modeling approach above described its limitations in this context of analysis.

*Table 4.1: Alternative Modeling Approaches*

| | **Advantages** | **Specification Problems (in this context)** |
|---|---|---|
| **MNP with no purchase Option** | ▪ Relax the assumption of IIA.<br>▪ It handles correlated unobserved factors over time problem. | ▪ Problems in the error terms' var. covar. matrix identification<br>▪ Arbitrary in building the variance covariance matrix structure |
| **Nested Logit** | ▪ Relax the assumption of IIA across groups | ▪ We cannot recover the true *IV* |
| **MNL with no purchase option** | ▪ The choice probabilities take a closed form and they are readily interpretable. | ▪ IIA assumption across channel choices and no-purchase option<br>▪ Potential convergence problems |

In the next paragraphs we present the multinomial channel selection switching model develop using a pure multinomial logit to handle channel choice. This model represents the best solution, taking into consideration the above mentioned limitations. Of course it has disadvantages as well, the principal disadvantage is that we do not consider incidence decision. Anyway, the advantages of its use are greater that the disavantages. Furthermore, we no not envisage strong conceptual limitations on its use in this context of analysis. In addition, it is well-supported in the literature even to model channel migration process (see. Thomas and Sullivan, 2005).

## 4.3 Multinomial Logit Channel Selection Switching Model

We conceptualize the customer decision process as a multinomial logit. The typical customer $h$ ($h$=1,…, H) observed over $t$=1, 2, …., $T_h$ purchase occasions decides which channel to select among the $J$ channels of the company during the $t^{th}$ purchase occasion. On any given purchase occasion, one among the elements of the vector $y_{ht}$=($y_{ht1}$, $y_{ht2}$, …$y_{htJ}$) takes the value 1. Equivalently this type of outcome may be represented by a categorical indicator $D_{ht}=j$ if $y_{htj}$=1 and the others are zero ($y_{htk}$=0 for $k\neq j$) where $j$ =1,2,…$J$.

Customer $h$ may move from a learning phase (*trial*) to a phase in which he or she has formed channel preferences (*post-trial*). The probability to switch to a *post-trial* stage and the length of the *trial* period is governed by a geometric distribution (see Aaker, 1971). The quickness of this transit varies for different customers. Thus, we allow for heterogeneity in the duration of the *trial* period among customers.

The modeling approach that we propose has two distinct components. The first models the probability to switch to the *post-trial* stage (we called it learning model). The seconds model the multinomial outcome $y_{ht}$ (multinomial logit model). Formally, we combine these components into an overall model that we called *multinomial logit channel selection switching model.* This overall model takes into account that customers might be "learning prone" or not associating to each customer a specific probability to switch to the *post-trial* model. It takes also into account that customers may use one *trial* multinomial logit when they are first acquired, but then migrate, after an heterogeneous number of purchase occasions, toward a *post- trial* multinomial logit.

### 4.3.1 The Learning Model

The purpose of this "learning model" is to estimate the number of purchase occasions that a generic customer *h* needs before switching to the *post-trial* stage. In other words, it estimates the length of the *trial* period in term of number of purchase occasions. As Aaker we use a geometric distribution to estimate the length of the *trial* period and we allow for heterogeneity among customers. Geometric distributions have been widely used in marketing to estimate count data models (e.g. Buchanan and Morrison, 1988; Morrison and Perry 1970; Fourt and Woodlock 1960).

Equation 12 and 13 represent the geometric distribution which governs the transition from the *trial* period multinomial logit to the *post- trial* multinomial logit:

$$q_h = \frac{1}{1+\exp(-(c_{0h}))} \tag{12}$$

$$X_{ht} = 1-(1-q_h)^{t-1} \tag{13}$$

$q_h$ represents the probability that the customer switches to the *post- trial* model. We can infer that a $q_h$ close to zero means that customer *h* has a very low probability to switch to the *post-trial* model (i.e. for customer *h* we do not observe a *trial* stage). $X_{ht}$ represents the probability that the customer is using the *post- trial* model during the purchase occasion *t*.

### 4.3.2 Multinomial Logit Channel Choice Models

We consider a market with utility-maximizing customers. The typical customer *h* in purchase occasion *t* may choose to select the channel *j*. $U_{hjt}$ represents the utility that customer *h* derives from choosing channel *j* in period *t*. We distinguish between two different utilities and we model these utilities following the classical multinomial logit structure. We indicate with $U_{0hjt}$ the utility of choosing channel *j* in the *trial* period at *trial* purchase occasion *t* and with $U_{1hjt}$ the utility of choosing channel *j* at "*after trial* purchase occasions *t'*.

$$U_{0hjt} = \alpha_{0hj} + \beta 1_{0hj} CS_{ht} + \beta 2_{0hj} ES_{ht} + \beta 3_{0h} LC_{hjt} \tag{14}$$

$$U_{1hjt'} = \alpha_{1hj} + \beta 1_{1hj1} CS_{ht'} + \beta 2_{1hj} ES_{ht'} + \beta 3_{1h} LC_{hjt'} \tag{15}$$

Four types of elements accommodate the substantive issues and our modeling assumptions: unobserved customers characteristics (random intercepts), catalogs sent (*CS*), emails sent (*ES*) and state dependence (*LC*). We have both alternative-specific and non-alternative specific independent covariates. For identification purposes we set one channel as base. Thus, we will have alternative specific coefficients only for *J*-1 channels. In the following line we describe each element considered in utilities 14 and 15.

### 4.3.2.1 Unobserved Customers Characteristics

Customers have different preferences over channels for exogenous reasons that are unrelated to the customers' past purchase histories. A customer who was observed to choose channel *j* at purchase occasion *t*-1 is more likely to have preferences such that he or she generally prefer channel J than a customer who chose channel *i* at purchase occasion *t*-1. For this reason alone, a customer who chose *j* at *t*-1 is more likely to chose *j* at *t* than is a customer who chose *i* at *t*-1. Keane (1997) in the brand choice context refers to such differences in exogenously given preferences of consumers as *heterogeneity*. Heterogeneity represents one of the explanations of the observed persistence in choice.

Unobserved heterogeneity is theoretically proven and well documented phenomena (Lancaster, 1979). Controlling for unobserved heterogeneity improves the accuracy of the parameter estimates of the covariates in the model, because, if heterogeneity is not accounted for, the omitted unobservable factors may be correlated with some of the covariates that cause aggregation bias (Gönul, Kim and Shi 2000). In addition, recent evidence in the marketing literature suggests that observed heterogeneity is not sufficient to capture differences across customers and that another measure to capture unobserved differences is called for (Kamakura & Russell, 1989; Gonul & Srinivasan, 1993). Channel preferences heterogeneity can be accounted for letting a certain parameters of the utility function differ across customers (see Elrod 1988, Jones and Landwehr 1988, Steckel and Vanhonacker 1988, McCulloch and Rossi, 1994; Keane, 1997).

We control for unobserved heterogeneity using individual level random intercepts which capture the effect of unobserved specific variables of customers. Specifically, $\alpha_{0hj}$ and $\alpha_{1hj}$ represent channel-specific and customer-specific intercepts (for identification purpose $\alpha_{0h1}$=0 and $\alpha_{1h1}$=0). These intercept terms can be interpreted as preferences for channel $j$.

#### 4.3.2.2 Catalogs Sent and Emails Sent

Previous works empirically support the idea that marketing influences channel choice and that the influence is heterogeneous across customers (Thomas and Sullivan, 2005; Knox, 2005; Ansari et al., 2008; Pauwels and Neslin, 2006; Venkatesan et al., 2006). Marketing instruments studied to date include emails, catalogs, and other direct forms of communications.

Despite the recent works which have considered maketing variables in their channel migration models, much more need to be learned about customer heterogeneity in order to understand what type of customers respond to marketing moving to one channel versus the other (Blattberg, Kim and Neslin, 2008 p.647). For these reason, we aim to examine the impact of different marketing communications in guiding customers to the channel choice of a particular firm.

We consider two different types of direct marketing communications: catalogs and emails. In line with Ansari, Mela and Neslin (2008) we define communication *c* a particular communication sent by the company at a particular time. Two different catalogs mailed at two different times are considered two different communications. The individual *h* in purchase occasion *t* may receive: nothing, *n* emails, *k* catalogs or both. Specifically, $C_{ht}$ indicates the number of catalogs that customer *h* received at purchase occasion *t* and $E_{ht}$ indicates the number of emails that customer *h* received at purchase occasion *t*.

The number of catalogs sent ($C_{ht}$) and emails sent ($E_{ht}$) variables are not alternative specific. The parameters which concerns direct marketing communication variables ($\beta 1_{0hj}$, $\beta 2_{0hj}$, $\beta 1_{1hj}$ and $\beta 2_{1hj}$) vary across channel alternatives.

***Selectivity and Endogeneity Bias in Direct Marketing CommunicationVariables***

Both the catalogs sent (*CS*) and emails sent (*ES*) variables included in our channel choice model might be determined by firm's marketing strategy variables and customers' demographic profile.

Consumers' choice behavior models (e.g. brand choice) often include marketing variables (e.g. price, advertising, marketing communications) among their independent variables. Many companies use RFM[11] values in order to determine their marketing policies (e.g. mailing decision, promotion, etc.). For example, if the company has to heed a mailing budget constraint, then not every customer with positive expected profit can be sent mail, the mailing decision might be based on an appropriate threshold (Gönül, Kim and Shi 2000).

These marketing policies may produce two main problems in the estimation of customer behavioral models. For example, if a customer is not selected to receive a marketing offer he or she has no way to respond to the offer. Consequently, the RFM values for this customer will deteriorate regardless of the true tendency to respond.

Similarly, Rhee and Russel (2003) clarify that the use of RFM information in targeting households creates major problems in empirically estimating a model of household purchase behavior. Specifically, it may cause selectivity[12] and endogeneity[13] bias. Blattberg, Kim and Neslin (2008) emphasize that selectivity bias due to target marketing is a concern for all database marketing models that include marketing among the independent variables and it represents a challenge in modeling customer channel migration.

[11] RFM stands for Recency, Frequency and Monetary. Recency (R) is defined as the number of periods since the last purchase. Frequency (F) is defined as the total number of orders placed over a standard period of time. Monetary value (M) is defined as the dollar amount that the household has spent in all purchases to date.

[12] If the firm selects households for mailings based on a non-random selection rule (such as the RFM code), a study that only analyzes the selected households generates biased results. This bias arises from the fact that the researcher does not observe the responses of non-selected households (Rhee and Russel, 2003).

[13] An explanatory variable is said to be endogenous if it is correlated with the error terms. In traditional usage, a variable is endogenous if it is determined within the context of a model (Wooldridge, 2002). Endogeneity usually refers to situations where observed explanatory variables are correlated with error terms, so that standard estimation procedures that rely on independent errors cannot be used directly. The classical form of endogeneity arises in random utility models if variables that enter systematic utility components are correlated with random utility components (Louviere et al., 2005).

The customers' marketing variables profile depends upon both the customers' inherent interest in the firm's product or services and the firm's assessment of which customers should receive a solicitation. For example, in the channel choice context if we include emails as an independent variable in the channel choice equation one might argue that there may be unobserved variables (e.g. internet orientation) that generate the receipt of emails and these same variable generate channel choice (Blattberg et al., 2008).

Selectivity bias and endogeneity are slightly different problems but it is highly likely that they come together. Selectivity bias is a special type of missing data problem but it should be noted that it is not a concern if customers can select the product/service without receiving a marketing solicitation because in this situation all customers responses are observable (Rhee and Russel, 2003).

The correlation between the marketing variables and the error term results in the endogeneity problem. Not accounting for this correlation will give incorrect estimates for the effects of the included marketing variables (Chintagunta, 2001; Shugan, 2004).

In formal statistical terms, it can be shown that endogeneity yields incorrect parameter estimates in a predictive model due to unobserved correlations between the marketing variables and the error in the model (see for example Davidson and MacKinnon, 1993).

Blattberg et al. (2008) list some ways to address these issues. First, to include all variables that generates marketing contacts (i.e. the same RFM marketing variables that companies use to target mailing). In this way, these variables are observed and accounted for. A second alternative is to specify a formal model which takes into accounts these issues, e.g. allowing the error terms between equations to be correlated (see Ansari, Mela and Neslin, 2008) or using a two least stage approach (see Gonul, Kim and She 2000).

As it emerges from this discussion, catalogs sent (*CS*) and emails sent (*ES*) variables in our channel choice model behave like an endogenous variable, since they might be a function of the firm's direct marketing strategy and of several intrinsic customers characteristics. Consequently, there may be a potential endogeneity bias in the estimates that using catalogs sent and emails sent as channel choice covariates.

Similarly to Gönul, Kim and Shi (2000), we use a two-stage least- squares approach to minimize this bias by using an instrument instead of the actual value of the catalogs sent and emails sent variables. This approach is in the same spirit as using instrumental variables in place of an endogenous variable on the right-hand side of an equation, in two-stage least-squares frameworks.

The two-stage least squares approach is the most common method used for estimating simultaneous-equation models (Greene, 2002) and it is often used in order to address endogeneity bias problems. It is frequently applied in traditional OLS regression models. This approach consists in two estimation stages:

1) In this stage endogenous variables became the depended variable of a new regression model. In this stage new variables are created (called instrumental variables) which replace the problematic endogenous variables. This is accomplished using a new OLS regression in which the problematic endogenous variable is the dependent and additional instrumental variables (called instruments) are the independents. The instruments are the exogenous variables which might cause the problematic endogenous variable. The predicted values of the dependent variable of this regression model are used in the second stage.
2) In this stage the original model is estimated, but using the predicted values of the newly created variables. In this way, the problem of the correlation among the endogenous variables and the error terms should be minimized.

There is ample literature which supports the 2SLS approach with traditional OLS regression models as a possible way to minimize potential endogeneity bias (see Greene 2002 and Wooldridge 2002). Similar applications in context other than traditional OLS regression and other than continuous endogenous variables have been developed and applied empirically (see Heckman 1978, Neslon and Olson 1978, Windmeijer and Santos Silva 1997, Gonul, Kim and Shi, 2000).

Specifically, we estimate the probability to receive a certain number of emails with a poisson model. We use as explanatory variables seasonality dummies, the number of purchases made lagged, and customer characteristics. Seasonality dummies take into consideration that during particular quarters, e.g. Christmas quarter or mothers' day, the firm might send more emails or catalogs. The number of purchases lagged should control for RFM strategy of the company which might sent more emails and catalogs to those customer who purchase more and more rapidly. Customer characteristics (e.g. age, gender, Internet orientation) account customers' inherent interest in the firm's channels. See appendix 2 for details on the models used and the results.

#### 4.3.2.3 State Dependence

State dependence might be defined as a causal link between past and present purchase behavior (Heckman 1981, Keane 1997). Thus, it investigates the effects of a customer's current choice on its future choices. In the context of brand choice Keane (1997) shows that it may be that purchase of a particular brand at time *t*-1 makes the consumer more likely to purchase that brand again at *t*. There are a myriad of plausible explanations for such a causal link between current and past behavior.

Similarly, Seetharaman (2004) distinguishes the different behavioral explanations related to the concept of state dependence. He identifies four different sources of state dependence. He indicates with the term *structural state dependence* the fact that a household's prior purchase experiences with specific brands typically influence the household's purchase propensities for the same brands in the future. Structural state dependence can be positive or negative, in which cases they are called inertia (Jeuland 1979) and variety seeking (McAlister 1982), respectively.

There has been a lot of empirical work in marketing over the past 20 years on the estimation of structural state-dependence effects using scanner panel data. The consensus that has emerged in this literature is that there is substantial evidence of structural state dependence in households' brand choices even after adequately controlling for unobserved heterogeneity across households (Keane 1997, Abramson et al. 2000, Moshkin and Shachar 2002). State dependence, as well as customer heterogeneity, represents an explanation for the observed persistence in customer choices. For this reason, Heckman (1981) underlines the importance of taking into consideration both state dependence and customer heterogeneity in choice models. If heterogeneity is present in the true model and one ignores it, estimating a model that only allows for state dependence, one will tend to overestimate the degree of state dependence (spurious state dependence).

State dependence can be accounted for by allowing past purchases to have an impact on current-period utility evaluations (Guadagni and Little 1983). Similarly, we include a state dependence variable in our channel choice model. $LC_{hjt}$ represents state dependence. It is a channel specific variable which indicates which channel customer $h$ chose at time $t$-1. $\beta 3_{0h}$ and $\beta 3_{1h}$ represent the state dependence parameters which work on all three channels

### 4.3.3 The Final Model

We compute the probability that customer $h$ chooses channel $j$ at time $t$ ($PrC_{hjt}$) as follow:

$$PrC_{hjt} = (1 - X_{ht}) * \left( \frac{\exp(U_{0hjt})}{\sum_{j=1}^{J} \exp(U_{0hjt})} \right) * \left( \frac{1}{1 + \exp(-Z_{0ht})} \right) + X_{ht} * \left( \frac{\exp(U_{1hjt})}{\sum_{j=1}^{J} \exp(U_{1hjt})} \right) * \left( \frac{1}{1 + \exp(-Z_{1ht})} \right) \tag{16}$$

This equation absorbs the customers switching behavior. Formally, it is built taking into consideration the switching regression models logic. The probability to select channel *j* at time *t* is computed taking into consideration that the customer *h* at time *t* might be switched to the *post-trial* stage. In other words, it takes into consideration that the customer *h* might have a different utility to select channel *j* in the *trial* ($U_{0hj}$) and in the *post-trial* stage ($U_{1hj}$). The moving from the *trial* stage to the *post-trial* stage is modeled along a probabilistic logic; we do not allow a strict change between the two phases. Equation 7 includes the probability that the customer *h* has to switch to the *post-trial* model at time t. We explain the underlying logic with the following example.

For each customer the learning model estimates the probability to switch to the *post-trial* stage. $q_h$ represents this probability ($0 \leq q_h \leq 1$). If $q_h$ I close to 1 it means that the customer *h* has a very short *trial* stage, if it is close to 0 it means that the customers *h* does not switch to the *post-trial* stage(i.e. it might be a *stayer*), otherwise it means that the customer needs a moderate or a long learning period. For example, let's say that the probability that the customer *h* switches to the steady state model is 0.5 (q=0.5). Thus, for the customer *h* we can estimate the probability that he is using the *post-trial* model in each period *t* ($X_{ht}$). Table 4.2 shows this example; we can notice that the customer *h* has a probability of 75% to be switched at the *post-trial* phase during the third quarter. Therefore, the probability that customer *h* will select channel *j* during the third purchase occasion is computed weighing 25% the utility $U_{0hj}$ which characterizes the *trial* period and 75% the utility $U_{1hj}$ which characterizes the *post-trial* period.

*Table 4.2: Example about the Switching Probability*

| Purchase Occasion | $X_{ht}$ if q=0.5 |
|---|---|
| 1 | 0.00 |
| 2 | 0.50 |
| 3 | 0.75 |
| 4 | 0.88 |
| 5 | 0.94 |
| 6 | 0.97 |
| 7 | 0.98 |
| 8 | 0.99 |
| 9 | 1.00 |
| 10 | 1.00 |
| 11 | 1.00 |
| 12 | 1.00 |
| 13 | 1.00 |
| 14 | 1.00 |
| 15 | 1.00 |
| 16 | 1.00 |
| 17 | 1.00 |
| 18 | 1.00 |

What happen if customer $h$ has a null probability to switch to the *post-trial* phase (i.e. $q_h$ =0)? In this situation $X_{ht}$ will be zero, doesn't matter what $t$. Therefore, the probability to select channel $j$ for customer $h$ will be computed taking into consideration only the fist part of equation 16, this because (1- $X_{ht}$) equals 1 and $X_{ht}$ equals 0 in each time period. In this situation the only relevant multinomial logit model for customer $h$ is the *trial* period model (i.e. equation 14).

## 4.4 Estimation Approach

We adopt a Bayesian approach to conducting inference in the *multinomial channel selection switching* model. Specifically, we add a hierarchical Bayesian structure in order to obtain individual-level estimates.

### 4.4.1 Why a Hierarchical Bayesian Estimation Approach?

Rossi, Allenby and McCulloch (2005, p. 4) argue that there are really no other approaches which can provide a unified treatment of inference and decision as well as properly account for parameter and model uncertainty, in addition they argue (p.129) that a hierarchical Bayesian approach is particularly useful in marketing practices which are designed to respond to consumer differences, therefore, require an inference method and model capable of producing individual or unit-level parameter.

The key reason which induces us to use a hierarchical Bayesian estimation approach is that we aim to estimate individual level parameter. Classical econometric methods do not allow for the estimation of individual-level parameters. As Rossi et al. (2003) point out the models of heterogeneity considered in the econometrics literature often restrict heterogeneity to subsets of parameters such as model intercepts. In this context (and in general in many marketing situations) there is no reason to believe that differences should be confined to the intercepts and differences in slope coefficients are critically important considering the aims of this research.

In addition, our data panel structure is characterized by a large number of unit (individuals) relative to the length of the panel (number of purchase occasions considered). These types of data set generally present a small amount of information about the decision unit compared, for example, to the amount of information of cross-sectional survey data set. For a variety of reasons this is a typical problem in panel data base with a large number of customers observed for a relatively large time horizon. However, a flexible model, combined with Bayesian inference methods, can produce accurate estimates at both aggregate and individual decision unit level. (Rossi et al., 2003).

The hierarchical structure is commonly used for individual-level parameter estimates. In the hierarchical model, we assume that each parameter is drawn from a superpopulation

(McCulloch and Rossi, 1994). It generally consists of two stages of priors (first-stage priors and second stage-priors); we specify a second-stage prior on the hyperparameters of the first-stage prior. Nonhierarchical models are usually inappropriate for models with large datasets with many parameters to estimate because they tend to "overfit" such data in the sense of producing models that fit the existing data well but lead to inferior predictions for new data. In contrast hierarchical models can have enough parameter to fit data well, while using a population distribution to structure some dependence into the parameters thereby avoiding problems of overfitting (Gelman et al. 2004, p. 117).

Since at individual level, in several instances, the amount of information available for many units is small, the specification of the functional for and hyperpararameter for the prior may be important in determining the inferences made for any one unit. For example, it may happen that some consumers do not choose all of the alterative available during the course of observation, e.g. some customers may choose mainly the Store. The specification of the priors, in these cases can be very important, due to the scarcity of data for some units. Both the form of the prior and the values of the hyperparameters are important and can have effects on the inferences. The situations in which the investigator has no strong prior beliefs about the location of the model parameters may be approximated by choosing extremely diffuse, but proper, priors. The diffusion of the prior relative to the likelihood determines how strong an influence the prior will have on the posterior distribution of the identified model parameters. It is common to specify a normal and diffuse prior for the parameters.

We hypothesize that the "first stage" priors of the considered parameters follow a normal distribution with mean $\mu$ and precision $\tau$. The "second stage" priors follow a normal distribution for $\mu$ and a gamma distribution for $\tau$ (i.e. and inverse gamma distribution on the variance). We aim to use a vague prior for the prior mean, e.g. $\mu$ ~Normal(0,0.00001), and a proper but vague inverse-gamma prior on the variance, e.g. $\tau$ ~Gamma(0.5,0.5).

In summary, as in many other marketing works (e.g. Rossi, Mcculloch and Allenby, 1996) we adopt a Bayesian approach to conducting inference in this model rather than a standard classical econometric approach for two main reasons:

1) Marketing actions require inference about the customer parameters directly and not just on the common parameters.
2) We are making inferences in many cases on the basis of the handful observations; therefore we need a method which properly accounts for parameter uncertainty.

# *Chapter 5*

# DATA, MODEL SELECTION AND RESULTS

## 5.1 Data

This research aims to study customer channel migration process; specifically we aim to obtain insights into how the channel decision process of newly acquired customer evolves over purchase occasions. Furthermore, we want to understand the role which marketing plays in this channel migration process. To test our model data should have at least three main characteristics. First, we need data on a cohort of new customers. In particular, customers' purchasing history should be tracked from the first purchase ever made in the product category. Second, we need longitudinal records tracking all the channels from which customers purchased. Third, we need information about the firm's marketing instruments across channels.

We use data from a major retailer in a European country market which operates in the book industry. The data set consists of several files. A transaction file indicates which channel was selected by each customer during each purchase occasion. It also contains data on how much was spent, on returns, and on others information about transactions. The direct marketing communication file indicates which customers received emails and catalogs, the number and type of communications received, and it tracks the exact time during which customers received them. Linking the marketing data set to the channel choice data set we can estimate the impact of marketing on channel selection.

The attractiveness of this dataset relies on the contractual nature of the relationship between the company and its customers. The company uses different recruiting strategies (see table 5.1) and it operates a subscription-oriented business model, thus each customer must become a member in order to purchase. This characteristic allows us to track every transaction in whatever channel of the firm without loss of information on the store choice. A code number is associated to each customer, tracking the customer each time he or she purchases an item from the various channels.

*Table 5.1: Frequency of the Different Recruiting Strategies*

| Recruiting Strategies | Freq % |
|---|---:|
| Mailing | 0.31% |
| Internet | 0.87% |
| Door to door agents | 60.87% |
| On the street agents | 35.80% |
| Member gets member | 1.30% |
| Others | 0.85% |

The considered period starts in October 2001 and it ends in June 2006. The order and marketing data span 18 quarters subsequent to the "recruiting period"[14].

We select a cohort of new customers following specific inclusion rules. We restrict the attention to customers who have bought at least two times per year (see Ansari, Mela and Neslin 2008), who live in stores' attraction area, and who subscribes a contract of membership with the company within a particular time range, i.e. the last quarter of 2001. We turned out with a cohort of 15,555 customers.

In this "contractual" setting the time at which customers become "inactive" is observed. Specifically, we observe that roughly 80% of the customers quitted the firm before

[14] The recruiting period represents the period in which customers were recruited and became members of the firm. In this dataset this period corresponds to the last quarter of 2001, i.e. October, December 2001.

the end of June 2006. Therefore, we have 3,134 customers with a "full life cycle" who, in other words, they never quit the firm (see table 5.2). In addition, during the considered period some new physical retail stores opened. A new store opening causes a modification in the customers' choice set. For analytic purposes we need to consider those customers who have a full life cycle[15] and a "full channel choice set" from the beginning of the relationship with the company till the end. For this reason, we restrict our attention only to customers who did not experience a new store introduction during the period under consideration and to customers who did not quit the company before the end of June 2006. The resulting sample size is made of 1018 customers.

*Table 5.2: Number of Active Customers Each Year*

| | *2001*[a] | *2002* | *2003* | *2004* | *2005* | *2006*[b] |
|---|---|---|---|---|---|---|
| Active customers | 15555 | 14947 | 9157 | 5780 | 3705 | 3134 |
| Quitters | - | 608 | 5790 | 3377 | 2075 | 571 |

*a the cohort starts during last quarter of 2001*
*b the observation period ends during the second quarter of 2006*

In summary, to test our model we use 1018 customers and 18 quarters. We aggregate data to the quarterly level, as the mean purchase frequency is about 2.5 purchases per year. Finer gradations yield an excess of observations with zero sales, and coarser gradations result in multiple purchases within a single interval[16]. That is, the quarterly sampling rate corresponds largely to the decision processes we model.

[15] Our purpose is to estimate two distinct channel selection models (one for the *trial* period and another for the *post-trial* period, see chapter 4 for details). For this reason we need to guarantee that customers continue to purchase during the overall selected period and that they do not quit the firm. Otherwise the *post-trial* parameter estimates will be biased from the presence of quitting or inactive customers. Hence, we restrict the attention only to the "active" customers.

[16] When using quarterly aggregation, multiple purchases are negligible. When there are multiple purchases in the same quarter, we classify the channel with the higher order-size as the channel of choice.

We have 16,003 observations available for estimation. However, the initialization period necessary to create lagged variables reduced our estimation sample to 14,985 observations.

The firm sells through different channels: phone, sms internet, mail, fax, and stores. We designate phone, mail, sms and fax into the catalog. Therefore, the customers in our dataset may choose among three main channels: physical retail stores, catalogs, and internet. Among the selected sample of 1018 customers 71% are single-channel-users (see table 5.3). Specifically, 27.6% use only the catalog, 42.4% use only the stores and 0.6% only internet (see table 5.4). We remark that the customer defined as single-channel-user do not always use strictly the same channel in each purchase occasion. Specifically, only 39% of the sample use always the same channel over time (18% tried once another channel and 15% tried twice others channels). 29.4% are multichannel customers (4.3% use three channel, and 25% two channels). Table 5.3 compares the proportion of multichannel customers in the different samples.

*Table 5.3: Proportion of Multichannel Customers*

| | **Sample Size** | **Multichannel** | **Single-channel-users**** |
|---|---|---|---|
| Total | 15555 | 3252* *(21%)* | 11806 *(76%)* |
| Full life cycle sample | 3134 | 1051 *(34%)* | 2083 *(66%)* |
| Full life cycle and full choice set sample | 1342 | 411 *(31%)* | 931 *(69%)* |
| Full life cycle, full choice set and more that 3 purchases per year | 1018 | 299 *(29%)* | 719 *(71%)* |

** 3% of the customers purchased only once*
*** by single-channel-users we mean that customers mainly use only one channel, they could have tried once or twice different channels.*

*Table 5.4: Number of Active Customers each Year Distinct by Channel Usage*

| | *Entire data set (15,555 customers)* | | *Selected sample (1,018 customers)* | |
|---|---|---|---|---|
| **Channel Usage** | Sample size | Percent | Sample size | Percent |
| Only one purchase | 497 | 3.2% | - | - |
| Mainly Catalog | 6692 | 43.0% | 281 | 27.6% |
| Mainly Internet | 222 | 1.4% | 6 | 0.6% |
| Mainly Store | 4892 | 31.4% | 432 | 42.4% |
| Catalog and Store | 1523 | 9.8% | 157 | 15.4% |
| Catalog and Internet | 1353 | 8.7% | 86 | 8.4% |
| Internet and Store | 137 | 0.9% | 12 | 1.2% |
| Catalog, Internet and Store | 239 | 1.5% | 44 | 4.3% |
| **Multiple-Channel-Users** | | | | |
| Yes | 3252 | 20.9% | 299 | 29.4% |
| Two channel Buyer | 3013 | 19.4% | 255 | 25.0% |
| Three Channel Buyer | 239 | 1.5% | 44 | 4.3% |
| No | 11806 | 75.9% | 719 | 70.6% |
| Total | 15555 | 100% | 1018 | 100% |

Table 5.5 compares the proportion of customers who quitted the firm distinguishing per channel usage. Interestingly, 32.3% of the multichannel customers in the entire data set of 15,555 customers do not quit the firm (49% if we consider the three-channel-users). Roughly 85% of the single-channel-users quitted the firm before June 2006. Only the 13% of the single catalog users stay. This evidence seems to suggest that multichannel customer exhibit a grater loyalty then single-channel-users. Some researchers argue that purchasing from multiple channels increases customer service and satisfaction, therefore loyalty. A higher loyalty seems to be a natural consequence of multichannel usage (see Blattberg, Kim and Neslin 2008).

*Table 5.5: Percentage of Customers who Quitted theFirm Distinct by Channel Usage*

| **Channel Usage** | *Percentage of customers who quitted the firm before June 2006* |
|---|---|
| Mainly Catalog | 87% |
| Mainly Internet | 82% |
| Mainly Store | 82% |
| Catalog and Store | 72% |
| Catalog and Internet | 71% |
| Internet and Store | 76% |
| Catalog, Internet and Store | 51% |

An emerging generalization in multichannel research is that multichannel shoppers purchase higher volumes (Neslin et al. 2006). Multichannel customers indeed appear to spend more on average during each visit than customers who only use one channel. Consistent with prior research (Kumar and Venkatesan, 2005; Thomas and Sullivan, 2005; Myers, Van Metre, and Pickersgill, 2004; Kushwaha and Shankar, 2005; Ansari et al., 2008) we further break down our sample distinguishing it by channel usage, and we describe the groups in terms of the average amount spent. Figure 5.1 shows the average amount that the average customer spends each quarter. It confirms the evidence that multichannel customers buy more.

*Figure 5.1: Average Amount Spent OverTime Break Down by Channel Usage*

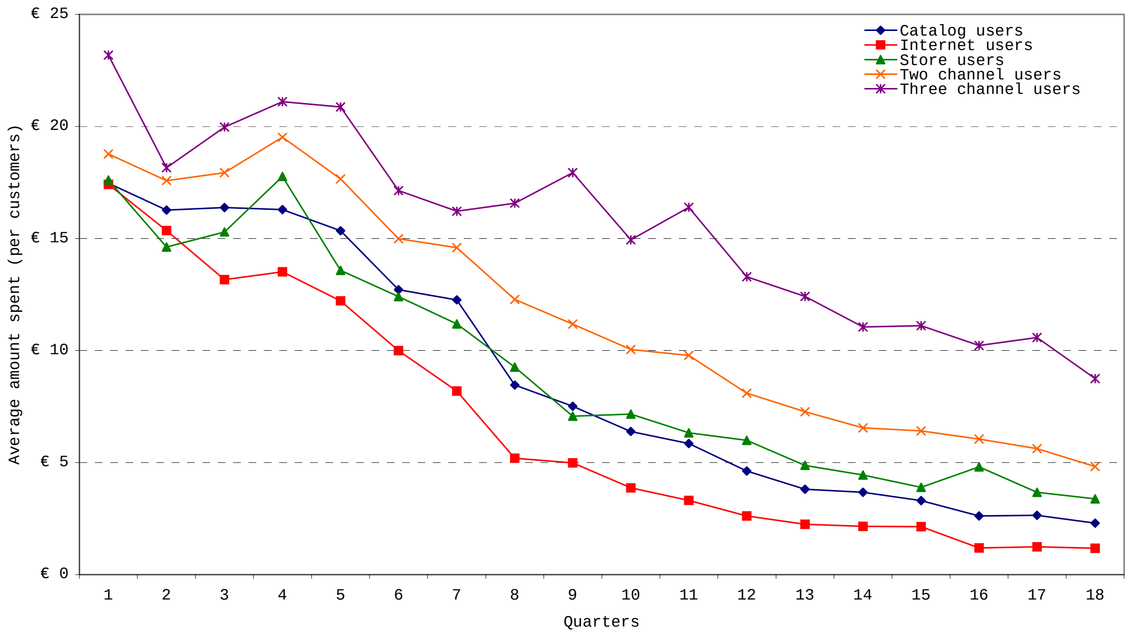



Table 5.6 and table 5.7 compare the entire data set and the selected sample of 1,018 customers. According Table 5.6 multiple-channel-users appear to spend more, on average, than single-channel-users. This result is strongly supported in the literature (see for example Blattberg, Kim and Neslin, 2008, p. 638 for a review). However, if we consider the selected sample this evidence is no more confirmed. In this case the combination of Internet and Store seems to be the most profitable. Actually, these customers spend more and purchase on average higher quantities of items. However, the second best is the group of customers who mainly use stores, followed by the mainly internet users. Specifically, it seems that the store and the internet-store combination are the bests, but in general, it is difficult to evaluate the profitability of the different channel combinations in this sample. However, it should be remarked that the differences among the amount spent and the average quantity purchased in the selected sample are less evident that the in the entire data set. For example, an average three-channel-user in the entire data set spends on average 21.8 € per quarter and an average

catalog user spend on average 8.8 € per quarter. The same average types of customers for the selected sample spend respectively 21.3 € and 20.6 €.

*Table 5.6: Descriptive Statistics and Comparison between the Entire Data set and the Selected Sample*

| | *Entire data set (15,555 customers)* | | | *Selected sample (1,018 customers)* | | |
|---|---|---|---|---|---|---|
| **Channel Usage** | € spent per quarter (mean per customer) | Total € spent over relationship (mean per customer) | Items purchased per quarter (mean per customer) | € spent per quarter (mean per customer) | Total € spent over relationship (mean per customer) | Items purchased per quarter (mean per customer) |
| Mainly Catalog | € 8.8 | € 159.3 | 10.7 | € 20.6 | € 371.5 | 24.8 |
| Mainly Internet | € 6.7 | € 120.8 | 8.0 | € 22.7 | € 407.7 | 27.2 |
| Mainly Store | € 9.2 | € 165.2 | 15.0 | € 25.2 | € 454.3 | 41.8 |
| Catalog and Store | € 11.4 | € 204.5 | 16.5 | € 21.3 | € 383.7 | 31.8 |
| Catalog and Internet | € 12.3 | € 220.8 | 14.8 | € 21.9 | € 394.8 | 26.2 |
| Internet and Store | € 11.2 | € 202.1 | 16.2 | € 27.6 | € 496.2 | 42.7 |
| Catalog, Internet and Store | € 21.8 | € 392.7 | 30.4 | € 21.3 | € 383.6 | 28.0 |
| **Multiple-Channel-User** | | | | | | |
| Yes | € 12.1 | € 217.2 | 16.1 | € 21.7 | € 391.3 | 30.0 |
| Two channel Buyer | € 11.8 | € 211.7 | 15.7 | € 21.8 | € 392.7 | 30.4 |
| Three Channel Buyer | € 21.8 | € 392.7 | 30.4 | € 21.3 | € 383.6 | 28.0 |
| No | € 8.6 | € 154.5 | 12.0 | € 23.4 | € 421.6 | 35.6 |
| **Total** | *€ 9.3* | *€ 167.6* | *12.8* | *€ 22.9* | *€ 412.7* | *34.0* |

Of course, we should consider that the selected sample contains customers who did not quit the firm; therefore it might be that the evidence that multichannel customers spend more depends on the length of the customer "lifetime" with the firm. For example it might be that customers with a short lifetime are more profitable if they are multichannel, but customers with a long lifetime are more profitable if they are single channel or if they use specific channels combinations (e.g. store-internet). Of course, these considerations depend on simple descriptive statistics, however more effort should be made to evaluate, for example, if the length of the customer lifetime has an impact on the more profitable types of channel users.

Anyway, our descriptive analysis supports at least one dimension of the profitability of multiple-channel-users versus single-channel-users: their loyalty. It seems in fact that they have a longer lifetime. Again this result should be confirmed with more rigorous analysis. Table 5.7 shows that multichannel customers present an higher average value of returns items if compared versus single-channel-users.

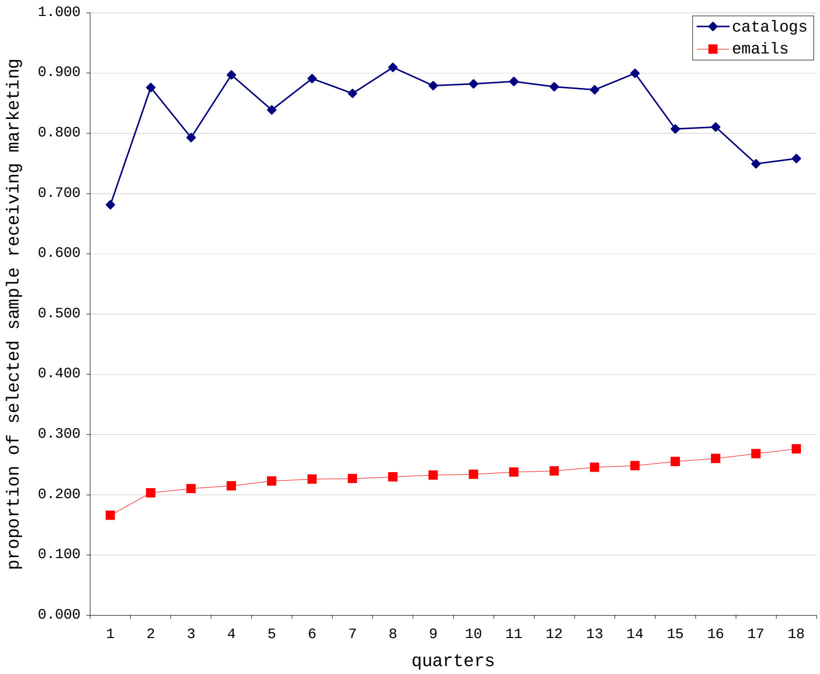
*Table5.7: Descriptive Statistics and Comparison about the Average Amount of Returns.*

| | *Entire data set (15,555 customers)* | *Selected sample (1,018 customers)* |
|---|---|---|
| **Channel Usage** | Total € returns over relationship (mean per customer) | Total € returns over relationship (mean per customer) |
| Mainly Catalog | € 8.0 | € 10 |
| Mainly Internet | € 5.5 | € 0 |
| Mainly Store | € 0.9 | € 3 |
| Catalog and Store | € 5.5 | € 9 |
| Catalog and Internet | € 9.5 | € 14 |
| Internet and Store | € 1.9 | € 3 |
| Catalog, Internet and Store | € 7.5 | € 8 |
| **Multiple-Channel-User** | | |
| Yes | € 7.2 | € 10 |
| Two channel Buyer | € 7.1 | € 10 |
| Three Channel Buyer | € 7.5 | € 8 |
| No | € 5.0 | € 6 |
| **Total** | *€ 5.3* | € 7 |

Figure 5.2 displays the proportion of customers receiving emails and catalogs over time. The firm uses different strategies to target catalogs and emails sending. For example, the firm catalog sending policy is based on RFM variables. Specifically, they send additional catalogs to the customer who purchase more rapidly. The variables used to target these

strategies are likely to be in included as independent variables in our model (see paragraph 4.3.2.2 and appendix 2 ).

*Figure 5.2: Fraction of Customers Receiving Emails and Catalogs over Time*




## 5.2 Model Selection

In this paragraph we discuss issues related to model selection of the developed model: the *multinomial logit channel selection switching model* (see chapter 4). The first goal of this section is to find the best-fitting model by estimating a series of models which vary along three different dimensions. The first dimension aims to test the existence of two distinct stages in the channel migration process. The second dimension whether there is heterogeneity across individuals or not in the length of the *trial* period. The third dimension aims to explore whether customers demographic (age and gender) affect the length of the *trial* period.

Specifically, we compare four different types of models: M1, M2, M3, and M4 (See the appendix for details on the four modeling syntaxes).

*M1 –Multinomial logit model*: This is a simple multinomial logit model which does not distinguish between *trial* and *post-trial* stages. A similar model has been estimated by Thomas and Sullivan (2005). It implicitly hypothesizes that for all the customers we can not envisage a learning phase.

*M2 –Multinomial logit model distinct in two periods*: This second model is a multinomial logit which distinguishes between two periods. We assume that the customer starts off using one multinomial logit model, and then she switches to another multinomial logit model after an *a priori* defined number of quarters[17]. This is equivalent to assume that the *trial* stage exists for all customers and that its length is homogeneous among customers, therefore the transition in to the *post-trial* phase happens after a fixed number of quarters.

*M3 –Multinomial logit channel selection switching model:* the third model is the multinomial logit channel selection switching model described in chapter 4. We assume that a customer starts off using one multinomial logit model, and then switches to another. A geometric distribution governs this transition from the "*trial* multinomial logit" to the "*post-trial* multinomial logit". A geometric distribution models the probability that the customer *h* switches to the *post-trial* model. This distribution has heterogeneous parameters. If the customer has a non null switching probability, two models describe her channel migration process. The first multinomial logit represents the decision process while the customer is learning her preferences. The second multinomial logit represents the *post-trial* decision process.

*M4 –Multinomial logit channel selection switching model (with age and gender):* this fourth model is again the multinomial logit channel selection switching model, but it assumes

[17] Specifically, we set the a priori period after nine purchase occasions.

that the probability that the customer *h* switches to the *post-trial* model is affected by some customers' demographic characteristics (gender and age). For example, younger customers may have a quick *trial* period.

We compare these four types of models using the deviance information criterion (DIC) statistic (Spielberg et al.,2002). It represents a Bayesian measure of model complexity and fit. M1, M2, M3, and M4 are estimated using hierarchical Bayesian models (see paragraph 4.4). Therefore, we used priors that themselves depend on other parameters not mentioned in the likelihood. Specifically, we hypothesize that the first stage priors of the considered parameters follow a normal distribution with mean μ and precision *prec*. The second stage priors follow a Normal distribution for μ and an Inverse Gamma distribution for *prec*. The deviance information criterion is a well-know statistic base on the log-likelihood estimation which suits the problem of comparing complex hierarchical models in which the number of parameters is not clearly defined. We used the following fomulas to estimate DIC (see Spielberg et al.2002):

$$DIC = \bar{D} + 2p_D \tag{5.1}$$

Where:

$$p_D = \mathrm{var}(\bar{D})/2 \tag{5.2}$$

$$\bar{D} = -2\log p(y\,|\,\theta) \tag{5.3}$$

$p_D$ represents an estimate of the effective number of parameters in a model. Adding $p_D$ to the posterior mean deviance ($\bar{D}$) gives a deviance information criterion for comparing models.

The DIC results for each model are presented in Tables 5.8.

*Table5.8: Model Comparison*

| *Model* | *Description* | *DIC* |
|---|---|---|
| M1 | Multinomial Logit | 6881.0 |
| M2 | Multinomial logit model distinct in two "macro" periods: | 6648.2 |
| M3 | Multinomial logit channel selection switching model | 6525.8 |
| M4 | Multinomial logit channel selection switching model (with age and gender in the geometric) | 6598.2 |

The best model is M3, which assumes the existence of two distinct models describing the channel choice process over time and that the length of the *trial* period is heterogeneous among customers. M3's superiority to M4 suggests that customers' demographics do not affect the length of the *trial* period. M3's superiority to M1 suggests that customer channel migration is characterized by two distinct stages: *trial* and *post-trial*. Finally, M3's superiority to M2 suggests that some customers have a probability to switch to a *post-trial* model and that the length of the *trial* period is heterogeneous among customers.

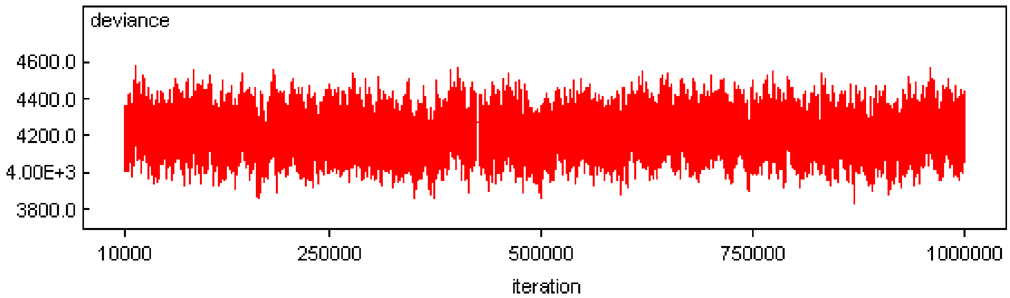
## 5.3 Model Convergence

Operationally, effective convergence of Markov chain simulation has been reached when inferences for quantities of interest do not depend on the starting point of the simulations (Brooks and Gelman, 1998). This suggests monitoring convergence by comparing inferences made from several independently sampled sequences with different starting points.

It is standard practice to discard observations within an initial transient phase (the burn in period). Specifically, we discarded the first 10,000 iterations. Most methods for inference are then based on the assumption that what remains can be treated as if the starting points had been drawn from the target distribution.

Convergence for multiple chains may be assessed using different approaches. In this study we follow two different methods in order to assess convergence: history plots diagnosis and the Brooks-Gelman-Rubin statistic. History graphs (iteration number on x-axis and parameter value on y-axis) are commonly used to assess convergence. If the plot looks like a horizontal band, with no long upward or downward trends, then we have evidence that the chain has converged. A clear sign of non convergence with a traceplot occurs when we observe some trending in the sample space. We assessed the convergence of the average parameters[18] of our model using one million iterations. The analysis of the history graphs suggests a strong evidence of the model convergence. For example figure 5.3 shows the deviance's history graph (see the appendix 3 for all the specific history graphs).

*Figure5.3: Deviance History Graph with 1 Million Iterations*

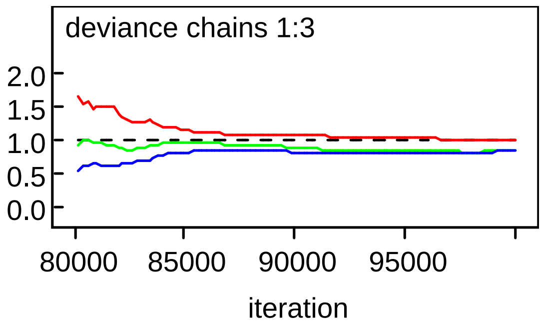



The Brooks-Gelman-Rubin statistic assesses the variability within parallel chains as compared to variability between parallel chains. The model is judged to have converged if the

[18] Our model estimates individual level parameters; therefore it estimates more that 16,000 parameters. For these reason we assessed the convergence only of the average parameters.

ratio of between to within variability is close to 1. Plots of this statistic can be obtained[19] by setting more than one chain (i.e. different initial values chains) during the specification model phase. Otherwise, one can save the parameter estimates for each iteration. In this way, the chains containing the "changing parameter estimates" can be saved into an ascii file so that more formal tests can be undertaken. We used both procedures to test convergence. The appendix contains the detail computations made. The Brooks-Gelman-Rubin statistics and plots confirmed the model convergence. Figure 5.4 shows for illustrative purposes the Brooks Gelman Rubin plot for the deviance.

*Figure5.4: Deviance Brooks Gelman Rubin Plot*




## 5.4 Empirical Results

Drawing on the Aaker's (1971) new-trier modeling logic we developed a *multinomial logit channel selection switching* model (see chapter 4). This modeling approach allows us to estimate the probability that the customer switches to the *post-trial* model and, at the same time, to estimate the length of the *trial* period at individual level.

We contend that the *trial* and *post-trial* stages are governed by a different set of parameters. Specifically, we seek to capture, at individual level, the impact of direct

[19] We used WINbugs 14.1 to estimate our model.

marketing communications (catalogs and emails), state dependence and intrinsic preferences on channel migration process. First, I present the parameter estimates which assess the probability to switch, and then I discuss the result about the *trial* and *post-trial* multinomial logit models.

### 5.4.1 The Learning Model, *Stayers*, and *Switchers*

We hypothesize that a geometric distribution governs the transition from the *trial* multinomial logit to the after *trial* multinomial logit (see equation 4.1 and 4.2 in chapter 4). The heterogeneous intercept can be interpreted as the customer's inner propensity to switch to the *post-trial* model. Table 5.9 summarizes the results. 7.96% of the customers in the data set have a 95% posterior interval which includes zero; therefore 92% of customers have a significant intercept.

*Table5.9: Learning Model Intercepts' Statistics*

Learning model intercept

| | Percentiles | Smallest | | |
|---|---|---|---|---|
| 1% | -6.275 | -6.661 | | |
| 5% | -6.205 | -6.661 | | |
| 10% | -6.153 | -6.661 | Obs | 14985 |
| 25% | -5.987 | -6.661 | Sum of Wgt. | 14985 |
| 50% | -5.6 | | Mean | -5.185078 |
| | | Largest | Std. Dev. | 1.207348 |
| 75% | -5.044 | -.1785 | | |
| 90% | -3.214 | -.1785 | Variance | 1.45769 |
| 95% | -2.375 | -.1785 | Skewness | 1.779459 |
| 99% | -.9528 | -.1785 | Kurtosis | 5.547144 |

Table 5.10 shows the descriptive statistics about the probability to switch. The average value is 0.043 (see table 5.10), which means that, for example, on average a generic customer of our sample has a probability to switch greater than 50% after the sixteenth purchase

occasion (see equation 4.3 for detail on this computation). This is plausible because we expect that not all the customers in our database are *switchers*, therefore this average result is affected by the presence of *stayers* whose probability to switch is close to zero.

*Table 5.10: Probability to Switch*

Probability to switch to the post-trial model

| | Percentiles | Smallest | | |
|---|---|---|---|---|
| 1% | .01078 | .006024 | | |
| 5% | .01275 | .006024 | | |
| 10% | .01392 | .006024 | Obs | 14985 |
| 25% | .01723 | .006024 | Sum of Wgt. | 14985 |
| 50% | .02832 | | Mean | .0435346 |
| | | Largest | Std. Dev. | .0597367 |
| 75% | .03302 | .4997 | | |
| 90% | .08609 | .4997 | Variance | .0035685 |
| 95% | .1496 | .4997 | Skewness | 4.46514 |
| 99% | .3653 | .4997 | Kurtosis | 26.78424 |

In chapter 3 we have discussed the theoretical reasons behind customers' learning proneness. Basing on this, we aim to split our sample in two groups of customers: customers with a low probability to switch (*stayers*) and customers with a high probability of changing their decision process, i.e. high probability to switch (*switchers*). To do so, we adopted the following decision rule. We classified as *stayers* those customers who, during the last period of observation (June, 30 2006), presented a probability to switch to the *post-trial* model lower than 50%. This approach might be questionable and it can be of course improved. For example, one might argue that some customers simply need more time in order to switch, in other words that these customers actually have a very long *trial* period, hence they are not real *stayers*. However, we consider a quite long observation period; its length is 18 quarters, i.e. four years and a half. This gives us support to our classification.

We found a high percentage of *stayers* in our sample (see table 5.11). This evidence is quite interesting because it tells us that there is a large proportion of customers using the same decision strategy over time with a very low probability to experiment distinct stages. This result is supported in the literature. Several authors argue that in many situations consumers use simplified heuristics to make their choices and sometimes simplified decision processes might be observed since the first purchase (Solomon, Bamossy and Askegaard, 2002; Olshavsky and Granbois, 1979; Blackwell, Miniard and Engel 2002). Table 5.11 compares *stayers'* and *switchers'* channel usage. *Stayers* are mainly single-channel-users[20]. This is plausible because the low probability to learn might induce these customers to rely mainly on the same channel usage over time. On the contrary, *switchers* experiment various channels during their purchasing history.

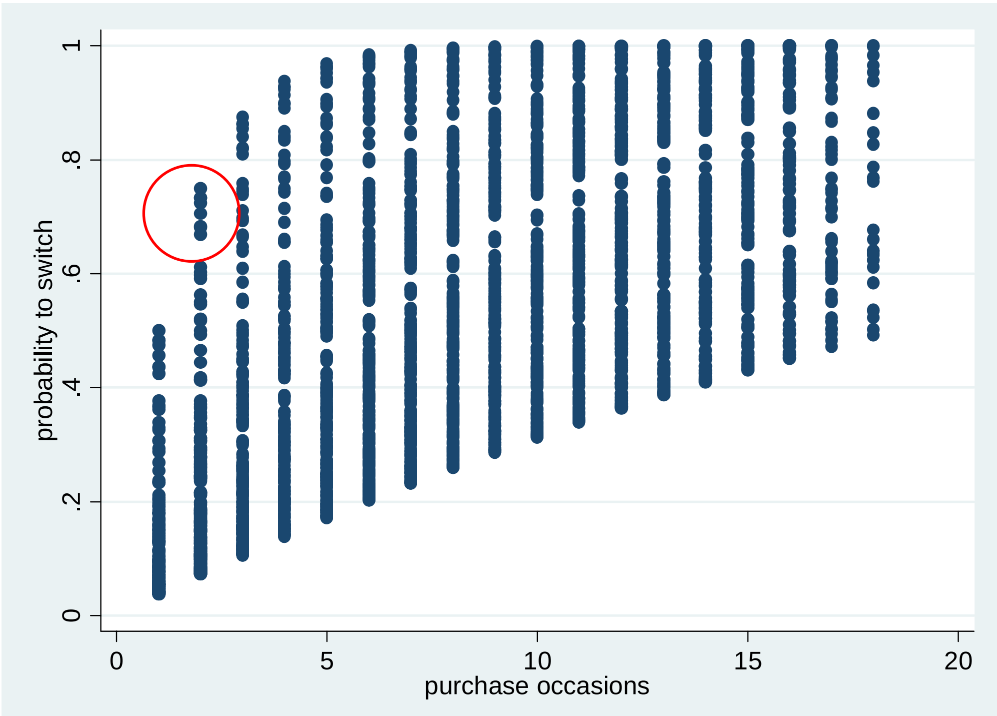
*Table 5.11: Stayers versus Switchers*

| | *Stayers* (787 customers) | *Switchers* (231customers) |
|---|---|---|
| **Channel Usage** | | |
| Mainly Catalog | 36% | 0% |
| Mainly Internet | 1% | - |
| Mainly Store | 55% | 1% |
| Catalog and Store | 4% | 55% |
| Catalog and Internet | 4% | 23% |
| Internet and Store | 0% | 4% |
| Catalog, Internet and Store | 1% | 16% |
| **Multiple-Channel-User** | | |
| Yes | 9% | 99% |
| Two channel Buyer | 8% | 82% |
| Three Channel Buyer | 1% | 16% |
| No | 91% | 1% |

[20] it does not mean that they strictly purchase always using the same channel over time, for example they might experiment other channels in few purchase occasions.

Figure 5.5 shows the scatter plot which represents switching probabilities over purchase occasions for the *switchers*' group. We can notice that there is large variance in these results which means that the length of the *trial* period varies among customers. For example, the red circle in figure 5.5 highlights that some customers have a probability grater than 60% to switch to the *post-trial* model after three purchase occasions.

*Figure5.5: Scatter Plot: Switchers' Probability to Move to the Post-Trial Model*




Given the high variance in the probability to move from the *trial* model among *switchers* over time, we further distinguish *switchers* with a short *trial* period from *switchers* with a long *trial* period[21] with the purpose to show that there are main differences in the

[21] The model considers 18 purchases occasions. We consider *quick stayers* customers for which the *trial* length is less than 9 purchase occasions. We used the same rule, that is we classified as *quick stayers* customers for

length of the *trial* period among *switchers*. We called the former group *quick switchers* and the latter group *long switchers*. Table 5.12 shows some descriptive information about these groups. Interestingly, 48% of the *switchers* have a quick *trial* stage and 52% have a long *trial* stage. This supports our idea of differences in the length of the *trial* period among customers. It is interesting to note that the percentage of three channels users is higher among *quick switchers*. It is also important to remark that customers using mainly the same channel over time might be certainly *switchers*. For example, customer *h* might use the Catalog initially without having achieved a “conscious” channel preference yet; she could maybe try once or twice different channels, but subsequently she might develop preferences for that channel. Hence, we can observe a learning process and different models describing his or her channel choice pattern over time. Actually, we observe only 3% of “switching” single-channel-users and these customers have a long *trial* period. Anyway, we should pay attention while interpreting this result because it is difficult to distinguish *stayers* from *switchers* who use mainly the same channel over time and who have a very long *trial* period.

*Table5.12: Quick Switchers versus Long Switchers*

| | | **Quick Switchers** | **Long Switchers** |
|---|---|---|---|
| *Channel Usage* | | *Freq %* | *Freq %* |
| Catalog-Store | | 49% | 60% |
| Catalog-Internet | | 27% | 20% |
| Internet -Store | | 5% | 4% |
| Three channels | | 20% | 13% |
| Single-Channel-users | | - | 3% |
| Total | 231 | 111 | 120 |
| | 100% | *48%* | *52%* |

which the probability to switch to the *post-trial* period is grater than 50% during the ninth purchase occasion. Similarly, we classified long *switchers* the remaining part of the sample, i.e. the customers for which is more probable that the switching will take place after the ninth purchase occasion.

### 5.4.2 Factors Influencing Learning

Literature on learning by doing, motivation to search and process information indicate several factors that might influence customers' learning. In chapter 3 we have identified several factors which, in this context of analysis, might influence switching behavior. Specifically, in this section we assess the impact of these factors on the probability to be *stayer* or *switcher*. In chapter 3 we have developed several hypotheses about factors' positive or negative influence on the probability to be *stayer*. In table 5.13 we summarize these hypotheses, the variables used, and their operationalization.

*Table5.13: Hypotheses Summary and Variables*

| | HYPOTHESES | | *Probability to be a switcher* | *Variable Name* | *Operalization* |
|---|---|---|---|---|---|
| Familiarity | $H_1$ | Customers with a strong channel familiarity are more likely to be *stayers* | - | Familiarity | Dummy variable which takes value 1 if the proportion of purchases made using the same channel during the earlier purchase occasions is high (i.e. grater than 80%). |
| Customer's motivation to learn | $H_{2a}$ | Early use of the Internet would be associated with switching decision processes | + | Early use of the Internet | Proportion of purchases made using the Internet channel during the earlier purchase occasions. |
| | $H_{2b}$ | Customers whose first channel used was the Internet would be more likely to switch decision processes | + | First choice Internet | Dummy variable which takes value 1 if the customer made his first purchase using internet. |
| | $H_3$ | Customer who provide emails are more likely to switch decision processes | + | Email address | Dummy variable which takes value 1 if the customer provides an email address to the firm |
| Demogr. | $H_4$ | Younger customers are more likely to be *switchers*. | - | Age | Customer's age |

| | | | | | |
|---|---|---|---|---|---|
| Negative experiences | $H_5$ | Consumers who return products are more likely to switch decision processes. | + | Returns | Amount ($) of returns the customers made during the earlier purchase occasions. |
| | $H_6$ | Customer acquisition method should be associated with switching decision process. | +/- | Customers' acquisition strategy | Dummy variable which takes value 1 if the customers where acquired using agents and 0 otherwise. |

To test these hypotheses we performed a discriminant analysis, using as dependent a variable which takes values 1 if the customer was classified as *stayer* and 2 if she was classified as *switcher*.

The discriminant analysis has received much theoretical attention in the marketing literature (Dillon, 1979; Dillon and Schiffman, 1978; Crask and Perreault, 1977; Morrison 1969; Frank, Massey and Morrison,1965).The purpose of a classification of observations of known grouping is merely to see how well the derived function predicts group membership using several factors. Specifically, we estimate a linear discriminant function describing the importance of the independent variables in differentiating observations of known group membership, where the independent variables are the seven variables in table 5.13, i.e. familiarity, early use of the Internet, first choice Internet, email address, age, returns and customers' acquisition strategy.

The results from the two-group multivariate discriminant analysis using the stepwise procedure are reported in table 5.14 and 5.15. The hypotheses tested pertain to the relationship between each independent variable and whether switching occurs. To facilitate evaluation of the relative roles of the seven independent variables, Table 5.14 gives their standardized function coefficients or weights, it also shows their partial F-values, their Wilks' lambda, and their Rao'V. The magnitude of the weights (i.e. standardized function coefficients) represents the relative importance of each variable in the discriminant function.

The result of the discriminant model supports H1 (negative association between strong familiarity with one channel and switching behavior). Customer's motivation to learn as element which enhances the probability to learn is supported by H2b (positive association between first choice the Internet and probability to switch) and H3 (positive association between providing an email address and switching behavior). Finally, H5 supports the idea of a positive association between negative experiences, specifically returns, and the decision to switch decision processes.

The overall effectiveness of discriminant function usually is examined by using two criteria: statistical significance and predictive accuracy. The results for three widely used tests of statistical significance are reported in the last column of table 5.14. The Rao'v total, chi square, and Wilks' lamba all show the discriminant function to be highly significant (p<=.01).

*Table5.14: Two-Group Discriminant Analysis Results*

| | *Function Coeff.* | *Statistic* | | | | |
|---|---|---|---|---|---|---|
| **Variable** | *Stand.* | *Wilks' Lambda* | *F value* | *Rao's V* | *Sig.* | *Total* |
| Familiarity | **-0.405**[a] | 0.895 | 59.8 [a] | 119.7 [a] | 0.00 | Rao's V Total |
| Early use of the Internet | 0.178 | 0.884 | 2.5 | 132.7 | 0.11 | 129.9 [a] |
| First choice Internet | **0.187**[a] | 0.887 | 32.4 [a] | 129.9 [a] | 0.00 | Chi-square |
| Email address | **0.824**[a] | 0.912 | 97.6 [a] | 97.6 [a] | 0.00 | 122.0 [a] |
| Age | -0.051 | 0.886 | 0.5 | 130.5 | 0.46 | Wilks' Lambda |
| Returns | **0.211**[a] | 0.890 | 41.7 [a] | 125.3 [a] | 0.00 | 0.887 [a] |
| Customers' acquisition strategy | 0.113 | 0.886 | 1.2 | 131.2 | 0.27 | |

*a Significant at p<0.01 level*
*b Significant at p<0.05 level*

The results for predictive accuracy are reported in the confusion matrix in Table 5.15. The classification accuracy is 79% and it greatly exceeds the proportional chance criterion (Aaker 1971). It should be note that the discriminant function classifies *stayers* more precisely than *switchers*.

*Table 5.15: Two-Group Discriminant Analysis Classification Matrices (all numbers are percentages)*

Classification Results(a)

| | | | Predicted Group Membership | | Total |
|---|---|---|---|---|---|
| | | | *switchers* | *stayers* | |
| Original | *Percent* | *switchers* | 30.3% | 69.7% | 100% |
| | | *stayers* | 6.5% | 93.5% | 100% |

a 79.2% of original grouped cases correctly classified.

The parameter estimates for the best model (M3) are presented in Table 3 (bold indicates that the 95% posterior interval excludes zero).

### 5.4.3 Multinomial Logit Parameter Estimates

As discussed in chapter four we distinguish between two different channel utilities and we model them following the classical multinomial logit structure. $U_{0hjt}$ represents the utility of choosing channel *j* in the *trial* period at *trial* purchase occasion *t* and $U_{1hjt}$ is the utility of choosing channel *j* during *post-trial* purchase occasion *t'*. These utilities jointly contribute to estimate the probability to select channel j during the purchase occasion t. Specifically, we suppose, for example, that the customer h has a probability to be in the *trial* period during his third purchase occasion of 0.6, therefore to be in the *post-trial* stage in the same purchase occasion of 0.40. Then, the probability that the customer *h* selects channel j during purchase occasion three is estimated accounting for the 60% the *trial* utility and for 40% the *post-trial* utility. Four types of elements where considered in the utility functions: unobserved customers preferences, catalog sent, email sent and state dependence.

Both the catalog sent (i.e. the number of catalogs that customer *h* received at purchase occasion *t*) and email sent (i.e. the number of emails that customer *h* received at purchase occasion *t*) included in our channel choice model might be determined by firm's marketing strategy variables and customers' demographic profile (see paragraph 4.2.2). This may cause an endogeneity bias. Similarly to Gonul, Kim and Shi (2000), we use a two-stage least-squares approach to minimize this bias by using an instrument instead of the actual value of the catalog sent and email sent variables (see appendix 3 for the results).

The parameter estimates for the best model (M3) are presented in Table 5.16.

*Table 5.16: Parameter Estimates Multinomial Logit Channel Selection Switching Model*

| **Channel Choice Model Results**[a] | | | | | |
|---|---|---|---|---|---|
| | | **Catalog vs Store** | | **Internet vs Store** | |
| | | *Coefficients (sd)* [b] | *Elasticity*[c] | *Coefficients (sd)* [b] | *Elasticity*[c] |
| Intercept | *Trial* | 0.85 (0.26) *[-0.2; 1.8]* | - | **-2.48** (0.39) *[-3.3; -1.8]* | - |
| | *Post-trial* | 0.38 (0.92) *[-1.0; 1.4]* | - | **-3.78** (0.46) *[-5.2; -2.3]* | - |
| Catalog sent | *Trial* | **-1.45** (0.36) *[-2.4; -0.5]* | ***-7.24*** | **-9.10** (0.97) *[-11.1; -7.5]* | ***-20.56*** |
| | *Post-trial* | **-2.12** (0.63) *[-3.1; -1.2]* | ***-0.33*** | 0.63 (0.59) *[-0.5; 1.8]* | *1.83* |
| Email sent | *Trial* | -0.03 (0.61) *[-1.2; 0.9]* | *-0.46* | **3.28** (0.75) *[2.3; 4.5]* | ***4.90*** |
| | *Post-trial* | **2.35** (0.34) | ***0.93*** | 0.41 (0.70) | *-0.49* |

| | | | |
|---|---|---|---|
| | | *[1.4; 3.5]* | *[-0.6; 1.0]* |
| State Dependence | *Trial* | | **4.09** (0.60) |
| | | | *[3.1; 5.1]* |
| | *Post-trial* | | **3.14** (0.62) |
| | | | *[2.5; 3.9]* |

*[a] [;] represents the 95% interval. Bold indicates that it exclude zero*
*[b] A positive coefficient means that a customer is more likely to choose channel j than base channel. The base channel is store*
*[c] we computed elasticities at the mean value of the continuous variables and the modal of the categorical variables*

The intercept estimates suggests a preference for the use of the Store over the Internet in the *trial* period which is reinforced in the post *trial* period. Interestingly, the intercept about Catalog over Store is not significant (by significant we mean the 95% posterior confidence interval excludes zero), suggesting that on average we cannot distinguish a stronger preference for the use of the Catalog over Store and vice versa.

Table 5.16 also shows how the customers (on average) respond to direct marketing communications. It reports both the parameter estimates and the elasticities. The computed elasticities measure the percentage change in the probability of choosing a particular channel alternative to a percentage increase in catalog sent and email sent (Louvier et al., 2000; Frances and Paap, 2001; Little and Guadagni,1983)[22].

The effect of catalog sent is always significant in the *trial* period, and in particular it has a positive association with store selection, suggesting that catalog sent "promote" the use of stores. This effect is particular strong in the choice of the Store over Internet. In the *post-*

[22] Specifically, we computed direct marketing communication elasticities as follow (see Little and Guadagni, appendix 2):$\varepsilon_k$=b*X*(1-$m_k$) where $\varepsilon$ represents the elasticity, $b_k$ is the coefficient of direct marketing communication, X represents the average direct marketing communication sent and $m_k$ the expected share of channel k. We computed elasticity at individual level, hence we assess the impact of a change in direct marketing communication on each customer response outcome. In table 5.18 we report the average elasticities.

*trial* period the magnitude of the catalog sent effect strongly decrease for the choice of the Internet over Store. In this case the effect of catalog sent is not significant suggesting that this direct marketing communication strategy is unable in moving customers from the Internet to the Store. However, the effect of catalog sent is still significant in *post-trial* for the choice of the Store over Catalog.

Emails do not influence the choice of the Catalog over the Store in the *trial* period. However, they increase the probability of the use of Internet over Store in the *trial*. It is interesting to note that the effect of emails is reversed in the *post-trial* period, i.e. they significantly affect the choice of the Catalog over the Internet but they are no more effective in increasing the probability to choose the Internet over Store.

Finally, it is very interesting to note that state dependence effect is reduced in the *post-trial* period suggesting that, over time, people get set in their ways. We noticed that preferences for the use of the Store over the Internet in the *trial* period are reinforced in the *post-trial* period. The increasing of individual intercepts effect combined with the decreasing of the state dependence parameter in the *post-trial* stage might suggest that channel choice is guided by channel preferences and it is less inertial.

My results show first the parameter estimates differ from *trial* and post *trial* phase, suggesting that customers change decision process over time. Second, it can be signed that the decision process change for two main reasons: i) channel preferences become more important over time, and they seem to drive channel choice, ii) marketing communication effect differs in the *trial* and *post-trial* stages suggesting that marketing communication influences the channel migration process. This implies that different decision strategies could characterize customers' channel choice process over time. Therefore, our next step will be to segment customers depending on the different decision strategies that might take place, in order to assess if migrations among different decision pattern over time occur.

### 5.4.4 How Channel Decision Patterns Evolve

We estimated the multinomial channel selection switching model at individual level; therefore for each customer in the data set we have individual parameter estimates. This model accounts for both heterogeneity in individual channel preferences and state dependence explanations of the observe persistence in channel choices over time, so it disentangles the contribution of each. This allows us to study different types of decision processes basing on these results and to observe how they evolve over time. In chapter 3 we argue that we can distinguish between two main types of decision strategies, specifically we can differentiate inertial versus preference-based decision-making strategies describing customers' channel choice. In particular, comparing channel preferences (i.e. individual random intercepts) with state dependence we can delineate two different types of decision strategies: preference-based versus inertial. A high and positive state dependence can be read as inertial behavior (Seetharaman et al. 1999). Therefore, we argue that if channel preference is greater than a positive state dependence we can suppose that a preference-based decision making takes place, in other words customers are committed in channel choice task and they consciously choose the channel that they prefer. Otherwise, an inertial decision making could be observe which means that customers rely their channel decisions on the previous channel chosen and they do not exhibit a strong commitment in this choice task. This distinction is of fundamental importance in marketing. Consider, for example, the decision to advertise a particular channel. If the true model of customer behavior is preference based and it is not inertial, such marketing communication will increase the probability of its channel choice only while it is in effect. If a strong inertia is present, some customers who choose the advertised channel will be persuaded to stay with the channel choice after the advertising period ends. Thus, as Keane (1997) point out a cost/benefit analysis of the marketing strategies will depend critically on the assumed forms of heterogeneity and state dependence, i.e. on the different types of decision strategies used in the considered population. For this reason we account for the role of marketing and we develop a general framework which takes into account that marketing indifferently might be effective or not in inertial or preference based situations. Basing on this we depict four possible conditions which discriminate four different customers' decision styles:

1) pattern 1: preference based decision making – high marketing responsiveness,
2) pattern 2: preference based decision making – low marketing responsiveness
3) pattern 3: inertial decision making– high marketing responsiveness,
4) pattern 4: inertial decision making – low marketing responsiveness.

In order to identify these patters we used three classes of parameters estimates: i) individual level intercepts that we can interpret as channel preferences, ii) individual level state dependence parameters, iii) direct marketing communications elasticities.

By comparing individual channel preferences with state dependence we can classify customers into preference-based or inertial groups. Specifically, we used the following rule to classify customers: if customers *h* exhibits a positive and high state dependence, higher than his or her preference for catalog over store or internet over store (in absolute values), we classify this customer as inertial (see table 5.17 for details).

*Table 5.17: Preference-Based versus Inertial Decision Strategy*

| Conditions | *Preference- based Decision Strategy* | | | *Inertial Decision Strategy* | | |
|---|---|---|---|---|---|---|
| 1 | State Dependence | < | Preference (Catalog versus Store) | State Dependence | > | Preference (Catalog versus Store) |
| | State Dependence | < | Preference (Internet versus Store) | State Dependence | > | Preference (Internet versus Store) |
| 2 | State Dependence | > | Preference (Catalog versus Store) | | | |
| | State Dependence | < | Preference (Internet versus Store) | | | |
| 3 | State Dependence | < | Preference (Catalog versus Store) | | | |
| | State Dependence | > | Preference (Internet versus Store) | | | |

After having classified consumers into preference-based or inertial groups we further distinguish them basing on their marketing responsiveness. In order to do so, we consider direct marketing communications elasticities. We compared individual marketing elasticities with their median value and we classify customers into high marketing responsiveness group if emails or catalog sent elasticities are greater than the respective median elasticities (see table 5.18 for details).

*Table 5.18: High Marketing versus Low Marketing Responsiveness*

| Conditions[a] | High marketing responsiveness[b] | | | Low marketing responsiveness | | |
|---|---|---|---|---|---|---|
| 1 | Mkt communications elasticity (C versus S) | > | Median Mkt communications elasticity (C versus S) | Mkt communications elasticity (C versus S) | < | Median Mkt communications elasticity (C versus S) |
| | Mkt communications elasticity (I versus S) | > | Median Mkt communications elasticity (I versus S) | Mkt communications sent elasticity (I versus S) | < | Median Mkt communications elasticity (I versus S) |
| 3 | Mkt communications elasticity (C versus S) | < | Median Mkt communications elasticity (C versus S) | | | |
| | Mkt communications elasticity (I versus S) | > | Median Mkt communications elasticity (I versus S) | | | |
| 3 | Mkt communications elasticity (C versus S) | > | Median Mkt communications elasticity (C versus S) | | | |
| | Mkt communications elasticity (I versus S) | < | Median Mkt communications elasticity (I versus S) | | | |

*a actually we consider two types of direct marketing communications (email sent and catalog sent), for each condition we evaluate four sub-conditions. Therefore we have a total of 16 possible outcomes. We classify customers as low responsive to marketing only if all their marketing elasticities are less then their median values (or are not significant).*
*b C stands for Catalog, I for Internet and S for Store*

In chapter 3 we mention that these four patterns can describe both the *stayers* and *switchers'* behavior. The only difference is that *stayers* remain always with the selected pattern; on the contrary, *switchers* change their decision pattern over time. Specifically, the change could pertain the marketing sensitiveness or the underling decision strategy (i.e. preference-based versus inertial). We start briefly describing *stayers'* behavior, and we present the results for the *switchers*.

Table 5.19 shows some descriptive statistics about *stayers*. Interestingly, *stayers*' behavior in this context can be described mainly by pattern 3. Essentially, 98% of *stayers* present an inertial behavior and high marketing sensitiveness. For these customers inertia plays a critical role in explaining channel choices, and at the same time marketing has a positive impact. For example, we can think of a consumption situation in which a customer does not pay attention to the channel choice, he or she may desire to simplify this decision task. A positive effect of marketing in this situation, as we highlighted in chapter 3, is plausible because it may serves as cue to reinforce channel choice. The presence of a large number of inertial customers is supported by the literature which shows that in many situations consumers use simplified heuristics to make their channel choices. Additionally it has been shown that sometimes simplified decision processes might be also observed since the first purchase (Granbois, 1977; Olshavsky and Granbois, 1979; Balasubramanian et al., 2005). The evidence that customers are responsive to marketing is important as inertial customers can be induced to try different channels more easily than customers using a preference-based decision strategy.

*Table5.19: Stayers Decision Patterns*

| | Preference based vs Inertial | H vs L mkt | Sample Size (%) | Average emails received over relation ship | Average Catalogs received over relation ship | Channel Usage | Sample Size (%) |
|---|---|---|---|---|---|---|---|
| pattern 1 | Preference | H | 2% | 36.9 | 19.0 | Two Channels | 88% |
| | | | | | | Three Channels | 13% |
| pattern 3 | inertial | H | 98% | 34.9 | 5.3 | Catalog | 39% |
| | | | | | | Internet | 1% |
| | | | | | | Store | 60% |
| | | | | | | Three Channels | 1% |

The four patters described can depict *switchers*' channel decision making as well. For *switchers* we can observe evolving patters between *trial* and *post-trial* stages. In chapter 3 we have identified twelve possible *trial* / *post-trial* combinations. *Switchers* can go from preference-based to inertial decision making, from inertial decision making to preference based, or stays in their current decision making but just change their marketing sensitiveness (see figure 3.1 in chapter 3). Therefore, for *switchers* we can identify four *trial* decision patterns and four *post-trial* decision patterns. For example, customer *h* behavior could be well represented by patter 1 during the *trial* stage, however pattern 3 could describe better his or her behavior during the *post-trial* stage which means that the customer *h* migrates from pattern 1 to pattern 3 over time. In other words, this customer maintains a high marketing responsiveness over time, but he or she switches from a preference-based to an inertial channel choice behavior.

Our purpose is to map all the possible customers' migrations which take place in our data set. To achieve this goal, we proceed as follow: first, using the *trial* multinomial logit parameter estimates we verify if customers fit into the four "*trial*" patterns, second, by using the *post-trial* multinomial logit parameter estimates we segment the customers into the four "*post-trial*" patterns, finally we map the migrations among the *trial* and *post-trial* patters.

We initially considered four possible *trial* patterns, but we found that in our sample *switchers* use only two of them during the *trial* period. Specifically, *switchers* are mainly inertial in their channel choice in the *trial* period, 94% are classified in pattern 3 (see table 5.20). However, a small percentage (6%) starts using pattern 1, i.e. high marketing responsiveness and preference based channel decision making. This small group of customers exhibit "conscious" channel preferences since the beginning. They are younger than pattern 2 *switchers* (see table 5.20). On average their amount of returns is smaller. This makes sense considering that they exhibit channel preferences.

*Table5.20: Switchers Trial Patterns*

| *Descriptive Information* | *Trial Patterns* | | | |
|---|---|---|---|---|
| | **Trial Pattern 1** | **Trial Pattern 2** | **Trial Pattern 3** | **Trial Pattern 4** |
| Sample Size<br>(%) | 13<br>(6%) | 0 | 218<br>(94%) | 0 |
| | *mean per customer* | | | |
| Age | 33.23 | - | 41.48 | - |
| Amount spent per purchase occasion during *trial* | € 23.93 | - | € 28.11 | - |
| Items purchased per purchase occasion during *trial* | 2.14 | - | 2.60 | - |
| Amount returns per purchase occasion during *trial* | € 0.22 | - | € 1.10 | - |
| Catalogs received per purchase occasion during *trial* | 1.98 | - | 1.83 | - |
| Emails received per purchase occasion during *trial* | 0.55 | - | 1.19 | - |
| Amount shipping costs per purchase occasion during *trial* | € 2.09 | - | € 2.35 | - |

Table 5.21 and 5.22 show descriptive statistics about channel choices made by these groups of customers. Pattern 1 *switchers* and Pattern 3 *switchers* do not differentiate in terms of the proportion of choices made using the Catalog, the Internet or the Store (see Table 5.21). However, we can notice some differences in their combination of channel used (Table 5.22). For example, *trial* pattern 3 presents roughly 26% of *switchers* who use mainly only one channel during the *trial* period. However, we should pay attention to make comparisons between these two groups of "*trial switchers*" because pattern 1 represents only 6% of the sample. Nevertheless, this result, combined with the large amount of inertial *stayers*, is interesting because it demonstrates that at the beginning a large number of customer is not particularly committed with channel choice, their channel choice decision are mainly inertial and we could even think that if "nothing" happen (e.g. negative experiences, marketing, etc.)

they could stay in their pattern for a long time and they could be very slow in developing "conscious" channel preferences.

*Table5.21: Switchers Trial Patterns Channel Proportion Statistics*

| *Channel selection descriptive information* | *Trial Patterns* | | | |
|---|---|---|---|---|
| | **Trial Pattern 1** | **Trial Pattern 2** | **Trial Pattern 3** | **Trial Pattern 4** |
| | *mean per customer* | | | |
| Catalog choices proportion | 0.47 | - | 0.48 | - |
| Internet choices proportion | 0.15 | - | 0.15 | - |
| Store choices proportion | 0.38 | - | 0.36 | - |

*Table5.22: Switchers Trial Patterns distinct by Channel Usage*

| *Channel Usage* | *Trial Patterns* | | | |
|---|---|---|---|---|
| | **Trial Pattern 1** | **Trial Pattern 2** | **Trial Pattern 3** | **Trial Pattern 4** |
| Mainly Catalog | - | - | 9.63% | - |
| Mainly Internet | - | - | 3.67% | - |
| Mainly Store | - | - | 12.39% | - |
| Catalog and Store | 69.23% | - | 40.37% | - |
| Catalog and Internet | - | - | 23.85% | - |
| Internet and Store | 7.69% | - | 4.13% | - |
| Catalog, Internet and Store | 23.08% | - | 5.96% | - |

Tables 5.23 and 5.24 report the results of the multinomial logit models of the *trial* stage. Results are similar in terms of marketing and intercept despite the state dependence parameter which is higher and significant for customers in pattern 3.

*Table5.23: Parameter Estimates Trial Multinomial Logit Pattern 1*

| **Channel Choice Model Results**[a] | **TRIAL PATTERN 1** | | | |
|---|---|---|---|---|
| | **Catalog vs Store** | | **Internet vs Store** | |
| | **Coefficients (sd)** [b] | **Elasticity**[c] | **Coefficients (sd)** [b] | **Elasticity**[c] |
| Intercept | **1.10** | - | **-3.55** | - |
| Catalog sent | -0.70 | -3.19 | **-10.81** | **-19.21** |
| Email sent | -1.28 | -0.44 | **8.97** | **4.64** |
| State Dependence | 2.12 | | | |

*[a] We have these parameters estimates at individual level. Here we represent the overall average estimates.*

*[b] A positive coefficient means that a customer is more likely to choose channel j than base channel. The base channel is store*

*[c] we computed elasticities at the mean value of the continuous variables and the modal of the categorical variables*

*Table5.24: Parameter Estimates Trial Multinomial Logit Pattern 3*

| **Channel Choice Model Results**[a] | **TRIAL PATTERN 3** | | | |
|---|---|---|---|---|
| | **Catalog vs Store** | | **Internet vs Store** | |
| | **Coefficients (sd)** [b] | **Elasticity**[c] | **Coefficients (sd)** [b] | **Elasticity**[c] |
| Intercept | 1.24 | - | **-3.08** | - |
| Catalog sent | -0.35 | -4.52 | **-10.10** | **-18.68** |
| Email sent | -1.11 | -0.42 | **9.15** | **4.68** |
| State Dependence | **5.09** | | | |

*[a] We have these parameters estimates at individual level. Here we represent the overall average estimates.*

*[b] A positive coefficient means that a customer is more likely to choose channel j than base channel. The base channel is store*

*[c] we computed elasticities at the mean value of the continuous variables and the modal of the categorical variables*

We identified four possible *post-trial* patterns. We found that in our sample *switchers* use three of them: pattern 1 (39%), pattern 2 (60%) and pattern 3 (1%). Table 5.25 shows some descriptive information about these *post-trial* patterns. Specifically, *switchers* in the *post-trial* stages mainly use a preference-based decision strategy to select channels. Actually almost all the customers use a preference-based strategy (99%). This is plausible remembering that *switchers* are learning prone customers and that the learning phase might be important to develop "conscious" channel preferences. For customers in pattern 1 marketing remains an important instrument and it reinforces and affects channel choice. However, for customers in pattern 2 the effect of direct marketing communications decreases. A very small percentage of *switchers* are classified in pattern 4, i.e. inertial strategy and low marketing responsiveness. The most profitable *post-trial* pattern seems to be pattern 2 which presents a higher average amount spent (€) per purchase occasion and a higher average quantity purchased.

*Table5.25: Post-Trial Patterns Descriptive Information*

| *Descriptive Information* | *Post-Trial Patterns* | | | |
|---|---|---|---|---|
| | **Post Pattern 1** | **Post Pattern 2** | **Post Pattern 3** | **Post Pattern 4** |
| Sample Size<br>(%) | 90<br>(39%) | 138<br>(60%) | 0 | 3<br>(1%) |
| | *mean per customer* | | | |
| Age | 41.27 | 40.77 | - | 44.67 |
| Amount spent per purchase occasion during *post-trial* | € 25.94 | € 30.59 | - | € 29.50 |
| Items purchased per purchase occasion during *post-trial* | 2.45 | 3.02 | - | 1.95 |
| Amount returns per purchase occasion during *post-trial* | € 0.56 | € 0.56 | - | € 0.00 |
| Catalogs received per purchase occasion during *post-trial* | 1.93 | 1.95 | - | 1.96 |
| Emails received per purchase occasion during *post-trial* | 3.12 | 1.75 | - | 0.60 |
| Amount shipping costs per purchase occasion during *post-trial* | € 2.40 | € 2.02 | - | € 4.76 |

Table 5.26 and 5.27 show descriptive statistics about channel choices made by these groups of customers during the *post-trial* period. Interestingly, pattern 1 presents a high percentage of customers who use jointly catalog and internet (44%) and also a higher percentage of three-channel-users (13%) compared to the three channels users percentage of pattern 2 (3.6%). By contrast pattern 2 presents an high percentage of customers who use jointly catalog and stores (44%). It also exhibits a high percentage of catalog-internet users (25%) and of mainly store users (19%).

*Table5.26: Switchers Post-Trial Patterns Channel Proportion Statistics*

| *Channel selection descriptive information* | *Post-Trial Patterns* | | | |
|---|---|---|---|---|
| | **Post Pattern 1** | **Post Pattern 2** | **Post Pattern 3** | **Post Pattern 4** |
| | *mean per customer* | | | |
| Catalog choices proportion | 0.45 | 0.40 | - | 0.73 |
| Internet choices proportion | 0.23 | 0.12 | - | 0.28 |
| Store choices proportion | 0.32 | 0.49 | - | 0.00 |

*Table5.27: Switchers Post-Trial Patterns distinct by Channel Usage*

| *Channel Usage* | *Post-Trial Patterns* | | | |
|---|---|---|---|---|
| | **Post Pattern 1** | **Post Pattern 2** | **Post Pattern 3** | **Post Pattern 4** |
| Mainly Catalog | - | 0.72% | - | - |
| Mainly Internet | 7.78% | 4.35% | - | - |
| Mainly Store | 14.44% | 18.84% | - | - |
| Catalog and Store | 17.78% | 44.2% | - | - |
| Catalog and Internet | 44.44% | 25.36% | - | 100% |
| Internet and Store | 2.22% | 2.9% | - | - |
| Catalog, Internet and Store | 13.33% | 3.62% | - | - |

Tables 5.28, 5.29 and 5.30 report the results of the multinomial logit models of the *post-trial* stage. Pattern 1 presents a low state dependence and significant intercepts. Interestingly, this group of customers is responsive to direct marketing communications. In particular catalog sent bring these customers to the Internet over the Store. Therefore, it seems that the company catalogs sending strategy advertises the use of the Internet. On the contrary emails significantly increase the probability of the use of the Catalog over the Store. Both catalog sent and emails significantly affect the use of channel which presents a different "technology" (e.g. we can assert that emails and the Internet or catalogs and the Catalog share the same "technology"). This result is interesting, it could be, for example, the receiving of an email or of a catalog trigger customer attention and induce him or her to select his or her favorite channel/s. This makes sense remembering that these customers have already developed "conscious" channel preferences. Marketing may reinforce them.

On the contrary, Pattern 2 customers are not responsive to marketing. They exhibit definite channel preferences; hence marketing communications do not affect channel choice.

Finally, the last group (pattern 3) represents a very small percentage of the sample (1%), for this reason one should interpret these results with caution. However, we can observe that this customers' channel behavior is inertial, they do not exhibit strong channel preferences and marketing does not have an impact on channel choice.

*Table5.28: Parameter Estimates Post-Trial Multinomial Logit Pattern 1*

| **Channel Choice Model Results**[a] | **POST-TRIAL PATTERN 1** | | | |
|---|---|---|---|---|
| | **Catalog vs Store** | | **Internet vs Store** | |
| | **Coefficients (sd)** [b] | **Elasticity**[c] | **Coefficients (sd)** [b] | **Elasticity**[c] |
| Intercept | **-2.93** | - | **-2.77** | - |
| Catalog sent | -0.39 | -0.66 | **1.32** | **1.58** |
| Email sent | **2.99** | **2.05** | -0.94 | -0.40 |
| State Dependence | | **2.26** | | |

*Table5.29: Parameter Estimates Post-Trial Multinomial Logit Pattern 2*

| **Channel Choice Model Results**[a] | **POST-TRIAL PATTERN 2** | | | |
|---|---|---|---|---|
| | **Catalog vs Store** | | **Internet vs Store** | |
| | **Coefficients (sd)** [b] | **Elasticity**[c] | **Coefficients (sd)** [b] | **Elasticity**[c] |
| Intercept | -1.37 | - | **-3.94** | - |
| Catalog sent | -0.13 | -0.23 | 0.11 | 0.76 |
| Email sent | 1.05 | 0.81 | -0.94 | -0.49 |
| State Dependence | **2.54** | | | |

*Table5.30: Parameter Estimates Post-Trial Multinomial Logit Pattern 4*

| **Channel Choice Model Results**[a] | **POST-TRIAL PATTERN 4** | | | |
|---|---|---|---|---|
| | **Catalog vs Store** | | **Internet vs Store** | |
| | **Coefficients (sd)** [b] | **Elasticity**[c] | **Coefficients (sd)** [b] | **Elasticity**[c] |
| Intercept | 1.36 | - | 1.54 | - |
| Catalog sent | 0.04 | -0.01 | 1.03 | 0.62 |
| Email sent | 1.27 | 0.67 | -0.89 | -0.33 |
| State Dependence | **2.70** | | | |

[a] *We have these parameters estimates at individual level. Here we represent the overall average estimates.*
[b] *A positive coefficient means that a customer is more likely to choose channel j than base channel. The base channel is store*
[c] *we computed elasticities at the mean value of the continuous variables and the modal of the categorical variables*

These results demonstrate that customers' channel choice behavior can be represented by different decision patterns. They also demonstrate that customers with a learning proneness might revise their decision patters over time and that *trial* decision patterns are not the same to *post-trial* decision patters. Specifically, we found that the *trial* period is mainly characterized by customers using an inertial-based decision strategy, highly responsive to direct marketing stimuli. By contrast, the *post-trial* stage is mainly characterized by preference-based decision strategies, marketing could be both effective or not on channel choice.

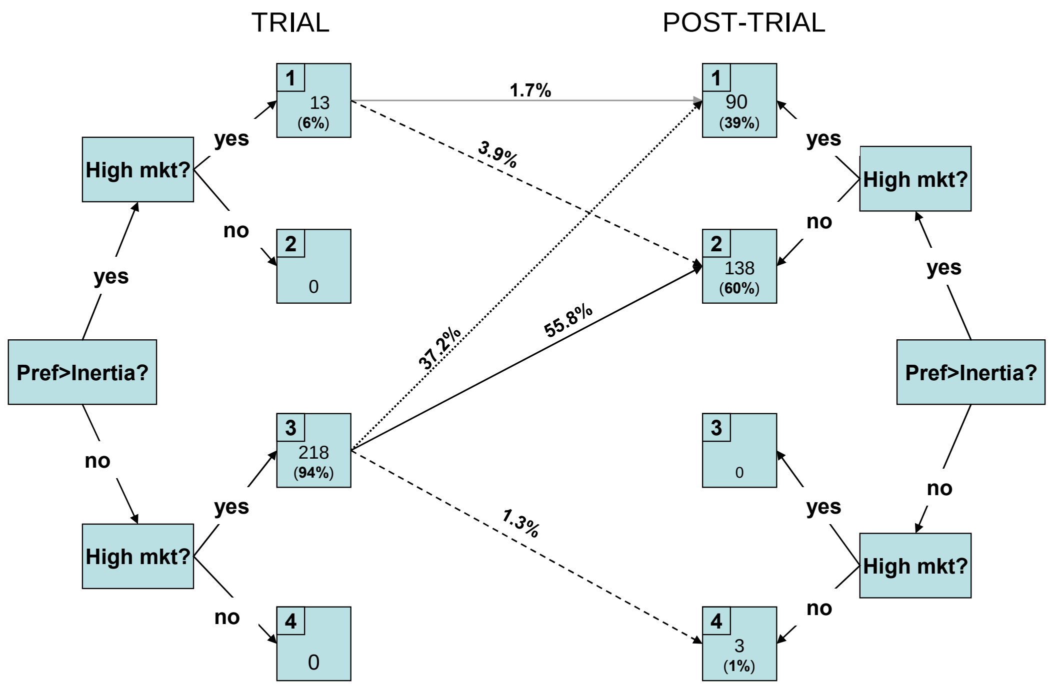
As final step of our analysis we map all the migrations which take place among *trial* and *post-trial* decision patterns. In this way, we can observe how the generic customer h moves from his or her trail-pattern to his or her *post-trial* pattern. Figure 5.6 shows these migrations.

The key finding of this analysis is the demonstration that customers switch decision pattern over time, in other words their channel decision-making strategy does not remain the same and it is possible to observe different types of migrations patterns. In particular figure 5.6 highlight two main types of migrations. The migration from *trial* pattern 3 to *post-trial* pattern 1 (37%) and the migration from *trial* pattern 3 to *post-trial* pattern 2 (56%). Both depict one interesting results: customers switch from an inertial-based decision strategy to a preference-based decision strategy, in other words customers at the beginning do not seems to be committed with the channel choice task, it might be that they not consider at the beginning the channel choice as an important choice and they simply choose the channel/s used in the previous purchase occasions. However, these customers are learning prone, in addition they are responsive to marketing stimuli; hence the acquisition of external information (e.g. marketing) and/or some events (e.g. negative experiences) over time might induce them to learn on the channels and develop "conscious" channel preferences.

The main difference in these two types of migration is the role of direct marketing communications. In the former (migration from 3 to 1) marketing maintains its effectiveness over time and it still has an impact on channel choice. However, it is interesting to note that its impact is not the same. In *trial* pattern 3 (see table 5.27) catalog sent strongly "promote" the use of store over the Internet and email sent push customers towards the use of Internet over Store. In *post-trial* pattern 1 (see table 5.31) catalog sent have a positive impact on the choice of the Internet over store and email sent on the choice of the Catalog over Store. In the latter (i.e. the migration from 3 to 2) marketing effect decreases over time. Therefore, customers switch from an inertial decision making strategy where marketing reinforces channel choices to a preference-based channel decision strategy where channel decisions are mainly driven by preferences.

Figure 5.6 shows other minor migrations patterns. In particular migration from *trial* decision pattern 1 to *post-trial* decision pattern 2 (4%) is interesting because it shows that high marketing responsive customers who reveal "conscious" channel preferences since the beginning gradually move to the same decision pattern (i.e. preference-based) where marketing does not influence channel choice, in other words channel preferences are established and marketing has no more a role on the channel decision process of these customers.

*Figure5.6: Migrations from trail to post-trail among different channel decision making patters*



Finally, Table 5.31 shows some descriptive information about customer *trial* period's length and channel usage grouping *switchers* by type of migration patterns. Interestingly, customers who switch from *trial* pattern 3 to *post-trial* pattern 1 on average have a shorter *trial* period which means that 3_1 migration takes place more rapidly than migration of the type 1_2 or 3_2. It is interesting to note that 13% of customers who migrate from patter 3 to pattern 1 used the Catalog during the *trial* period but then in the *post-trial* stage there is not a group of customers who mainly use catalog. On the contrary we can observe that the percentage of customers who mainly use the store is increased (from 5% to 15%) and also the percentage of customers using mainly the Internet (from 2% to 7%). Therefore the customers who mainly use the catalog during the *trial* period migrate to other channels or to other combination of channels.

*Table5.31: Migrations Descriptive Information*

| *Code*[a] | *Sample percentage* | *Trial Length Purchase Occasions* | *Channel Usage distinct by trial and post-trial stages* | *trial* | *post-trial* |
|---|---|---|---|---|---|
| 1_2 | *4%* | 8.2 | Mainly Catalog | - | - |
| | | | Mainly Internet | - | 22.2% |
| | | | Mainly Store | - | 22.2% |
| | | | Catalog and Store | 77.8% | 55.6% |
| | | | Catalog and Internet | - | - |
| | | | Internet and Store | 11.1% | - |
| | | | Catalog, Internet and Store | 11.1% | - |
| 3_1 | *37%* | 6 | Mainly Catalog | 12.8% | - |
| | | | Mainly Internet | 2.3% | 7.0% |
| | | | Mainly Store | 4.6% | 15.1% |
| | | | Catalog and Store | 22.1% | 17.4% |
| | | | Catalog and Internet | 39.5% | 46.5% |
| | | | Internet and Store | 4.6% | 2.3% |
| | | | Catalog, Internet and Store | 13.9% | 11.6% |
| 3_2 | 56% | 8.5 | Mainly Catalog | 7.0% | 0.8% |
| | | | Mainly Internet | 4.6% | 3.% |
| | | | Mainly Store | 17.8% | 18.6% |
| | | | Catalog and Store | 53.5% | 43.4% |
| | | | Catalog and Internet | 12.4% | 27.1% |
| | | | Internet and Store | 3.9% | 3.% |
| | | | Catalog, Internet and Store | 0.8% | 3.9% |

*a We describe only migration patterns which represents more that 3% of the total*

# *Chapter 6*

# CONCLUSIONS

## 6.1 General Conclusion

This research was triggered by an emergent trend in customer behavior: customers have rapidly expanded their channel experiences and preferences beyond traditional channels (such as stores) and they expect the company with which they do business to have a presence on all these channels. This evidence has produced an increasing interest in multichannel customer behavior and it has motivated several researchers to study the customers' channel choices dynamics in multichannel environment.

This dissertation is positioned on the customer channel "migration" process literature. In particular, we analyzed how the channel decision process of newly acquired customer evolves over purchase occasions. In this field several authors (Ansari, Mela and Neslin, 2008; Thomas and Sullivan, 2005; Knox, 2005; Venketesan and Kumar, 2007) modeled customer channel migration, in other words they recently start to study how channel choice evolves over time. However, limited effort has been made to investigate the learning process per se, i.e., how customers' decision process changes over time as they learn their preferences and become familiar with the firm's marketing activities.

The main contribution of this dissertation can be disentangled in three main aspects: i) the modeling approach contribution ii) the theoretical contribution, and iii) the managerial contribution.

The modeling approach is new and represents one of the first attempts to investigate how a learning phase has an impact on the development of customers channel decision strategies. This represents the first work in the channel choice literature and in particular in the context of the channel migration models which hypothesizes the existence of two distinct stages in the evolution of customers' channel decisions over time (*trial* sand a *post-trial* stages). We developed a modeling approach which allowed us: first to estimate the probability that the customer is learning prone (i.e. the probability the customer switches to a *post-trial* model), secondly to estimate how many *trial* purchase occasions customers need to go through before switching to the *post-trial* stage model, and, finall, to distinguish between two different channel choice utilities. The first governs the *trial* stage and the second the *post-trial* stage. In this way, we obtained two distinct set of parameters, one set for the *trial* stage and the other for the *post-trial* stage. We estimated individual level parameters ending up with more than 15,000 parameter estimates.

Five key aspects define the original theoretical contribution of this dissertation. We summarize these aspects below:

1) *Some people switch channel decision processes while others don't*. We argued that not all the customers have the same probability to switch channel decision process. We supposed that a geometric distribution represents the time (i.e. number of purchase occasions) to switch and we estimated for each customer the probability that he or she changes decision process over time. We argued that some customers might have a very low probability to switch. This might happen if customers are not "committed" with the channel choice task and they do not exhibit a learning proneness. Using the parameter estimates which arise from the geometric model we classified customers as: *switchers* or *stayers*. We called *stayers* the customers with a low or null probability to switch, and *switchers* the customers with a high probability of changing their decision process, i.e. high probability to switch.
2) *Some factors might influence customers' motivation to switch*. We identified several factors that might enhance the probability that customers will move to a *post-trial*

decision process. We argued that these factors might have a negative (e.g. the customers' early familiarity with the domain), a positive (e.g. negative experiences), or both (e.g. customers characteristics) impact on the switching behavior. Some factors might trigger an extended-problem solving task which might induce customers to evaluate different channel alternative attributes and induce them to switch to a *post-trial* decision strategy. By contrast, others might contribute to reinforce the currently used decision process. We operazionalized these factors using different variables (e.g. returns for negative experiences). We performed a discriminant analysis, using as dependent a variable which takes values 1 if the customer was classified as *stayer* and 2 if he or she was classified as *switcher*, to test the effect of these factors on the group membership.

3) *People who switch exhibit two distinct stages in their channel decision process over time: trial and post-trial.* Drawing on Aaker (1971) new-trier model we contended that customers may move from a *trial* stage to a stage in which he or she has changed channel decision strategy (*post-trial*).
4) *The trial and the post-trial stages strongly differ along channel preferences, marketing responsiveness and state dependence*. We showed that the *post-trial* stage differs in term of customers' channel preferences, dependence upon previous channel choices, and responsiveness to the marketing with respect to the *trial* phase.
5) *Customers migrate towards different types of decision making "styles"; therefore we can depict different migration patterns.*. We contended the existence of different decision-making patterns in the channel choice context. We used these patterns to describe different *trial* and *post-trial* decision making strategies. We demonstrated that different evolution patterns, consequntly different types of migration exist. Each type of migration presents different characteristics in term of marketing

responsiveness, channel preferences and other aspects. This leads us to take an overall view on the channel choice migration process and to ground our work within a general decision-making framework which aims to describe the evolution of channel decision behavior of new customers to the firm.

Our results and our modeling approach have a managerial relevance because they could help managers to tailor specific marketing strategies which takes into account the existence of a *trial* and a *post-trial* stage in the channel decision process. Managers may also have insights on the timing of the direct marketing communications. They can predict the duration of the *trial* phase at individual level detecting the customers with a quick, long or even absent *trial* phase. They can even predict if the customer will change or not his decision process over time, and they can influence the switching process using the marketing tools.

## 6.2 Modeling Contribution

The only work in the literature which proposed a modeling approach which takes into account the existence of a learning phase is the work of Knox (2005). He considers that customers at the beginning of the relationship with the firm learn about channel alternatives and then migrate towards different inner state (e.g. online, offline and multichannel).

As Knox we modeled the adoption and the channel migration process of a cohort of new customers; however our work differs both in the theoretical contribution and in the modeling approach.

We use data of a cohort of new customers of a major multi-channel retailer in this study. We observe the channel migration process of this cohort of new customer from October 2001 until June 2006. We use as covariates in the channel choice model marketing communications, state dependence and heterogeneous channel preferences.

We postulated a *trial* period in the customers channel choice behavior. We captured, at individual level, the impact of direct marketing communications on channel migration process taking into consideration the existence of two stages in the customers channel choice's history. In order to capture this phenomenon, we distinguished between *trial* and *post-trial* periods in channel choice behavior. We conceptualize the customer decision process as a multinomial logit. We estimated the probability that the customer is learning prone (i.e. the probability the customer switches to the *post-trial* model) using a geometric distribution. During each purchase occasion the model estimates the probability that the customer switches to the *post-trial* period. Formally, we estimated a *multinomial logit channel selection switching model* which takes into account that customers might use one "*trial*" multinomial logit when they are first acquired, but then migrate toward an "after *trial*" multinomial logit after a certain period of time. The geometric distribution governs the transition from the *trial* period multinomial logit to the *post-trial* multinomial logit. We adopted a Bayesian approach to conducting inference in this model. Specifically, we added a hierarchical Bayesian structure in order to obtain individual-level estimates.

## 6.3 Theoretical Contribution

I summarize below our main results distinguishing them in two main classes, each of which answers to two general questions: 1) Why do some people switch decision processes while others don't? 2) How does the decision process change over time among the people who switch?

### 6.3.1 Customers' Propensity to Switch Decision Processes

Overall we find that:

- **There is a high presence of *stayers* in our sample, i.e. people who do not switch decision process and remain always with their initial channel decision making strategy**. This evidence is quite interesting because it tells us that a large proportion of customers use the same decision strategy over time. These customers exhibit a very low probability to experiment a *trial* and a *post-trial* stage. This result is supported in the literature. Several authors argue that in many situations consumers use simplified heuristics to make their choices and sometimes simplified decision processes might be observed since the first purchase (Solomon, Bamossy and Askegaard, 2002; Olshavsky and Granbois, 1979; Blackwell, Miniard and Engel 2002). For example, if the customer is not committed with the channel choice he or she chooses a specific channel alternative merely because less effort is required or for some reason it is easier to choose. Similarly, we can think to customers who exhibit since the beginning a strong preference for a particular channel or for a combination of channels. A nice example is given by John Mason, the 63-year-old widower described in Balasubramaninan et al. (2005), who always shops using a specific grocery store on his weekly shopping trip. This store seems like a second home to him, he perfectly knows where the items are, and he likes to chat with Susan Dillinger (one of the cashiers) at the checkout line. In these two examples we certainly can not distinguish *trial* and *post-trial* stages. This results is particularly important if we consider that all the customer migration models in the literature (e.g. Knox, 2005) implicitly suppose that customers new to the firm are likely to be learning about the firm's channel options available, in other words they hypothesize that at the beginning of their relationship customers undertake an extended-problem solving, that they are committed and involved with the channel choice task, without considering that a large number of customers might desire simplified decision processes. What is really interesting is that we found that the majority of *stayers* are responsive to the marketing.

- **The trial length is not the same among switchers**. We found evidence of a large variance in the number of purchase occasions that the customers need before switching to a *post-trial* stage. Specifically, 48% of the *switchers* have a quick *trial* period and 52% a long *trial* period. This supports our idea of differences in the length of the *trial* period among customers. Furthermore, we found that "quick *switchers*" exhibit differences in their channel usage. For example, the percentage of multichannel customers, specifically three-channel users is higher among "quick *switchers*".
- **Several factors significantly increase or decrease the probability to be *stayer* or *switcher*.** We found that a strong familiarity at the beginning with a particular channel decreases the probability to be a *switcher*. Customer motivation to search for information and learning proneness increase the probability to be *switchers*. Finally negative experiences (in particular returns) have a positive association with the probability to switch. This last result, for example, might explain the reason of a long *trial* phase. For example, if customers are not particularly involved with the channel choice task and they never experience "problems" or unsatisfying situations they might continue using their initial channel decision strategy, but if something "negative" happens this might trigger their motivation to learn and lead them to switch to a *post-trial* decision strategy.

### 6.3.2 How Choice Decision Patterns Evolve

Overall we find that:

- **Marketing communications have an impact in the channel migration process**. This result is consistent with previous works in this field. Ansari et al. (2008) found that emails were strongly associated with choice of the internet. Knox (2005) as well found that the online segment was highly responsive to emails. Pauwels and Neslin (2006) found that emails influenced catalog and Internet sales equally, while having little impact on store sales. We found that marketing communications' effect is significant both for *stayers* and *switchers*. In particular for *switchers* it significantly impact on the *trial* and post *trial* stages. Specifically, it differs systematically between these stages.
- **Marketing communications effect strongly differs in the trial and in the post-trial stages.** The effect of catalogs sent is always significant in the *trial* period but it strongly decreases in the *post-trial* period. Emails sent significantly affect channel choice both in the *trial* and *post-trial* period. In particular:
  - **Catalogs sent** have a positive association with store selection, suggesting that catalogs sent "advertise" the use of stores during the *trial* period. In the *post-trial* period the magnitude of the catalogs sent effect strongly decrease.
  - **Emails sent** increase the probability of the use of Internet over Store in the *trial* but their effect is reversed in the *post-trial* period, i.e. they significantly bring customers towards the Catalog over the Internet and they do not significantly affect anymore the probability to choose the Internet over Store

- **Initially customers are mainly inertial, but highly responsive to direct marketing stimuli. In the *post-trial* stage customers have developed consciuous channel preferences and the marketing could be both still effective or not on channel choice.**
- **Different types of migrations patterns between *trial* and *post-trial* stages exist. In particular we found evidence of two main types of migrations:**
    - **Inertial (High mkt)→Preference-Based (High mkt).** 37% of the customers switched from an inertial decision-making strategy characterized by high marketing responsiveness to a preference-based decision-making strategy which is still influenced by the marketing communication stimuli. This migration is quick; customers take on average 6 purchase occasions to switch. The customers who undertake this type of migration are mainly Catalog-Internet users, maybe with a preference for channels without contact personnel. There is also an high percentage of three-channel-users.
    - **Inertial (High mkt)→Preference-Based (Low mkt).** 56% of the customers migrate from an inertial decision strategy where marketing is effective, but they end up with a preference-based decision strategy where the marketing do not impact channel choice. These customers are mainly Catalog-Store users and there is an high percentage of customers who mainly use only the Store. This migration is slower than the first. It takes on average more than 9 purchase occasions.

Both migrations depict an interesting result: customers switch from an inertial-based decision strategy to a preference-based decision strategy, in other words customers at the beginning do not seem to be committed with the channel choice task, but external information (e.g. marketing) and/or some events (e.g. negative experiences) over time might induce them to learn on the channels alternatives attributes and to develop "conscious" channel preferences. The main difference in these two types of migration is the role of direct marketing communications. In the former the marketing maintains its effectiveness over time and it still has an impact on channel choice. In the latter, the marketing effect decreases over time.

## 6.4 Managerial Relevance

We believe that our results might help managers in several ways. First of all, we demonstrate the existence of *stayers* and *switchers* types of "decision-makers". *Stayers* might have conscious channel preferences or they might be inertial *stayers*, not committed with the channel choice task. The majority of *stayers* are positively marketing responsive, therefore managers could think to specific marketing strategies tailored for *stayers*. For example, companies should not "waste money" trying to induce a "conscious" *stayer* (e.g. John Mason, the 63-year-old widower described in Balasubramaninan et al., 2005) to change channel. Companies are probably aware of this type of behavior but it is important for them to predict the customers which have a high probability to be conscious *stayers*. Similarly, companies should not try to induce an "inertial *stayers*" to change channel, for example sending him complex and detailed marketing information (such as the typical catalog format). Simple and "attention capturing" marketing communications (see the literature on peripheral route and effective advertising) might be more effective for customers using simple heuristics to select channels.

The main managerial contribution is the identification of two stages in the customers channel choice history. Using our model companies could predict the length of the *trial* period. They should consider that during the *trial* period certain tools are effective. For example, catalogs during the *trial* period impact on channel choice and in particular they seem to “advertise” the use of the Store, but their impact strongly decreases during the *post-trial* phase. On the contrary, emails’ positive impact on channel choice is longer but it changes over time. These considerations could help managers to delineate a direct marketing strategy which take into account that the effect of marketing tools is different in the *trial* and in the *post-trial* phase.

Furthermore, we found evidence of two main types of migration patterns. Some customers migrate from an inertial channel decision strategy, when the marketing is effective, to a preference-based decision strategy which is not influenced by marketing. Therefore during the *trial* period marketing might help these customers to form their channel preferences. This *trial* period, on average, is long 9 purchase occasions. After that period these customers have developed their channel preferences, they will act as habitual in their channel decision process and they difficultly could be induced to change channels. Therefore, the company should be aware that during the *trial* period of these customers it has a big opportunity because it might lead customers towards the most profitable channels.

## 6.5 Limitation and Future Research

We use secondary data of one major retailer and a specific industry, channel migration can be affected by industry and differ among retailers. Further replications in other industries would be required to obtain empirical regularities on the relative influence of the variables.

Such replications would be beneficial for developing theories that can improve the effectiveness of multichannel marketing.

We used catalogs sent and email sent as independent variables in our model, however they are a very gross measures, given that firms send so many different kinds of catalogs and emails. Future research should account for different types of catalogs and emails.

We sample only customers with a "full life cycle" to perform our analysis. We need to do that in order to guarantee that the customers do not purchase only during the *trial* stage. However, in doing that we are aware that we might sample only the "best customers"; therefore our sample might be not representative of a "typical" customer. Further research should investigate the effect of *trial* versus *post-trial* stages in the channel decision strategies using a shorter time period. We used more than four years in our analysis, maybe in the context of frequently purchased goods a shorter time horizon might be enough.

Another limitation is that we do not allow for the possibility that a single purchase can be related to two channels. For example, many mail-order-catalog firms also have online channels. As Venkatesan et al. (2007) suggest this might lead to a significant "flowback" issue in which customers receive the catalog in the mail and then go online to place the order.

Finally, we use data on a subscription-oriented business model which might present distinctive characteristics. The advantages to use this kind of dataset are several (for example, we are sure to track all the purchase ever made by the customers in whatever channel). However, further research is needed to investigate the multichannel phenomenon distinguishing between subscription oriented business model and other type of business strategies.

Our descriptive analysis suggests some interesting evidences which might trigger further research. For example, we believe that further research is needed in order to investigate deeply multichannel customers lifetime value. Specifically, one should investigate if multichannel customers have a “longer life” with the firm. This might be important and it might add insight to the well-know result that multichannel customers buy more and are “best-customers”

It also might be interesting to understand if multichannel customers require more managerial efforts. In others words, whether they require more assistance (e.g. they returns often items). We believe that a cost/benefit analysis might enrich the well-know results that multichannel customers are more profitable.

Finally, another under-research area concerns the probability to “quit” the firm and the channel usage. It might be interesting to formally model the quitting behavior together with the customers channel migration behavior.

# REFERENCES

Aaker D. A. (1971), “The New-Trier Stochastic Model of Brand Choice”, Management Science, 17 (8)

Alba J. W. and Hutchinson J.W., 1987, Dimension of customer expertise. Journal of Consumer Research,13, 411-454

Alba J., Lynch J., Weitz B. Janiszewsky C., Lutz R., Sawyer A. and Wood S., Interactive Home Shopping: Cunsomer, Retailer, Manufacturer Incentives to Participate in Electronic Market Places, Journal of Marketing; Jul 1997; 61, 38-53s

Amemiya T., 1985, Advanced Econometrics, Cambridge, MA: Harvard University Press

Ansari, A., Mela, B., and e Neslin, S.A. (2008). Customer channel migration. Journal of Marketing Research, forthcoming February. 45(1), 60-76

Ariely, D. (2000), “Controlling the Information Flow: Effects on Consumers’ Decision Making and Preference.” Journal of ConsumerResearch, Vol. 27, No. 2: 233-248.

Bakos Y. 1991, ”A Strategical Analysis of Electronic Market-Places”, MIS Quarterly 15 (September): 295-310.

Bakos Y.,1997, Reducing buyer search costs: implications for electronics marketplaces. Management Science 43(12) 1676-1692

Balasubramanian S., Raghunathan R., Mahajan V. (2005), “Consumers in a Multichannel Environment”

Barnard P. O. and Hensher D. A., 1992, The spatial Distribution of Retail Expenditures, Journal of Transport Economics and Policy, 26 (September), 299-312.

Belk R., 1974, “An Exploratory assessment of situational effects in buyer behavior”, Journal of Marketing Research, 11(May), 156-163.

Bell, D.R. and J.M. Lattin (1998). Shopping Behavior and Consumer Preferences for Store Price Format: Why Large Basket Shoppers Prefer EDLP, Marketing Science 17(1), 66–88.

Bell, D.R., T. Ho, and C.S. Tang (1998) “Determining Where to Shop: Fixed and Variable Costs of Shopping” Journal of Marketing Research, 35, 3, pp. 352-369

Bell, David R., Jeongwen Chiang, and V. Padmanabhan (1999), “The Decomposition of Promotional Response: An Empirical Generalization,” Marketing Science, 18 (4), 504-526.

Ben-Akiva M. and Lerman S. R., 1985, Discrete Choice Analysis, Cambridge, MA, MIT Press

Berger, P.D., J. Lee, and B.D. Weinberg (2005), "Optimal Cooperative Advertising Integration Strategy for Organizations Adding a Direct Online Channel", Journal of the Operational Research Society, 56 (1), 15-24.

Berman B., J.R. Evans, 1995 "Retail Manangement. A Strategic Approach" 6th ed.Prentice Hall, New Jersey.

Bettman J.R., 1979, An information processing theory of consumer choice. Reading, MA:Addison-Wesley Publishing Company.

Black N. J., Lockett A., Ennew C., Winklhofer H., McKechnie S. (2002), "Modelling consumer choice of distribution channels: An illustration from financial services", The International Journal of Bank Marketing, 20, 4/5.

Blackwell R. D., Miniard P. W. and Engel J.F., 2001, "Consumer Behavior- 9th Edition", Harcourt College Publishers

Blattberg, R. C., Kim, B-D., and Neslin, S. A. (2008). Database marketing: analyzing and managing customers. New York: Springer.

Bloch P. H. and Richins M. L., 1983, A theoretical model for the study of product importance perceptions. Journal of marketing, 47(summer), 69-81

Boulding, William, Richard Staelin, Michael Ehret, and Wesley J. Johnston (2005), "A Customer Relationship Management Roadmap: What Is Known, Potential Pitfalls, and Where to Go," Journal of Marketing, 69 (4), 155-166.

Brooks, S. and A. Gelman. 1998. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7: 434-55.

Brucks M., 1985, The effect of productbclass knowledge on information search behavior. Journal of Consumer Research 12, 1-16

Brynjolfsson, Erik and Michael D. Smith (2000), "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," Management Science, 46 (4), 563-585.

Buchanan, B.S. and Morrison, D.G., (1988)"A Stochastic Model of List Falloff with Implications for Repeat Mailings," Journal of Direct Marketing, 7-15

Bucklin R. E. and Gupta S., 1992, "Brand Choice, Purchase Incidence and Segmentation: An Integrated Modeling Approach, Journal of Marketing Research, 29 (May), 201-215.

Burnkrant R. E. and Sawyers A.G., 1983, effects of involvements and message content on information-processing intensity, in information processing research in advertising, ed. Richard j. harris, hillsdale, NJ: Lawrence Erlbaum associates, 43-64

Chan, T., Y. Ma, C. Narasimhan and V. Singh. (2005). A Store Choice Model with Basket Shopping, Working Paper, Washington University in St. Louis.

Chiang J., (1991), “A Simultaneous Approach to the Whether, What and How Much to buy Questions”, Marketing Science, 10 (Fall), 297-315

Chib S., Seetharaman P.B., Strijnev A., (2004), Model of Brand Choice with a No-Purchase Option Calibrated to Scanner-Panel Data, Journal of Marketing Research, 41 (May) 184-196

Chintagunta P. K., (1993), “Investigating Purchase Incidence, Brand Choice and Purchase Quantity Decisions of Households”, Journal of Marketing Research, 39(May), 141-154.

Chintagunta P. K., (2001), Endogeneity and Heterogeneity in a Probit Demand Model: Estimation Using Aggregate Data (2001), Marketing Science, 20, 4, 442-456

Chintagunta P. K., (2002), “Investigating Category Pricing Behavior in a Retail Chain”, Marketing Science, 12(1), 184-208.

Chintagunta P. K.. (1992), “Estimating a Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments”, Marketing Science, 11 (4), 386-407.

Coelho, Filipe, Chris Easingwood, and Arnaldo Coelho (2003), "Exploratory Evidence of Channel Performance in Single vs. Multiple Channel Strategies," International Journal of Retail and Distribution Management, 31(11), 561-573.

Cohen J.B., 1983, “Involvment and You: 1000 great ideas”, in advances in consumer research, vol.10, eds R.P. Bagozzi and A. M. Tybout, Ann Arbor, MI: Association for consumer research, 325-328.

Congdon, P. (2005), “Bayesian Models for Categorical Data” John Wiley& Sons: New York.

Cox D.F., 1967, Risk-taking and information handling in consumer behavior. Boston: Harvard University

Davidson, R. and J.G. MacKinnon (1993), Estimation and Inference in Econometrics, 1st Edition, Oxford University Press

Deleersnyder, Barbara, Inge Geyskens, Katrijn Gielens, and Marnik G. Dekimpe (2002), “How Cannibalistic is the Internet Channel?” International Journal of Research in Marketing, 19(4), 337–348.

Doubleclick (2004a), "Retail Details: Best Practices in Multi-Channel Integration" (March), New York: DoubleClick, Inc.

DoubleClick (2004b), "Multi-Channel Shopping Study – Holiday 2003," New York: DoubleClick, Inc.

Dube' J.P. and Manchanda P., 2005, Dynamic Brand Competition across Markets: An empirical analysis. Marketing Science, 24(1), 81-95

Dutta S., Bergen, M., Heide, J.B., John, G. (1995), "Understanding dual distribution: the case of reps and house accounts", Journal of Law, Economics, and Organisation, Vol. 11 No. 1, pp. 189-205.

Easingwood C., Storey, C. (1996), "The value of multi- channel distribution systems in the financial services sector", The Service Industries Journal, Vol. 16 No. 2, pp. 223-41

Elrod T., 1988, "Choice Map: Inferring a Product Market Map From Panel Data", Marketing Science, 7, 21-40.

Erdem T.e Keane M. P. (1996), "Decision-Making under Uncertainty: Capturing Dynamic Choice Processes in Turbulent Consumer Goods Markets," Marketing Science, 15 (1), 1-20

Evans J.S.B.T. and Over D.E. ,1996, Rationality and Reasoning. Hove, UK: Psychology Press

Forster S., 2004, "When One Hand Doesn't …", The Wall Street Journal, (March 22), R3

Fourt L.A., Woodlock J.W., 1960. "Early prediction of market success for new grocery products" Journal of Marketing 25: 31–38

Fox , E. J., Montagomery A. L., Lodish L. M., 2004 "Consumer shopping and spending across retail formats", Journal of Business, 77, 2.

Frances P.H. and Paap R., Quantitative Models in Marketing Research, Cambridge University Press, 2001

Frazier, G.L. (1999), "Organizing and managing channels of distribution", Journal of the Academy of Marketing Science, 27, 2, pp. 226-40

Gelman, A. "Applied Bayesian Modelling" 2003. Wiley.

Gelman, A. and D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences (with discussion). Statistical Science 7: 457-511.

Gelman, A., Carlin, B.P., Stern, H.S. and D.B. Rubin. 1995. Bayesian Data Analysis. Chapman and Hall, London.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. pp. 169-93. In: Bayesian Statistics 4 (eds J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith.) Oxford University Press, Oxford

Gönül, Fusun and Kannan Srinivasan (1993) "Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues," Marketing Science, 12 (3), 213-229.

Gönül, Füsun, Byung-Do Kim, and Mengze Shi (2000), "Mailing Smarter to Catalog Customers," Journal of Interactive Marketing, 14 (2), 2-16.

Granbois D., 1977, "Shopping Behavior and preferences", in Selected Aspects of Consumer Behavior- A Summary from the perspective of different disciplines, Washington D.C., U.S. Printing Office, 259-298.

Greene (2002), Econometric Analysis, Prentice Hall

Greenwald A. G. and Leavitt C., 1984, "Audience involvement  in Advertising: Four Levels", Journal of Consumer Research, 11 (June), 581-592

Guadagni P. M, Little J.D.C. (1983), "A Logit Model of Brand Choice Calibrated On Scanner Data", Marketing Science, 2 (3), 203-237

Gupta, Alok, Bo-Chiuan Su, and Zhiping Walter (2004b), "An Empirical Study of Consumer Switching from Traditional to Electronic Channels: A Purchase-Decision Process Perspective," International Journal of Electronic Commerce, 8(3), 131-161

Ha Y.W. and Hoch S. J., 1988, Ambiguity processing strategy, and advertising-evidence interactions". Working paper, center for decision research, graduate school of business, university of Chicago.

Heckman, J. J. (1981), "Heterogeneity and state dependence", in Studies in Labor Markets, ed. S. Rosen, Chicago, Chicago Press.

Heckman, J., "Simple Statistical Models for Discrete Panel Data Developed and Applied to Tests of the Hypothesis of True State Dependence against the Hypothesis of Spurious State Dependence," Annales de l'INSEE, 1978, 30-31, pp. 227-269

Ho T.-H., Tang C. S. and Bell D. R., 1998, Rational Shopping behavior and the option value of pricing. Management Science 44 (December), S145-S160

Hoch S. J. and Deighton J., (1989), "Managing what consumers learn from experience" Journal of marketing, 53(2), 1-20

Hoffman D. and Novak T., 1996, Marketing in hypermedia computer-mediated environments: conceptual foundations. Journal of  Marketing, 60, July, 50-68

Hoffman D. L., Novak T. P., Peralta M., 1999, "Building Customer Trust Online", Communications of the ACM, 42 (4), 80-85

Holsapple. C. W., and Singh, M. (2000), Electronic commerce: From a definitional taxonomy toward a knowledge-management view. Journal of Organizational Computingand Electronic Commerce, 10(3), 149-170

Hoque A.Y. and Lohse G. L., 1999, "An Information Search Cost Perspective for Designing Interfaces for Electronic Commerce", Journal of Marketing Research, 36(3):387-394

Houston M.J. and Rothshild M.L., 1978, “Conceptual and Methodological perspectives in involvement”, in Research Frontiers in Marketing: Dialogues and directions, ed. S. Jain, Chicago, IL, American Marketing Association, 184-187

Inman, Jeffery J., Venkatesh Shankar, and Rosellina Ferraro (2004), “The Roles of Channel-Category Associations and Geodemographics in Channel Patronage Decision,” Journal of Marketing, 68 (April), 51-71.

Inman, Jeffery J., Venkatesh Shankar, and Rosellina Ferraro (2004), “The Roles of Channel-Category Associations and Geodemographics in Channel Patronage Decision,” Journal of Marketing, 68 (2), 51-71

Jeuland A. P. (1978). “Brand Preferences Over Time: A Partially Deterministic Operationization of The Notion of Variety-Seeking” in Subhash Jain (Ed.), AMA Research Frontiers in Marketing: Dialogues e Directions, Educators’Proceedings Series 43, Chicago, IL:American Marketing Association, 33-37.

Jeuland A.P, 1979, “Brand Choice Inertia as One Aspect of the Notion of Brand Loyalty”, Management Science, 25(3), 671-682

Johnson, Eric J. and J. Edward Russo (1981), “Product Familiarity and Learning New Information,” in Advances in Consumer Research, Vol. 8, ed. Kent B. Monroe, Ann Arbor, MI: Association for Consumer Research

Jones M.J. and Landwehr J., 1988, “Removing Heterogeneity Bias from Logit Model Estimation”, Marketing Science, 7 (1), 41-59.

Kahan B. E., Kalwani M. U., Morrison D. G. 1986, “Measuring Variety-Seeking and Reinforcement Behaviors Using Panel Data”, Journal of Marketing Research, 23(2), May, 89-100

Kamakura W. A. e Russell G. L. (1989), “A Probabilistic Choice Model for Market Segmentation and Elasticity Structure” Journal of Marketing Research, 26 (4) 379-390

Kau, A.K. and Ehrenberg, A.S.C. (1984) ‘Patterns of store choice’, Journal of MarketingResearch, 21, 399-409

Keane, Michael P. (1997), “Current Issues in Discrete Choice Modeling,” Marketing Letters, 8 (3), 307-322.

Keen, Cherie, Martin Wetzels, Ko de Ruyter, and Richard Feinberg (2004), “E-Tailers versus Retailers. Which Factors Determine Consumers Preferences,” Journal of Business Research, 57, 685-695.

Knox, George (2005), “Modeling and Managing Customers in a Multichannel Setting,” Working paper, University of Tilburg, The Netherlands.

Kuehn A. A. (1962), “Consumer Brand Choice: A Learning Process?” Journal of Advertising Research, 2, 10-17.

Kumar V. and Rajkumar Venkatesan (2005), “Who Are Multichannel Shoppers and How Do they Perform?: Correlates of Multichannel Shopping Behavior,” Journal of Interactive Marketing, 19(Spring), 44-61.

Kushwaha, Tarun L. and Venkatesh Shankar (2005), “Multichannel Shopping Behavior: Antecedents and Implications for Channel and Customer Equity,” Working Paper, Texas AandM University, College Station, TX 77845.

Lancaster T. (1979) Econometric methods for the duration of unemployment, Econometrica,47, 4, 939–956

Lastovicka J. L. and Gardner D.M., 1979, Components of involvement, in attitudes play for high stakes , john c. Maloney and Bernard silverman, eds. Chicago: American marketing association, 53-73

Lilien, Moorthy and Kotler (1992). Marketing Models. Englewood Cliffs: Prentice-Hall.

Lindley D and Smith. A F M (1972), Bayes estimates for the linear model (with discussion) J. Royal. Statistics Society Ser B. 34 1-41

Louviere, J. J., Hensher, D.A., and Swait, J. D. (2000). Stated Choice Methods: Analysis and Application. Cambrige, UK: Cambrige University Press.

Louviere, J., Train, K. Ben-Akiva, M., Bhat, C., Brownstone, D. Cameron, T.A., Carson, R. T., DeShazo J.R., Fiebig, D., Greene, W., Hensher, D., and Waldman, D.( 2005)Recent Progress on Endogeneity in Choice Modeling. Working Paper

Luce, R. D.: 1959, Individual Choice Behavior: A Theoretical Analysis, John Wiley and Sons,New York, NY

Lynch, John and Dan Ariely (2000), “Wine Online: Search Costs Affect Competition on Price, Quality, and Distribution,” Marketing Science, 19 (1), 83-103.

MacInnis D.J. and Jaworsky B.J. 1989, information processing from advertisements: toward an integrative framework. Journal of marketing, 53(October), 1-23.

Maddala, G. S. (1983), Limited-Dependent and Qualitative Variables in Econometrics, Cambridge, UK: Cambridge University Press.

Mahajan, Vijay, Raji Srinivasan, and Jerry Wind (2002), “The Dot.com Retail Failures of 2000: Were There Any Winners?” Journal of the Academy of Marketing Science, 30 (4), 474-486.

Marschak, J.(1960) “Remarks on the Economics of Information”, Cowles Foundation paper 146, repr. Contributions to Scientific Research in Management, University of California, UCLA.

Martineau P., 1958, “Social Classes and Spending Behavior” , Journal of Marketing, 23 (October), 121-130.

Mathwick, Charla, Naresh K. Malhotra, and Edward Rigdon (2002), "The Effect of Dynamic Retail Experiences on Experiential Perceptions of Value: An Internet and Catalog Comparison," Journal of Retailing, 78, 51-60.

McAlister L. (1982) “A Dynamic Attribute Satiation Model Of Variety-Seeking Behavior” Journal of Consumer Research, 9, 141-151

McAlister L. and Pessemier E., 1982. Variety seeking behaviour: An interdisciplinary review. Journal of consumer Research, 9, 311-322

McCulloch, Robert E., Peter E. Rossi. 1994. An exact likelihood analysis of the multinomial probit model. Journal of. Econometrics 64, 217–228.

McKnight D. H., Choudhury V. and Kacmar C., 2002Developing and validating trust measures for e-commerce: an integrative typology. Information systems research, 13(3), 334-359

McLean, R. e Blackie, N. M. (2004), “Customer and Company Voices in e-Commerce: A Qualitative Analysis”, Qualitative Market Research: An International Journal, 7 (4), 243-249.

Messinger, Paul R and Narasimhan, Chakravarthi (1997), "A Model of Retail Formats Based on Consumers' Economising on Shopping Time", Marketing Science, Vol. 16, No. 1, pp. 1 23.

Miller D., Jackson P., Thrift N., Holbrook B. and Rowlands M. 1998, Shopping Place and Identity, London: Routledge.

Mitchell A. A. and Olson J.C. 1981, Are product attribute beliefs the only mediator of advertising effects on brand attitude? Journal of marketing research, 18 (August), 318-332

Miyake N. and Norman D., 1979, “To ask a question, one must know enough to know what is not known”. Journal of verbal learning and verbal behavior, 18, 357-364.

Montoya-Weiss, Mitzi M., Glenn B. Voss, and Dhruv Grewal (2003), “Determinants of Online Channel Use and Overall Satisfaction with a Relational Multichannel Service Provider,” Journal of the Academy of Marketing Science, 31(4), 448-458.

Moorthy S., Ratchford B.T. and Talukdar D., 1997, “Consumer Information Search Revisited” Journal of Consumer Research, 23(march), 263-277

Moriarty, R., Moran, U. (1990), “Managing hybrid marketing systems”, Harvard Business Review, November-December, 146-55.

Morrison, D.G. and Perry, A., (1970), "Some Data Based Models for Analyzing Sales Fluctuations," Decision Sciences, 4, 258-274.

Moshkin N. and Shachar R., 2002, The asymmetric information model of state dependence. Marketing Science, 21(4), 435-454.

Murray K. B., 1991, A test of services marketing theory: consumer information acquisition activities., Journal of Marketing, 55(1), 10-25

Myers, Joe, Evan Van Metre, and Andrew Pickersgill (2004), "Steering Customers to the Right Channels," The McKinsey Quarterly: McKinsey on Marketing, (October).

Nelson FD, Olson L (1978) Specification and estimation of a simultaneous equation model with limited dependent variables. Int Econ Rev 19:695–710

Neslin, Scott A., Dhruv Grewal, Robert Leghorn, Venkatesh Shankar, Marjie L. Teerling, Jacquelyn S. Thomas and Peter C. Verhoef (2006a), “Challenges and Opportunities in Multichannel Customer Management,” Journal of Service Research, 9 (2), 95-112.

Neslin, Scott A., Thomas P. Novak, Kenneth R. Baker, and Donna L. Hoffman (2006b) “An Optimal Contact Model for Maximizing Online Panel Response Rates,” Working Paper, Tuck School of Business, Dartmouth College, Hanover, NH.

Nicholson, Michael, Ian Clarke, and Michael Blakemore (2002), “One Brand, Three Ways to Shop: Situational Variables and Multichannel Consumer Behavior,” The International Review of Retail, Distribution and Consumer Research, 12 (2), 131-148.

Paap, R. and Franses, P. H. (2000) "A dynamic multinomial probit model for brand choice with different long-run and short-run effects of marketing-mix variables," Journal of Applied Econometrics, John Wiley and Sons, Ltd., vol. 15(6), pages 717-744

Pauwels, K., and Neslin S.A. (2006). Building with bricks and mortar: the revenue impact of opening physical stores in a multichannel environment. Working paper: Dartmouth College, Hanover, NH.

Payne, Adrian, and Pennie Frow (2005), “A Strategic Framework for Customer Relationship Management, Journal of Marketing, 69 (4), 167-176.

Peterson, Robert A., Sridhar Balasubramanian, and Bart J. Bronnenberg (1997), “Exploring the Implications of the Internet for Consumer Marketing,” Journal of the Academy of Marketing Science, 25 (4), 329-346.

Petty R.E. and Cacioppo J.T., 1981, Attitudes and Persuasion: Classic and contemporary approaches. Dubuque, IA: William Brown.

Rangaswamy, Arvind and Gerrit van Bruggen (2005), “Opportunities and Challenges in Multichannel Marketing: An Introduction to the Special Issue,” Journal of Interactive Marketing, 19(2), 5-11.

Ratchford B. T., Lee M.-S., Talukdar D. 2003, The impact of the internet on information search for automobiles. Journal of marketing research, 40 (may), 193-209

Rhee, S. and Russel, G. (2003) “Measuring Household Response in Database Marketing: A Latent Construct Approach” Working Paper, The University of Iowa

Richard W. Olshavsky; Donald H. Granbois, The Journal of Consumer Research, Vol. 6, No. 2, Special Issue on Consumer Decision Making. Consumer Decision Making-Fact or Fiction? (Sep., 1979), pp. 93-100.

Rossi P. E. and G. M.Allenby, (1993), “A Bayesian Approach to Estimating Household Parameters”, Journal of Marketing Research, 30 May, 171-182.

Rossi P., McCulloch R.E. and Allenby G.M., 1996, The Value of Purchase History Data in Target Marketing,”, Marketing Science, 15(4), 321-340.

Rossi, P., and Allenby G. (2003). Bayesian statistics and marketing. Marketing Science, 22(3), 304-329.

Rossi, P.E., Allenby, G.M., and McCulloch, R. (2005), “Bayesian Statistics and Marketing” NJ: Wiley.

Rossi, Peter E., Greg M. Allenby. 1993. A Bayesian approach to estimating household parameters. Journal of Marketing Research. 30 171–182.

Schoenbachler, Denise D. and Geoffrey L. Gordon (2002), “Multi-Channel Shopping: Understanding What Drives Channel Choice”, Journal of Consumer Marketing, 19 (Winter), 42 51.

Seetharaman, P. B. (2004) “Modeling Multiple Sources of State Dependence in Random Utility Models: A Distributed Lag Approach,” Marketing Science, 23 (2), 263-271.

Seetharaman, P. B., A. Ainslie and P.K. Chintagunta, Investigation Household State Dependence Effects Across Categories, Journal of Marketing Research, Nov 1999, 36(4), 488-500.

Shugan S.M., 2005, “Brand Loyalty Programs: Are They Shames?”, Marketing Science, Vol. 24, No. 2, Spring 2005, pp. 185–193.

Shugan S.M., 2006, “Are Consumers Rational? Experimental Evidence?”, Marketing Science, Vol. 25, No. 1, January–February 2006, pp. 1–7.

Shugan, Steven M. (1980), “The Cost of Thinking,” Journal of Consumer Research 7, 99-111

Smith R. E.1993, integrating information from advertising and *trial*: processes and effects on consumer response to product information. Journal of marketing research, 30 (may), 204-219.

Smith R.E. and Swinyard W.R., 1988. cognitive response to advertising and *trial*: belief strength, belief confidence and product curiosity. Journal of advertising, 17 (fall), 3-14.

Solomon M., Bamossy G. and Askegaard S., 2002, Consumer Behavior: A European Perspective 2nd Edition, Prentice Hall Europe.

Spinnewyn, F. (1981), “Rational Habit Formation”, European Economic Review, 15 p. 91-109.

Sriram and Kalwam, Management Science, 53(1), 2007, P.98.

Steckel J, Vanhonacker W. 1988. A heterogeneous conditional logit model of choice. Journal of Business and Economic Statistics 6: 391-398.

Stern, Louis W., and Adel I. El-Ansary (1972), Marketing Channels, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Thomas, Jacquelyn S. and Ursula Y. Sullivan (2005a), “Managing Marketing Communications with Multichannel Customers,” Journal of Marketing, 69 (4), 239 - 251.

Thomas, Jacquelyn S. and Ursula Y. Sullivan (2005b), “Investigating Best Customers in a Multi-Channel Setting,” Working Paper, Medill School, Northwestern University.

Thurstone, L. (1927): “A Law of Comparative Judgement,” Psychological Review, 34,273–286.

Train, K. (2003) “Discrete Choice Methods with Simulation” Cambridge University Press.

Tschirgi J.E. 1980, Sensible reasoning: a hypothesis about hypotheses. Child Development, 51, 1-10

Uncles, M. D., and Ehrenberg, A. S. C. (1988). "Patterns of Store-choice: New evidence from the USA." Store Choice, Store Location, and Market Analysis., N. Wrigley., ed., London: Routledge.

Urban, Glen L. (2004), Digital Marketing Strategy, Upper Saddle River, NJ: Pearson Prentice-Hall.

Vakratsas D. and Ambler T. 1999, How Advertising Works: What Do We Really Know?” Journal of Marketing, 63, jan., 26-43.

Van Baal, Sebastiaan and Christian Dach (2005), “Free Riding and Customer Retention Across Retailers Channels,” Journal of Interactive Marketing, 19(2), 75-85.

Venkatesan, R., and Kumar, V. (2007). Multichannel shopping: causes and consequences. Journal of Marketing, 71(2), 114-132.

Venkatesan, Rajkumar, V. Kumar, and Nalini Ravishanker (2006), “Multi-Channel Shopping: Causes and Consequences,” forthcoming, Journal of Marketing.

Verhoef, Peter C., Scott A. Neslin, and Bjorn Vroomen (2007), “Multichannel Customer Management: Understanding the Research-Shopper Phenomenon,” International Journal of Research in Marketing, 24 (2), 129-148.

Wall Street Journal “Shopper Who Blend Store, Catalog, Web, Spend More”, September 3, 2004.

Ward, Michael R. (2001), "Will Online Shopping Compete More with Traditional Retailing or Catalog Retailing," Netnomics, 3(2), 103-117.

Webster, Frederick E., Jr. (1991), Indus*trial* Marketing Strategy, 3rd Edition, New York: John Wiley and Sons.

Weiner B., 1985, Spontaneous causal thinking. Psychological Bulettin, 97(1), 74-84

Wendel, S. and B.G.C. Dellaert (2005), “Situation Variation in Consumers‘ Media Channel Consideration”, Journal of the Academy of Marketing Science, 33 (4), p.575-584.

Windmeijer, F.A.G., Santos Silva, J.M.C. (1997) Endogeneity in Count Data Models: An Application toDemand for Health Care, Journal of Applied Econometrics, 12 (3), 281-294.

Wooldridge, Jeffrey M. (2002), Econometric Analysis of Cross Section and Panel Data, Cambridge, MA: MIT Press.

Wright P. L. (1974),The harassed decision maker: time pressures, distractions and the use of evidence. Journal of applied psychology, 59 (October), 555-561.

Wrigley N. (1988), “Store Choice, Store Location and Market Analysis” London: Routledge.

Wrigley, N. and Dunn, R. (1984) “Stochastic panel-data models of urban shopping behaviour: 2 Multi-store purchasing patterns and the Dirichlet model”, Environment and Planning A, 16, 759-778.

Young S.J., 2001, “Make the Most of Every Channel”, Catalog Age, 18 (March), 61-62.

Zaichowsky J.L., 1985, Measuring the involvement Construct. Journal of Consumer Research, 12, 341-352.

# APPENDIX

## **APPENDIX 1a**: MODEL SYNTAXES

### Multinomial Probit

```
model
{
for (n in 1:Nobs) {

U[n,1] <- b[h[n],3]*CS[n,1]+b[h[n],6]*ES[n,1] +
b[h[n],9]*LC[n,1]+error[h[n],1]
U[n,2] <- b[h[n],1] +b[h[n],4]*CS[n,2]+b[h[n],7]*ES[n,2] +
b[h[n],9]*LC[n,2]+error[h[n],2]
U[n,3] <- b[h[n],2] +b[h[n],5]*CS[n,3]+b[h[n],8]*ES[n,3] +
b[h[n],9]*LC[n,3]+error[h[n],3]
U[n,4] <-   error[h[n],4]

Maximum12[n] <- max(U[h[n],1],U[h[n],2])
Maximum3[n] <- max(U[h[n],3],Maximum12[n])
Maximum4[n]   <- max(U[h[n],4],Maximum3[n])
for (k in 1:4) {
Z[n,k]  <- equals(U[h[n],k],Maximum4[n])
Y[n,k] ~ dbern(Z[n,k])
                  }
            }

for(k in 1:H) {
b[k,1:9] ~ dmnorm(b.mu[1:9], b.tau[1:9, 1:9])
}
b.mu[1:9] ~ dmnorm(m[1:9], prec[1:9, 1:9])
b.tau[1:9, 1:9] ~ dwish(R[1:9, 1:9], 9)
b.sig[1:9, 1:9] <- inverse(b.tau[,])


for (i in 1:100) {
for (k in 1:4)  {
 error[i,k]  ~ dnorm(0,precer[k])
 }
}
precer[1]<-1
for (k in 2:4)  {
precer[k]~dgamma(0.5,0.5)
}
}
# b from 1 to 2 are intercepts (b0)
# b from  3 to 5 are catalog sent coefficients (b1)
# b from  6 to 8 are email sent coefficients (b2)
# b 9 are state dependence coefficients (b3)
```

## Nested Logit Channel Selection Switching Model

```
model
{

# LEARNING MODEL
# THIS IS THE MODEL THAT GOVERNS THE TRANSITION FROM THE "TRIAL PERIOD"
NESTED
# LOGIT TO THE "post-trial" NESTED LOGIT.
# THIS IS A GEOMETRIC DISTRIBUTION WITH
# Q[i] = PROBABILITY THE CUSTOMER SWITCHES TO THE STEADY STATE MODEL.
# X1[i,s] = PROBABILITY THE CUSTOMER IS USING THE STEADY STATE MODEL IN
PERIOD T.

for (i in 1:NHH){
q[i]<- max(0.0000, min(0.98, 1/(1+exp(-(c0[i]+c1*sex[i]+c2*age[i])))))
for (s in 1:19) {
  X1[i,s] <- 1-(pow((1-q[i]),(s-1)))
}
}

# CHOICE MODELS
# TRIAL=0 STEADY STATE=1
# THERE ARE THREE ALTERATIVES - CATALOG, INTERNET, STORE
# CS = CATALOGS; SAME FOR ALL ALTERNATIVES; COEF VARIES ACROSS ALTS.
# ES = EMAILS; SAME FOR ALL ALTERNATIVES; COEF VARIES ACROSS ALTS.
# LC = STATE DEPENDENCE; VARIES ACROSS ALTS; COEFF SAME ACROSS ALTS.
# ALL COEFS VARY ACROSS HOUSEHOLDS.
# UTILITY FOR ALTERNATIVE 3 IS SET = JUST b31[h[n]]*LC[n,3] BECAUSE THE
COEFS
# FOR CS AND ES ARE ALTERNATIVE SPECIFIC SO THE COEF FOR ONE OF THE ALTS
# IS NOT IDENTIFIED.  WE ARBITRARILY MAKE THIS ALTERNATIVE 3.
# LOOP IS OVER OBSERVATIONS ("NOBS"); 21 OBS PER HH; H[N] SIGNIFIES THE HH
# FOR OBSERVATION N.


for (n in 1:NOBS) {

z0[n]<-alpha0[h[n]]+ lambda0[h[n]]*IncVal0[n]
IncVal0[n] <-log(vbot0[n])
z1[n]<-alpha1[h[n]] + lambda1[h[n]]*IncVal1[n]
IncVal1[n] <-log(vbot1[n])


U0[n,1] <- b00[h[n],1]+b10[h[n],1]*CS[n]+b20[h[n],1]*ES[n] +
b30[h[n]]*LC[n,1]
U0[n,2] <- b00[h[n],2]+b10[h[n],2]*CS[n]+b20[h[n],2]*ES[n] +
b30[h[n]]*LC[n,2]
U0[n,3] <- b30[h[n]]*LC[n,3]
U1[n,1] <- b01[h[n],1]+b11[h[n],1]*CS[n]+b21[h[n],1]*ES[n] +
b31[h[n]]*LC[n,1]
U1[n,2] <- b01[h[n],2]+b11[h[n],2]*CS[n]+b21[h[n],2]*ES[n] +
b31[h[n]]*LC[n,2]
U1[n,3] <- b31[h[n]]*LC[n,3]

vbot0[n]<-exp(U0[n,1])+exp(U0[n,2])+exp(U0[n,3])
vbot1[n]<-exp(U1[n,1])+exp(U1[n,2])+exp(U1[n,3])
```

```
pr1[n,1]<- (1-X1[h[n],t[n]])*((exp(U0[n,1])/vbot0[n])*(1/(1+exp(-z0[n]))))+
X1[h[n],t[n]]*((exp(U1[n,1])/vbot1[n])*(1/(1+exp(-z1[n]))))
pr[n,1]<-max(.00000000, min(.98, pr1[n,1]))
pr1[n,2]<- (1-X1[h[n],t[n]])*(exp(U0[n,2])/vbot0[n]*(1/(1+exp(-z0[n]))))+
X1[h[n],t[n]]*(exp(U1[n,2])/vbot1[n]*(1/(1+exp(-z1[n]))))
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))
pr1[n,3]<-(1-X1[h[n],t[n]])*((exp(U0[n,3])/vbot0[n])*(1/(1+exp(-z0[n]))))+
X1[h[n],t[n]]*((exp(U1[n,3])/vbot1[n])*(1/(1+exp(-z1[n]))))
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))
pr1[n,4]<-(1-X1[h[n],t[n]])*(1-(1/(1+exp(-z0[n]))))+X1[h[n],t[n]]*(1-
(1/(1+exp(-z1[n]))))
pr[n,4]<-max(.00000000, min(.98, pr1[n,4]))


Y[n,1:4] ~ dmulti( pr[n,1:4] , 4)

}

# PRIORS
for (n in 1:NHH)  {

b30[n] ~ dnorm(mu30,prec30)
b31[n] ~ dnorm(mu31,prec31)
c0[n] ~ dnorm(muc0,precc0)
alpha0[n] ~ dnorm(mua0,preca0)
lambda0[n] ~ dnorm(mul0,precl0)
alpha1[n] ~ dnorm(mua1,preca1)
lambda1[n] ~ dnorm(mul1,precl1)

for (k in 1:2) {
  b00[n,k] ~ dnorm(mu00[k],prec00)
  b10[n,k] ~ dnorm(mu10[k],prec10)
  b20[n,k] ~ dnorm(mu20[k],prec20)
  b01[n,k] ~ dnorm(mu01[k],prec01)
  b11[n,k] ~ dnorm(mu11[k],prec11)
  b21[n,k] ~ dnorm(mu21[k],prec21)
}
}

c1 ~dnorm(muc1,precc1)
c2 ~dnorm(muc2,precc2)

preca0~dgamma(0.5,0.5)
precl0~dgamma(0.5,0.5)
prec00~dgamma(0.5,0.5)
prec10~dgamma(0.5,0.5)
prec20~dgamma(0.5,0.5)
prec30~dgamma(0.5,0.5)
preca1~dgamma(0.5,0.5)
precl1~dgamma(0.5,0.5)
prec01~dgamma(0.5,0.5)
prec11~dgamma(0.5,0.5)
prec21~dgamma(0.5,0.5)
prec31~dgamma(0.5,0.5)
precc0~dgamma(0.5,0.5)
precc1~dgamma(0.5,0.5)
precc2~dgamma(0.5,0.5)

for (k in 1:2) {
mu00[k]~dnorm(0,.00001)
```

```
mu10[k]~dnorm(0,.00001)
mu20[k]~dnorm(0,.00001)
mu01[k]~dnorm(0,.00001)
mu11[k]~dnorm(0,.00001)
mu21[k]~dnorm(0,.00001)
}

mua0~dnorm(0,.00001)
mul0~dnorm(.5,00001)
mu30~dnorm(0,.00001)
mua1~dnorm(0,.00001)
mul1~dnorm(.5,00001)
mu31~dnorm(0,.00001)
muc0~dnorm(0,.00001)
muc1~dnorm(0,.00001)
muc2~dnorm(0,.00001)

}
```

**MNL with no purchase Channel Selection Switching Model**

```
model
{

# LEARNING MODEL
# THIS IS THE MODEL THAT GOVERNS THE TRANSITION FROM THE "TRIAL PERIOD"
# LOGIT TO THE "STEADY STATE" LOGIT.
# THIS IS JUST A GEOMETRIC DISTRIBUTION WITH
# Q[i] = PROBABILITY THE CUSTOMER SWITCHES TO THE STEADY STATE MODEL.
# X1[i,s] = PROBABILITY THE CUSTOMER IS USING THE STEADY STATE MODEL IN
PERIOD T.

for (i in 1:NHH){
q[i]<- max(0.0000, min(0.98, 1/(1+exp(-(c0[i])))))
for (s in 1:18) {
  X1[i,s] <- 1-(pow((1-q[i]),(s-1)))
}
}

# CHOICE MODELS: LOGIT WITH NO-PURCHASE OPTION.
# TRIAL=0 POST-TRIAL=1
# FOUR ALTERATIVES (j): j=1=CATALOG, j=2=INTERNET, j=3=STORE, j=4=NO-
PURCHASE
# CS = CATALOGS; SAME FOR ALL ALTERNATIVES; COEF VARIES ACROSS ALTS.
# ES = EMAILS; SAME FOR ALL ALTERNATIVES; COEF VARIES ACROSS ALTS.
# LC = STATE DEPENDENCE; VARIES ACROSS ALTS; COEFF SAME ACROSS ALTS.
# ALL COEFS VARY ACROSS HOUSEHOLDS.

# UTILITY FOR ALTERNATIVE 4 IS SET =
JUSTeta0[h[n]]*trend[n]+iota40[h[n]]*Q4[n]+ b30[h[n]]*LC[n,4]
# BECAUSE THE COEFS FOR CS AND ES ARE ALTERNATIVE SPECIFIC SO THE COEF FOR
ONE OF THE ALTS
# IS NOT IDENTIFIED.  WE ARBITRARILY MAKE THIS ALTERNATIVE 4.
# LOOP IS OVER OBSERVATIONS ("NOBS"); 21 OBS PER HH; H[N] SIGNIFIES THE HH
# FOR OBSERVATION N.


for (n in 1:NOBS) {
```

```
U0[n,1] <- b00[h[n],1]+b10[h[n],1]*CS[n]+b20[h[n],1]*ES[n] +
b30[h[n]]*LC[n,1]
U0[n,2] <- b00[h[n],2]+b10[h[n],2]*CS[n]+b20[h[n],2]*ES[n] +
b30[h[n]]*LC[n,2]
U0[n,3] <- b00[h[n],3]+b10[h[n],3]*CS[n]+b20[h[n],3]*ES[n] +
b30[h[n]]*LC[n,3]
U0[n,4] <- eta0[h[n]]*trend[n]+iota40[h[n]]*Q4[n]
U1[n,1] <- b01[h[n],1]+b11[h[n],1]*CS[n]+b21[h[n],1]*ES[n] +
b31[h[n]]*LC[n,1]
U1[n,2] <- b01[h[n],2]+b11[h[n],2]*CS[n]+b21[h[n],2]*ES[n] +
b31[h[n]]*LC[n,2]
U1[n,3] <- b01[h[n],3]+b11[h[n],3]*CS[n]+b21[h[n],3]*ES[n] +
b31[h[n]]*LC[n,3]
U1[n,4] <- eta1[h[n]]*trend[n]+iota41[h[n]]*Q4[n]

vbot0[n]<-exp(U0[n,1])+exp(U0[n,2])+exp(U0[n,3])+exp(U0[n,4])
vbot1[n]<-exp(U1[n,1])+exp(U1[n,2])+exp(U1[n,3])+exp(U0[n,4])

pr1[n,1]<- (1-X1[h[n],t[n]])*((exp(U0[n,1])/vbot0[n]))+
X1[h[n],t[n]]*((exp(U1[n,1])/vbot1[n]))
pr[n,1]<-max(.00000000,min(.98, pr1[n,1]))

pr1[n,2]<- (1-X1[h[n],t[n]])*((exp(U0[n,2])/vbot0[n]))+
X1[h[n],t[n]]*((exp(U1[n,2])/vbot1[n]))
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))

pr1[n,3]<- (1-X1[h[n],t[n]])*((exp(U0[n,3])/vbot0[n]))+
X1[h[n],t[n]]*((exp(U1[n,3])/vbot1[n]))
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))

pr1[n,4]<- (1-X1[h[n],t[n]])*((exp(U0[n,4])/vbot0[n]))+
X1[h[n],t[n]]*((exp(U1[n,4])/vbot1[n]))
pr[n,4]<-max(.00000000, min(.98, pr1[n,4]))


Y[n,1:4] ~ dmulti( pr[n,1:4] , 4)

}


# PRIORS
for (n in 1:NHH)  {

b30[n] ~ dnorm(mu30,prec30)
b31[n] ~ dnorm(mu31,prec31)

c0[n] ~ dnorm(muc0,precc0)

eta0[n] ~ dnorm(mue0,prece0)
eta1[n] ~ dnorm(mue1,prece1)
iota40[n] ~ dnorm(mui40,preci40)
iota41[n] ~ dnorm(mui41,preci41)

for (k in 1:3) {
  b00[n,k] ~ dnorm(mu00[k],prec00)
  b10[n,k] ~ dnorm(mu10[k],prec10)
  b20[n,k] ~ dnorm(mu20[k],prec20)
  b01[n,k] ~ dnorm(mu01[k],prec01)
  b11[n,k] ~ dnorm(mu11[k],prec11)
  b21[n,k] ~ dnorm(mu21[k],prec21)
```

```
}
}


for (k in 1:3) {
mu00[k]~dnorm(0,.00001)
mu10[k]~dnorm(0,.00001)
mu20[k]~dnorm(0,.00001)
mu01[k]~dnorm(0,.00001)
mu11[k]~dnorm(0,.00001)
mu21[k]~dnorm(0,.00001)
}

mu30~dnorm(0,.00001)
mu31~dnorm(0,.00001)

# no-purchase option priors

mue0~dnorm(0,.00001)
mue1~dnorm(0,.00001)
mui40~dnorm(0,.00001)
mui41~dnorm(0,.00001)

prece0~dgamma(0.5,0.5)
prece1~dgamma(0.5,0.5)
preci40~dgamma(0.5,0.5)
preci41~dgamma(0.5,0.5)

# learning priors
muc0~dnorm(0,.00001)
precc0~dgamma(0.5,0.5)


# Choice model  priors precisions

prec00~dgamma(0.5,0.5)
prec01~dgamma(0.5,0.5)
prec10~dgamma(0.5,0.5)
prec11~dgamma(0.5,0.5)
prec20~dgamma(0.5,0.5)
prec21~dgamma(0.5,0.5)
prec30~dgamma(0.5,0.5)
prec31~dgamma(0.5,0.5)
```

**APPENDIX 1b**: M1, M2, M3, and M4 SYNTAXES

---

### M1 –Multinomial logit model

```
model
{


for (n in 1:NOBS) {

U[n,1] <- b0[h[n],1]+b1[h[n],1]*CS[n]+b2[h[n],1]*ES[n] + b3[h[n]]*LC[n,1]
```

```
U[n,2] <- b0[h[n],2]+b1[h[n],2]*CS[n]+b2[h[n],2]*ES[n] + b3[h[n]]*LC[n,2]
U[n,3] <- b3[h[n]]*LC[n,3]


vbot[n]<-exp(U[n,1])+exp(U[n,2])+exp(U[n,3])


pr1[n,1]<- exp(U[n,1])/vbot[n]))
pr[n,1]<-max(.00000000,min(.98, pr1[n,1]))

pr1[n,2]<- exp(U[n,2])/vbot[n]
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))

pr1[n,3]<- exp(U[n,3])/vbot[n]
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))


Y[n,1:3] ~ dmulti( pr[n,1:3] , 1)

}

# first stage PRIORS
for (n in 1:NHH)  {

b3[n] ~ dnorm(mu3,prec3)


for (k in 1:2) {
  b0[n,k] ~ dnorm(mu0[k],prec0)
  b1[n,k] ~ dnorm(mu1[k],prec1)
  b2[n,k] ~ dnorm(mu2[k],prec2)

}
}

# Choice model  2d stage priors

for (k in 1:2) {
mu0[k]~dnorm(0,.00001)
mu1[k]~dnorm(0,.00001)
mu2[k]~dnorm(0,.00001)

}
mu3~dnorm(0,.00001)


prec0~dgamma(0.5,0.5)
prec1~dgamma(0.5,0.5)
prec2~dgamma(0.5,0.5)
prec3~dgamma(0.5,0.5)
```

**M2 – Multinomial logit model distinct in two periods**

```
model
{
for (n in 1:NOBS) {
```

```
Y[n,1:3] ~ dmulti( pr[n,1:3] , 1)

U[n,1] <- b0[h[n],1]+  b0p[h[n],1]*dummy1[n] + b1[h[n],1]*CS[n]+
b1p[h[n],1]*dummy1[n]*CS[n]+b2[h[n],1]*ES[n] +b2p[h[n],1]*dummy1[n]*ES[n] +
b3[h[n]]*LC[n,1] + b3p[h[n]]*dummy1[n]*LC[n,1]
U[n,2] <- b0[h[n],2]+  b0p[h[n],2]*dummy1[n] + b1[h[n],2]*CS[n]+
b1p[h[n],2]*dummy1[n]*CS[n]+b2[h[n],2]*ES[n] +b2p[h[n],2]*dummy1[n]*ES[n] +
b3[h[n]]*LC[n,2] + b3p[h[n]]*dummy1[n]*LC[n,2]
U[n,3] <- b3[h[n]]*LC[n,3] + b3p[h[n]]*dummy1[n]*LC[n,3]

vbot[n]<-exp(U[n,1])+exp(U[n,2])+exp(U[n,3])

pr1[n,1]<- exp(U[n,1])/vbot[n]
pr[n,1]<-max(.00000000,min(.98, pr1[n,1]))

pr1[n,2]<- exp(U[n,2])/vbot[n]
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))

pr1[n,3]<- exp(U[n,3])/vbot[n]
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))

}

# first stage PRIORS
for (n in 1:NHH)  {
b3[n] ~ dnorm(mu3,prec3)
b3p[n] ~ dnorm(mu3p,prec3p)
for (k in 1:2) {
  b0[n,k] ~ dnorm(mu0[k],prec0)
  b0p[n,k] ~ dnorm(mu0p[k],prec0p)
  b1[n,k] ~ dnorm(mu1[k],prec1)
  b2[n,k] ~ dnorm(mu2[k],prec2)
  b1p[n,k] ~ dnorm(mu1p[k],prec1p)
  b2p[n,k] ~ dnorm(mu2p[k],prec2p)
  }
}


# Choice model  2d stage priors
for (k in 1:2) {
mu0[k]~dnorm(0,.00001)
mu1[k]~dnorm(0,.00001)
mu2[k]~dnorm(0,.00001)
mu0p[k]~dnorm(0,.00001)
mu1p[k]~dnorm(0,.00001)
mu2p[k]~dnorm(0,.00001)
}

mu3~dnorm(0,.00001)
mu3p~dnorm(0,.00001)
prec0~dgamma(0.5,5)
prec1~dgamma(0.5,5)
prec2~dgamma(0.5,5)
prec3~dgamma(0.5,5)
prec0p~dgamma(0.5,5)
prec1p~dgamma(0.5,5)
prec2p~dgamma(0.5,5)
prec3p~dgamma(0.5,5)

}
```

## M3 –Multinomial logit channel selection switching model

```
model
{

for (i in 1:NHH){
q[i]<- max(0.0000, min(0.98, 1/(1+exp(-(c0[i])))))
for (s in 1:Tpo[i]) {
X1[i,s] <- 1-(pow((1-q[i]),(s-1)))
}
}

for (n in 1:NOBS) {

U0[n,1] <- b00[h[n],1]+b10[h[n],1]*CS[n]+b20[h[n],1]*ES[n] +
b30[h[n]]*LC[n,1]
U0[n,2] <- b00[h[n],2]+b10[h[n],2]*CS[n]+b20[h[n],2]*ES[n] +
b30[h[n]]*LC[n,2]
U0[n,3] <- b30[h[n]]*LC[n,3]

U1[n,1] <- b01[h[n],1]+b11[h[n],1]*CS[n]+b21[h[n],1]*ES[n] +
b31[h[n]]*LC[n,1]
U1[n,2] <- b01[h[n],2]+b11[h[n],2]*CS[n]+b21[h[n],2]*ES[n] +
b31[h[n]]*LC[n,2]
U1[n,3] <- b31[h[n]]*LC[n,3]

vbot0[n]<-exp(U0[n,1])+exp(U0[n,2])+exp(U0[n,3])
vbot1[n]<-exp(U1[n,1])+exp(U1[n,2])+exp(U1[n,3])

pr1[n,1]<- (1-X1[h[n],po[n]])*((exp(U0[n,1])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,1])/vbot1[n]))
pr[n,1]<-max(.00000000,min(.98, pr1[n,1]))

pr1[n,2]<- (1-X1[h[n],po[n]])*((exp(U0[n,2])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,2])/vbot1[n]))
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))

pr1[n,3]<- (1-X1[h[n],po[n]])*((exp(U0[n,3])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,3])/vbot1[n]))
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))



Y[n,1:3] ~ dmulti( pr[n,1:3] , 1)

}

# first stage PRIORS
for (n in 1:NHH)  {

b30[n] ~ dnorm(mu30,prec30)
b31[n] ~ dnorm(mu31,prec31)

c0[n] ~ dnorm(muc0,precc0)

for (k in 1:2) {
  b00[n,k] ~ dnorm(mu00[k],prec00)
```

```
  b10[n,k] ~ dnorm(mu10[k],prec10)
  b20[n,k] ~ dnorm(mu20[k],prec20)
  b01[n,k] ~ dnorm(mu01[k],prec01)
  b11[n,k] ~ dnorm(mu11[k],prec11)
  b21[n,k] ~ dnorm(mu21[k],prec21)
}
}

# Choice model  2d stage priors

for (k in 1:2) {
mu00[k]~dnorm(0,.00001)
mu10[k]~dnorm(0,.00001)
mu20[k]~dnorm(0,.00001)
mu01[k]~dnorm(0,.00001)
mu11[k]~dnorm(0,.00001)
mu21[k]~dnorm(0,.00001)
}
mu30~dnorm(0,.00001)
mu31~dnorm(0,.00001)

prec00~dgamma(0.5,0.5)
prec01~dgamma(0.5,0.5)
prec10~dgamma(0.5,0.5)
prec11~dgamma(0.5,0.5)
prec20~dgamma(0.5,0.5)
prec21~dgamma(0.5,0.5)
prec30~dgamma(0.5,0.5)
prec31~dgamma(0.5,0.5)


# learning 2d stage priors
muc0~dnorm(0,.00001)
precc0~dgamma(0.5,0.5)


}
```

**M4 –Multinomial logit channel selection switching model (with age and gender)**

```
model
{

for (i in 1:NHH){
q[i]<- max(0.0000, min(0.98, 1/(1+exp(-(c0[i] +c1*sex[i]+c2*age[i])))))
for (s in 1:Tpo[i]) {
X1[i,s] <- 1-(pow((1-q[i]),(s-1)))
}
}

for (n in 1:NOBS) {

U0[n,1] <- b00[h[n],1]+b10[h[n],1]*CS[n]+b20[h[n],1]*ES[n] +
b30[h[n]]*LC[n,1]
U0[n,2] <- b00[h[n],2]+b10[h[n],2]*CS[n]+b20[h[n],2]*ES[n] +
b30[h[n]]*LC[n,2]
U0[n,3] <- b30[h[n]]*LC[n,3]
```

```
U1[n,1] <- b01[h[n],1]+b11[h[n],1]*CS[n]+b21[h[n],1]*ES[n] +
b31[h[n]]*LC[n,1]
U1[n,2] <- b01[h[n],2]+b11[h[n],2]*CS[n]+b21[h[n],2]*ES[n] +
b31[h[n]]*LC[n,2]
U1[n,3] <- b31[h[n]]*LC[n,3]

vbot0[n]<-exp(U0[n,1])+exp(U0[n,2])+exp(U0[n,3])
vbot1[n]<-exp(U1[n,1])+exp(U1[n,2])+exp(U1[n,3])

pr1[n,1]<- (1-X1[h[n],po[n]])*((exp(U0[n,1])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,1])/vbot1[n]))
pr[n,1]<-max(.00000000,min(.98, pr1[n,1]))

pr1[n,2]<- (1-X1[h[n],po[n]])*((exp(U0[n,2])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,2])/vbot1[n]))
pr[n,2]<-max(.00000000, min(.98, pr1[n,2]))

pr1[n,3]<- (1-X1[h[n],po[n]])*((exp(U0[n,3])/vbot0[n]))+
X1[h[n],po[n]]*((exp(U1[n,3])/vbot1[n]))
pr[n,3]<-max(.00000000, min(.98, pr1[n,3]))



Y[n,1:3] ~ dmulti( pr[n,1:3] , 1)

}

# first stage PRIORS
for (n in 1:NHH)  {

b30[n] ~ dnorm(mu30,prec30)
b31[n] ~ dnorm(mu31,prec31)

c0[n] ~ dnorm(muc0,precc0)

for (k in 1:2) {
  b00[n,k] ~ dnorm(mu00[k],prec00)
  b10[n,k] ~ dnorm(mu10[k],prec10)
  b20[n,k] ~ dnorm(mu20[k],prec20)
  b01[n,k] ~ dnorm(mu01[k],prec01)
  b11[n,k] ~ dnorm(mu11[k],prec11)
  b21[n,k] ~ dnorm(mu21[k],prec21)
}
}

# Choice model  2d stage priors

for (k in 1:2) {
mu00[k]~dnorm(0,.00001)
mu10[k]~dnorm(0,.00001)
mu20[k]~dnorm(0,.00001)
mu01[k]~dnorm(0,.00001)
mu11[k]~dnorm(0,.00001)
mu21[k]~dnorm(0,.00001)
}
mu30~dnorm(0,.00001)
mu31~dnorm(0,.00001)
c1 ~dnorm(muc1,precc1)
c2 ~dnorm(muc2,precc2)
```

```
prec00~dgamma(0.5,0.5)
prec01~dgamma(0.5,0.5)
prec10~dgamma(0.5,0.5)
prec11~dgamma(0.5,0.5)
prec20~dgamma(0.5,0.5)
prec21~dgamma(0.5,0.5)
prec30~dgamma(0.5,0.5)
prec31~dgamma(0.5,0.5)


# learning 2d stage priors
muc0~dnorm(0,.00001)
muc1~dnorm(0,.00001)
muc2~dnorm(0,.00001)
precc0~dgamma(0.5,0.5)
precc1~dgamma(0.5,0.5)
precc2~dgamma(0.5,0.5)

}
```

**APPENDIX 2**: **Endogeneity Bias in Direct Marketing Communication (2sls approach)**

---

We used a two stage least squares approach to handle the endogeneity problem which concern the variable CS and ES (catalogs and e-mails sent). We perform two regressions (an OLS regression for CS and a poisson regression for ES). Instead of entering the actual values of catalogs sent and emails sent in the multinomial logit channel choice model we enter their predicted values computed by using the predicted catalogs sent and emails sent variables that came from the above mentioned regressions.

$$CS_{ht} = a_0 + a_1 age_h + a_2 gender_h + a_3 Q2_{ht} + a_4 Q3_{ht} + a_5 Q4_{ht} + a_6 Npolag_{ht} + a_7 Corientation_h + \xi_{ht} \quad \text{(a)}$$

$$ES_{ht} = b_0 + b_1 age_h + b_2 gender_h + b_3 Q2_{ht} + b_4 Q3_{ht} + b_5 Q4_{ht} + b_6 Npolag_{ht} + b_7 Iorientation_h + \zeta_{ht} \quad \text{(b)}$$

Where:

Q4= fourth quarter (oct-dec)

Q3= third quarter (jul-sep)

Q2= second quarter (apr-jun)

Npolag= number of purchase occasions lagged per quarter.

Gender= 1 (female) 0 (male)

Iorientation= variable which assumes value 1 if the household has used at least once Internet as channel. It starts to assume value 1 after the first time that the internet choice happened.

Corientation= variable which assumes value 1 if the household has used at least once Catalog as channel. It starts to assume value 1 after the first time that the internet choice happened.

*Parameter Estimates Poisson Regression with emails sent as dependent variable*

```
Poisson regression                                 Number of obs   =       2079
                                                   LR chi2(7)      =    4858.20
                                                   Prob > chi2     =     0.0000
Log likelihood =  -2509.268                        Pseudo R2       =     0.4919

------------------------------------------------------------------------------
  EMAILS SENT|      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     Npo_lag |   .1109095   .0228795     4.85   0.000     .0660665    .1557526
         sex |    .243712   .0412121     5.91   0.000     .1629379    .3244862
         age |   .0046191   .0017445     2.65   0.008        .0012    .0080383
      ever_I |   3.006481    .053177    56.54   0.000     2.902256    3.110706
          q1 |  -.0337873   .0563565    -0.60   0.549     -.144244    .0766694
          q2 |  -.1651153   .0569886    -2.90   0.004    -.2768108   -.0534197
          q3 |  -.1332509   .0597119    -2.23   0.026    -.2502841   -.0162177
       _cons |  -1.486225   .0648627   -22.91   0.000    -1.613354   -1.359097
------------------------------------------------------------------------------
```

*Parameter Estimates OLS Regression with catalogs sent as dependent variable*

```
OLS regression

      Source |       SS       df       MS              Number of obs =  342210
-------------+------------------------------           F(  5.342204) =59532.74
       Model |  135495.03     5  27099.006             Prob > F      =  0.0000
    Residual |  155769.56342204  .455195029            R-squared     =  0.4652
-------------+------------------------------           Adj R-squared =  0.4652
       Total |  291264.59342209  .851130712            Root MSE      =  .67468

------------------------------------------------------------------------------
CATALOGS SENT|      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          Q4 |   .0895978   .0034214    26.19   0.000      .082892    .0963037
          Q3 |   -.120479   .0032774   -36.76   0.000    -.1269026   -.1140554
          Q2 |  -.1130488   .0032758   -34.51   0.000    -.1194693   -.1066282
     npo_lag |   .9825338   .0018251   538.35   0.000     .9789567    .9861109
      gender |   .0091336   .0024818     3.68   0.000     .0042694    .0139977
       _cons |   .5436449   .0030458   178.49   0.000     .5376752    .5496145
------------------------------------------------------------------------------
```

## APPENDIX 3: Convergence

In order to assess the convergence I have considered the mean values of the first stage priors parameters which I remember below (n indicates individuals and k channel alternative):

b00[n,k] ~ dnorm(mu00[k],prec00)

b01[n,k] ~ dnorm(mu01[k],prec01)

b10[n,k] ~ dnorm(mu10[k],prec10)

b11[n,k] ~ dnorm(mu11[k],prec11)

b20[n,k] ~ dnorm(mu20[k],prec20)

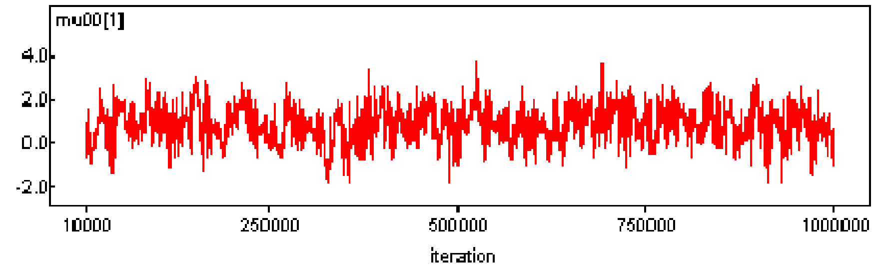
b21[n,k] ~ dnorm(mu21[k],prec21)

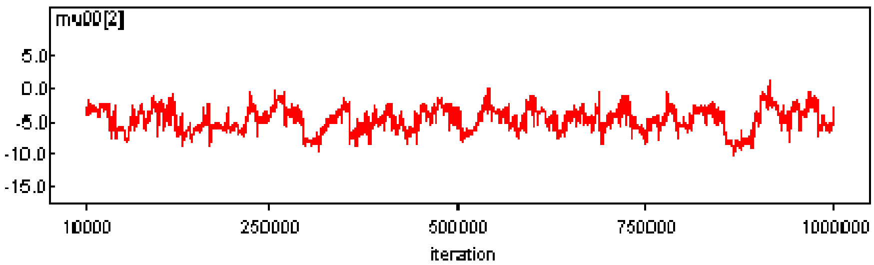
b30[n] ~ dnorm(mu30,prec30)

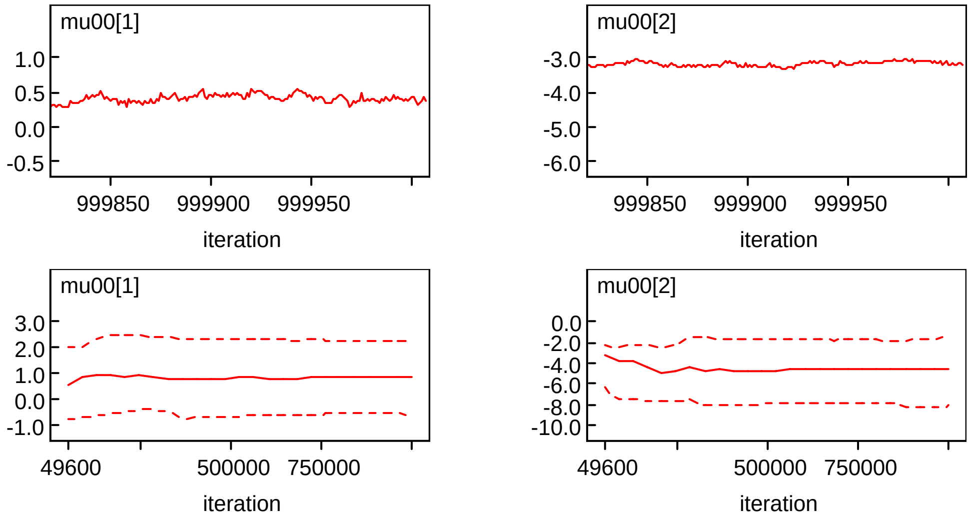
b31[n] ~ dnorm(mu31,prec31)

### Intercept Trial

Graphs: History, Trace and Quantiles and with 1 million iterations






Comparison of the Mean Results with 10.000 iterations and 100.000 iterations

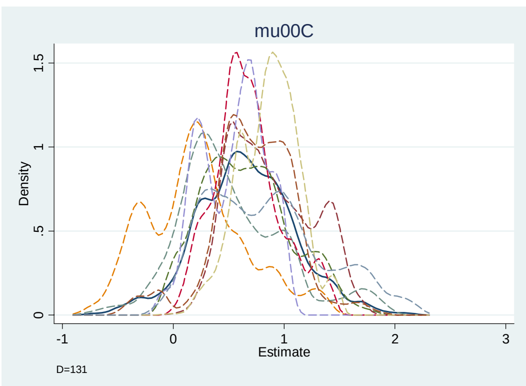
| node | mean | sd | MC error | 2.50% | median | 97.50% | Start (iter. n) | sample |
|---|---|---|---|---|---|---|---|---|
| mu00[1] | 0.96 | 0.30 | 0.03 | 0.34 | 0.99 | 1.57 | 4000 | 10001 |
| mu00[1] | 0.83 | 0.53 | 0.05 | 0.03 | 0.79 | 1.98 | 90000 | 10001 |
| mu00[2] | -3.05 | 1.13 | 0.11 | -5.62 | -2.58 | -1.54 | 4000 | 10001 |
| mu00[2] | -5.38 | 0.36 | 0.03 | -6.05 | -5.37 | -4.67 | 90000 | 10001 |

Tests: COMPARISON OF DIFFERENT CHAINS

In order to compute the B.G.R. statistic multiple chains should be available. Anyway, this is not possible using 100.000 iterations because the computer "exceeds the RAM limits" if we set more than one chain as initial values.
However, it's possible to compute this statistic using the results from a single (long) chain by dividing the chain into a number of pieces and treating each piece as if it were a different chain. Thus, I divided the long chain (100.000 iterations) in 9 pieces. I dropped the first 10.000 iterations (burn-in period).
Thus, in the following graphs we have 9 different chains of 10.000 iterations.
(see Brooks and Gelman 1998 for details)

Catalog

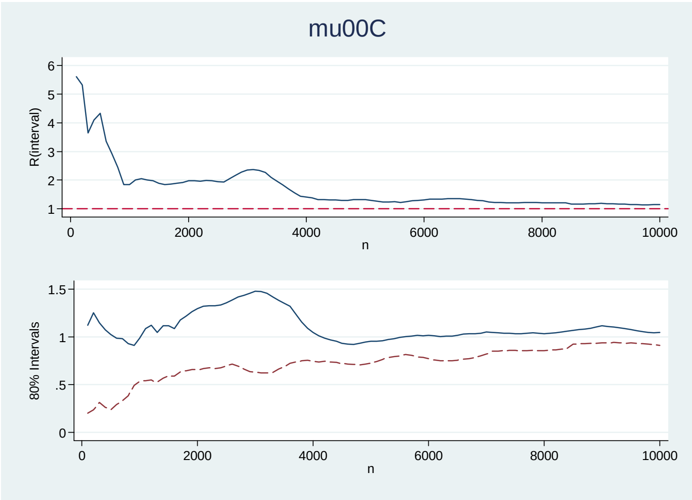
Density Graph Intercept *Trial* (catalog): comparison of nine density graphs (mean graph – bold line).



**B.G.R. statistic (R)**:
R is calculated for increasing chain sizes. For subchains consisting of the 100 values. R should approach 1 if the chains have converged.

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | Average mu00C | R Brooks-Gelman-rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.10 | 0.11 | 0.28 | 0.37 | 0.37 | 0.22 | 0.07 | 0.15 | 0.13 | 0.20 | **5.6** |
| 200 | 0.23 | 0.11 | 0.42 | 0.18 | 0.12 | 0.18 | 0.16 | 0.21 | 0.52 | 0.24 | **5.3** |
| 300 | 0.45 | 0.38 | 0.26 | 0.25 | 0.16 | 0.20 | 0.31 | 0.28 | 0.53 | 0.31 | **3.7** |
| 400 | 0.55 | 0.23 | 0.17 | 0.24 | 0.20 | 0.30 | 0.22 | 0.15 | 0.30 | 0.26 | **4.1** |
| 500 | 0.29 | 0.23 | 0.23 | 0.21 | 0.25 | 0.21 | 0.17 | 0.24 | 0.32 | 0.24 | **4.3** |
| 600 | 0.30 | 0.23 | 0.50 | 0.32 | 0.28 | 0.18 | 0.19 | 0.30 | 0.35 | 0.29 | **3.4** |
| 700 | 0.27 | 0.26 | 0.60 | 0.34 | 0.27 | 0.30 | 0.19 | 0.40 | 0.39 | 0.33 | **2.9** |
| 800 | 0.27 | 0.35 | 0.52 | 0.63 | 0.28 | 0.28 | 0.26 | 0.44 | 0.41 | 0.38 | **2.4** |
| 900 | 0.27 | 0.36 | 0.48 | 0.81 | 0.49 | 0.30 | 0.33 | 0.97 | 0.43 | 0.49 | **1.8** |
| 1000 | 0.53 | 0.35 | 0.46 | 0.77 | 0.49 | 0.30 | 0.39 | 1.10 | 0.41 | 0.53 | **1.9** |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 9000 | 0.95 | 0.96 | 1.27 | 1.70 | 0.85 | 0.54 | 0.56 | 0.71 | 0.91 | 0.94 | **1.2** |
| 9100 | 0.95 | 0.96 | 1.27 | 1.69 | 0.87 | 0.56 | 0.57 | 0.71 | 0.89 | 0.94 | **1.2** |
| 9200 | 0.95 | 1.00 | 1.26 | 1.66 | 0.90 | 0.57 | 0.56 | 0.70 | 0.85 | 0.94 | **1.2** |
| 9300 | 0.95 | 1.04 | 1.25 | 1.64 | 0.90 | 0.56 | 0.56 | 0.70 | 0.83 | 0.94 | **1.2** |
| 9400 | 0.95 | 1.07 | 1.25 | 1.62 | 0.90 | 0.56 | 0.56 | 0.70 | 0.81 | 0.93 | **1.2** |
| 9500 | 0.95 | 1.10 | 1.24 | 1.61 | 0.90 | 0.57 | 0.56 | 0.70 | 0.80 | 0.94 | **1.2** |
| 9600 | 0.96 | 1.12 | 1.24 | 1.59 | 0.89 | 0.56 | 0.55 | 0.69 | 0.81 | 0.94 | **1.1** |
| 9700 | 0.97 | 1.12 | 1.23 | 1.57 | 0.89 | 0.56 | 0.53 | 0.68 | 0.80 | 0.93 | **1.1** |
| 9800 | 0.98 | 1.14 | 1.21 | 1.56 | 0.89 | 0.55 | 0.54 | 0.67 | 0.78 | 0.92 | **1.1** |
| 9900 | 0.98 | 1.15 | 1.16 | 1.55 | 0.88 | 0.55 | 0.55 | 0.67 | 0.76 | 0.92 | **1.1** |
| 10000 | 0.97 | 1.15 | 1.13 | 1.55 | 0.88 | 0.54 | 0.56 | 0.68 | 0.74 | 0.91 | **1.1** |

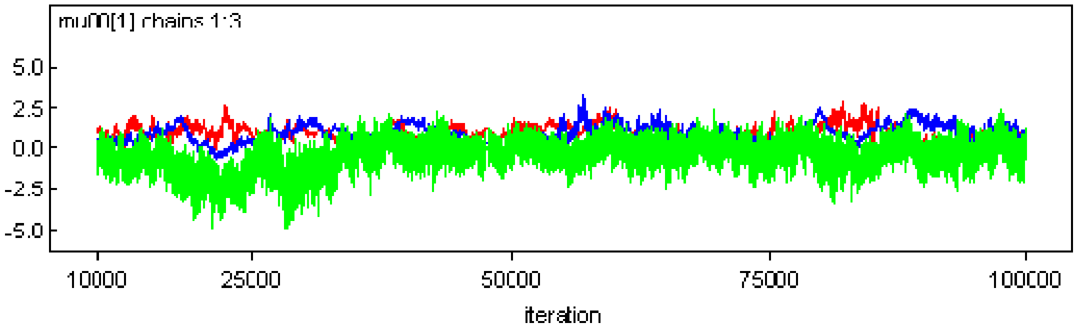
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**.



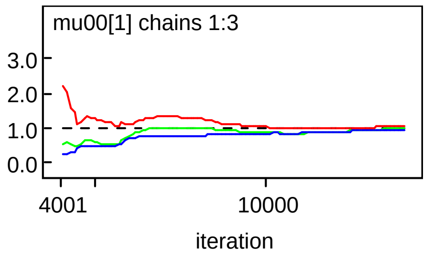

The plots assume that the chains start from different and over-dispersed initial values. As the chains come closer into agreement the variability of the pooled chains should be similar to the average variability of the individual chains that is R becames close to one.
The second plot shows the variance or interval based on all chains pooled (solid line) and on the average of the subchains (dashed line). These should stabilise into horizontal lines.

Even if it is not possible to obtain B.G.R. statistic using winbugs with 100.000 iterations, it is possible to obtain this statistic if we ask only 10.000 iterations. In order to reinforce the above results I run the model with 3 different sets of initial values. I obtained these initial values from a *trial* run with null parameters, then using the state space command after few iterations in order to obtain different initial value chains (see Congdon 2003 – Applied Bayesian Statistic p. 18). Then , I run winbugs with 10.000 iterations.

**History Graph: comparison of three different chains**



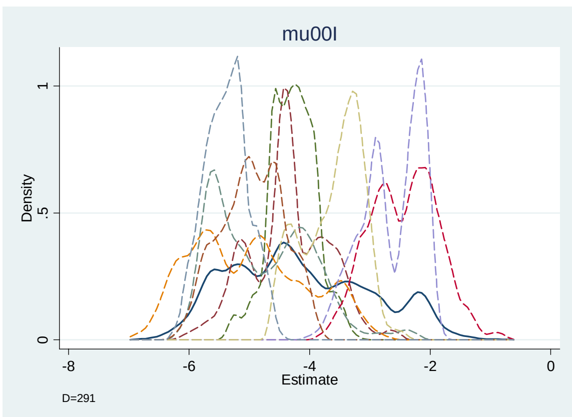

**Brooks-Gelman-Rubin plot**




Green line: the width of the central 80% interval of the pooled runs.
Blue line: The average width of the 80% interval within the individual runs.
The pooled and within interval widths are normalized to be less than 1 for plotting purpose.
Red line: the ratio of pooled/within (= R) calculated in bins of length 50

<u>Internet</u>

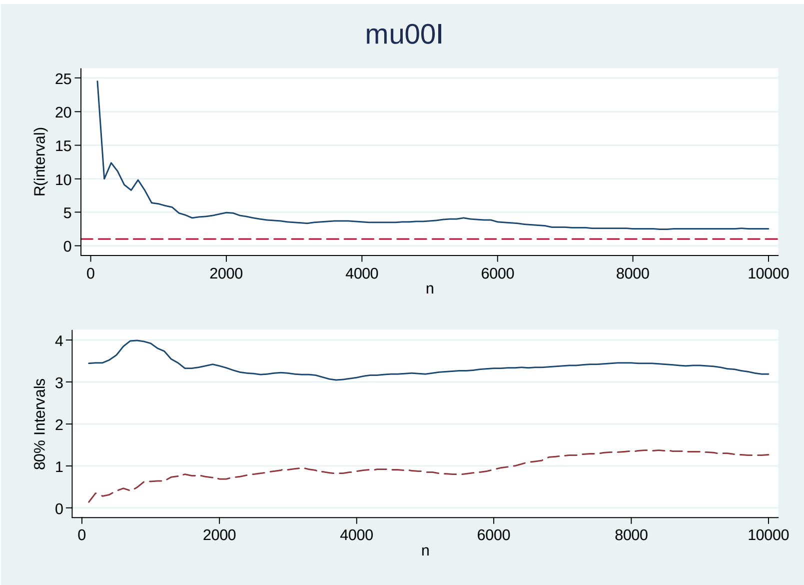
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



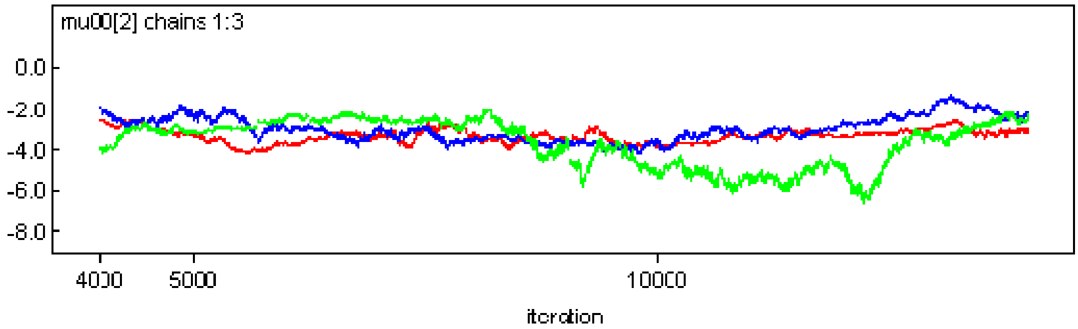

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | Average mu00I | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.23 | 0.06 | 0.10 | 0.10 | 0.18 | 0.23 | 0.10 | 0.08 | 0.17 | 0.14 | **24.5** |
| 200 | 0.77 | 0.40 | 0.21 | 0.14 | 0.29 | 0.24 | 0.16 | 0.23 | 0.68 | 0.35 | **10.0** |
| 300 | 0.28 | 0.22 | 0.37 | 0.17 | 0.29 | 0.32 | 0.16 | 0.26 | 0.45 | 0.28 | **12.4** |
| 400 | 0.28 | 0.33 | 0.59 | 0.14 | 0.40 | 0.33 | 0.22 | 0.28 | 0.29 | 0.32 | **11.1** |
| 500 | 0.41 | 0.47 | 0.73 | 0.16 | 0.65 | 0.29 | 0.28 | 0.30 | 0.30 | 0.40 | **9.1** |
| 600 | 0.35 | 0.54 | 1.07 | 0.21 | 0.54 | 0.22 | 0.53 | 0.42 | 0.30 | 0.46 | **8.3** |
| 700 | 0.27 | 0.38 | 0.91 | 0.25 | 0.40 | 0.21 | 0.52 | 0.41 | 0.30 | 0.41 | **9.8** |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 9000 | 1.13 | 0.97 | 1.49 | 2.03 | 1.22 | 1.39 | 1.33 | 1.29 | 1.19 | 1.34 | **2.5** |
| 9100 | 1.06 | 0.96 | 1.45 | 2.02 | 1.24 | 1.36 | 1.32 | 1.28 | 1.18 | 1.32 | **2.6** |
| 9200 | 1.01 | 0.95 | 1.43 | 2.01 | 1.25 | 1.36 | 1.32 | 1.27 | 1.18 | 1.31 | **2.6** |
| 9300 | 0.98 | 0.97 | 1.40 | 1.98 | 1.28 | 1.34 | 1.32 | 1.26 | 1.18 | 1.30 | **2.6** |
| 9400 | 0.95 | 0.99 | 1.36 | 1.96 | 1.31 | 1.33 | 1.31 | 1.26 | 1.18 | 1.30 | **2.6** |
| 9500 | 0.92 | 1.02 | 1.30 | 1.94 | 1.32 | 1.33 | 1.30 | 1.23 | 1.18 | 1.28 | **2.6** |
| 9600 | 0.89 | 1.05 | 1.22 | 1.90 | 1.32 | 1.30 | 1.29 | 1.21 | 1.18 | 1.26 | **2.6** |
| 9700 | 0.88 | 1.10 | 1.19 | 1.90 | 1.35 | 1.27 | 1.28 | 1.18 | 1.18 | 1.26 | **2.6** |
| 9800 | 0.87 | 1.12 | 1.18 | 1.91 | 1.38 | 1.25 | 1.28 | 1.15 | 1.17 | 1.26 | **2.6** |
| 9900 | 0.87 | 1.12 | 1.16 | 1.95 | 1.40 | 1.22 | 1.27 | 1.12 | 1.17 | 1.25 | **2.5** |
| 10000 | 0.87 | 1.20 | 1.14 | 2.00 | 1.44 | 1.19 | 1.25 | 1.11 | 1.17 | 1.26 | **2.1** |

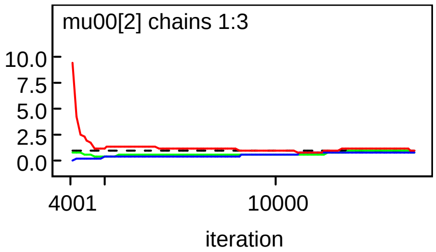
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**.




**History Graph: comparison of three different chains**



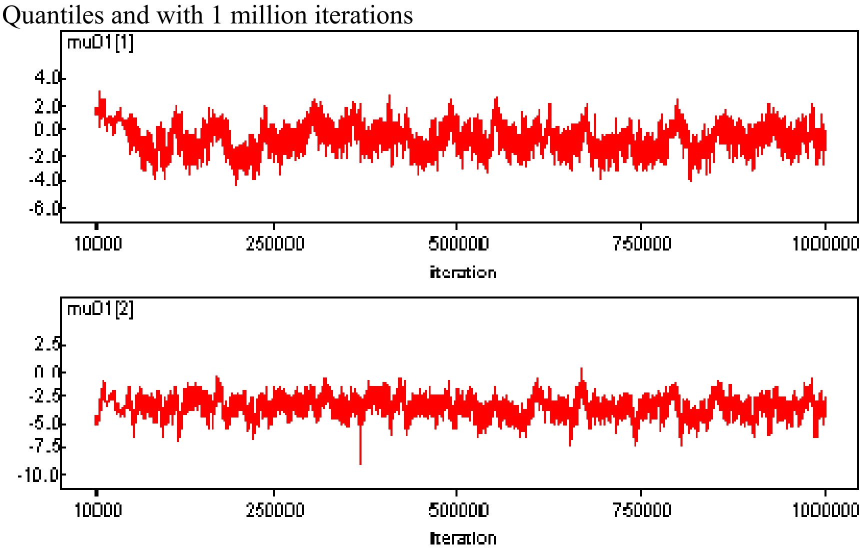

**Brooks-Gelman-Rubin plot**

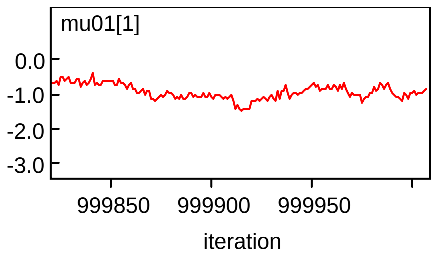


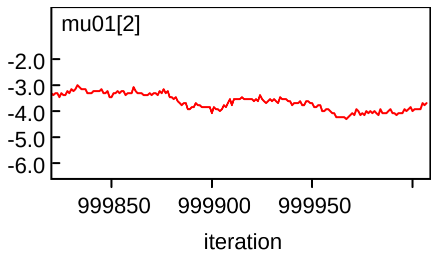

Red line: the ratio of pooled/within (= R) calculated in bins of length 50.

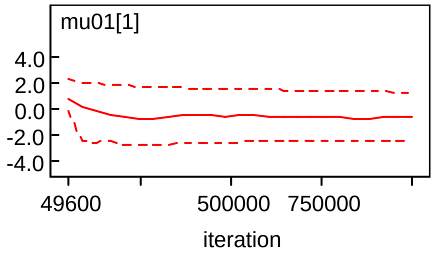
## Intercept Post-trial

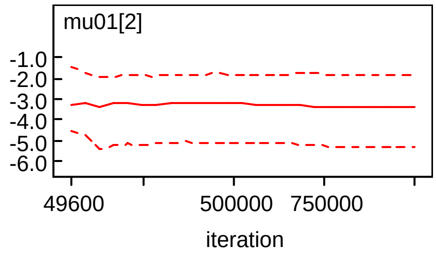
History, Trace and Quantiles and with 1 million iterations





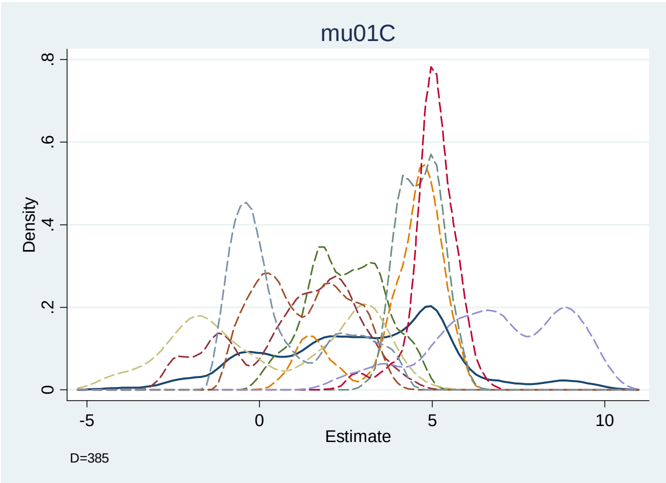

Comparison of the Mean Results with 10.000 iterations and 100.000 iterations

| **node** | **mean** | **sd** | **MC error** | **2.50%** | **median** | **97.50%** | **start** | **sample** |
|---|---|---|---|---|---|---|---|---|
| mu01[1] | -0.80 | 1.15 | 0.11 | -2.51 | -0.97 | 1.70 | 4000 | 10001 |
| mu01[1] | 0.72 | 1.56 | 0.16 | -1.05 | 0.03 | 3.90 | 90000 | 10001 |
| mu01[2] | -4.06 | 0.70 | 0.07 | -5.60 | -4.00 | -2.81 | 4000 | 10001 |
| mu01[2] | -2.11 | 0.85 | 0.08 | -3.39 | -2.37 | 0.22 | 90000 | 10001 |

**Formal Tests: COMPARISON OF DIFFERENT CHAINS**

Catalog

**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**

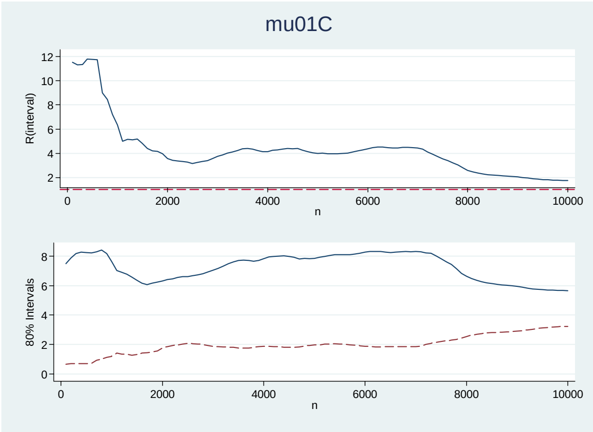



**B.G.R. statistic (R)**

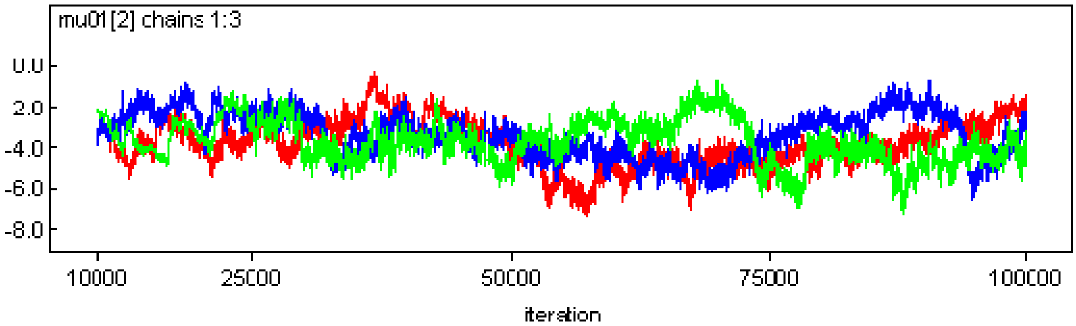
| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.16 | 0.51 | 1.16 | 0.75 | 0.38 | 0.80 | 1.04 | 0.71 | 0.35 | 0.65 | **11.5** |
| 200 | 0.24 | 1.06 | 1.37 | 0.71 | 0.52 | 0.57 | 0.70 | 0.73 | 0.37 | 0.70 | **11.3** |
| 300 | 0.59 | 0.89 | 1.09 | 0.69 | 0.62 | 0.97 | 0.69 | 0.63 | 0.32 | 0.72 | **11.4** |
| 400 | 0.53 | 0.61 | 0.72 | 0.64 | 0.67 | 0.91 | 1.33 | 0.67 | 0.24 | 0.70 | **11.8** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9500 | 2.69 | 2.67 | 1.32 | 1.81 | 1.35 | 5.98 | 5.88 | 2.92 | 3.57 | 3.13 | **1.8** |
| 9600 | 2.69 | 2.68 | 1.34 | 1.81 | 1.36 | 6.21 | 5.88 | 2.86 | 3.66 | 3.17 | **1.8** |
| 9700 | 2.68 | 2.72 | 1.35 | 1.81 | 1.36 | 6.38 | 5.87 | 2.80 | 3.74 | 3.19 | **1.8** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 9800 | 2.71 | 2.74 | 1.33 | 1.80 | 1.38 | 6.42 | 5.85 | 2.79 | 3.88 | 3.21 | **1.8** |
| 9900 | 2.75 | 2.75 | 1.32 | 1.79 | 1.39 | 6.37 | 5.84 | 2.81 | 3.99 | 3.22 | **1.8** |
| 10000 | 2.75 | 2.78 | 1.33 | 1.77 | 1.39 | 6.35 | 5.83 | 2.84 | 4.06 | 3.23 | **1.8** |

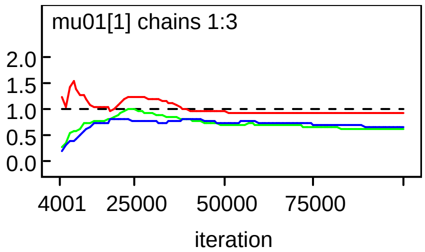
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**.




**History Graph: comparison of three different chains**

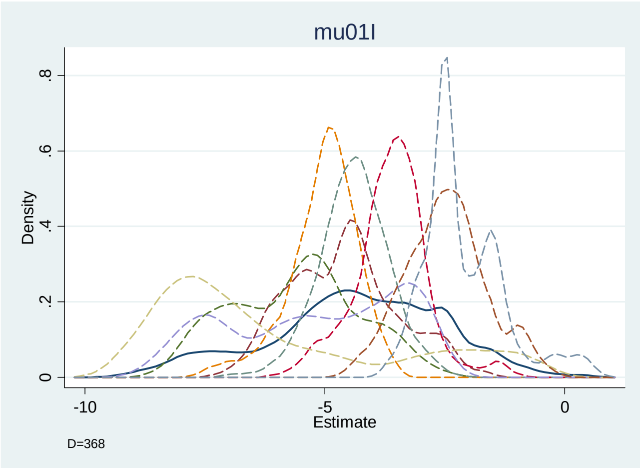



**Brooks-Gelman-Rubin plot**




<u>Internet</u>

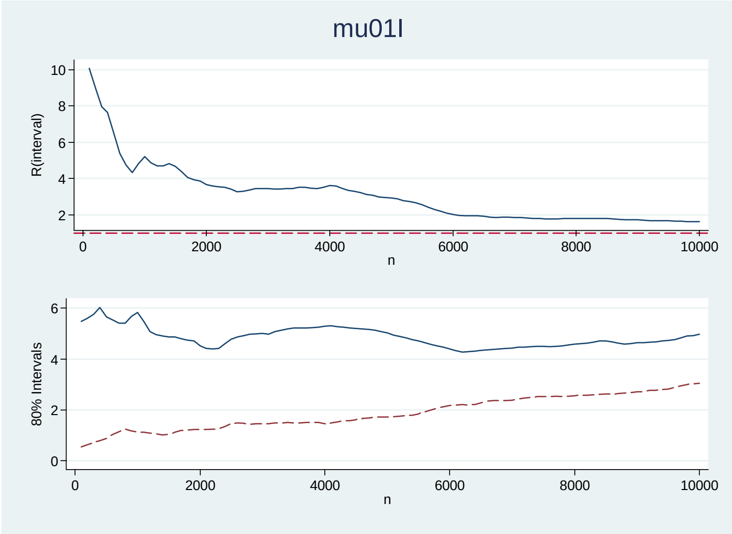
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



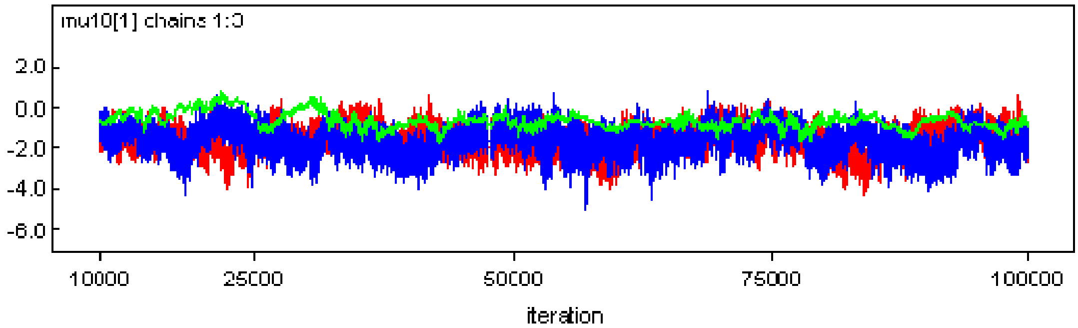

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.10 | 0.45 | 0.93 | 0.72 | 0.61 | 0.73 | 0.63 | 0.50 | 0.24 | 0.55 | **10.1** |
| 200 | 0.18 | 0.77 | 1.11 | 0.53 | 0.60 | 0.49 | 0.85 | 0.75 | 0.33 | 0.62 | **9.0** |
| 300 | 0.33 | 0.69 | 0.93 | 0.85 | 0.83 | 0.88 | 0.74 | 0.89 | 0.36 | 0.72 | **8.0** |
| 400 | 0.47 | 0.65 | 1.01 | 0.73 | 1.07 | 0.85 | 0.71 | 0.75 | 0.85 | 0.79 | **7.6** |
| 500 | 0.30 | 0.69 | 0.98 | 1.25 | 0.67 | 0.78 | 0.93 | 1.16 | 1.02 | 0.87 | **6.5** |
| 600 | 0.58 | 1.03 | 0.97 | 1.45 | 0.95 | 0.64 | 2.04 | 1.10 | 0.49 | 1.03 | **5.4** |
| 700 | 0.71 | 0.95 | 2.07 | 1.34 | 1.00 | 0.68 | 2.02 | 1.12 | 0.40 | 1.14 | **4.7** |
| 800 | 0.93 | 0.90 | 2.36 | 1.25 | 1.01 | 1.16 | 1.98 | 1.30 | 0.39 | 1.25 | **4.3** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 3.66 | 2.77 | 1.45 | 2.32 | 1.34 | 3.32 | 5.19 | 2.35 | 1.95 | 2.71 | **1.7** |
| 9100 | 3.69 | 2.80 | 1.44 | 2.30 | 1.45 | 3.31 | 5.30 | 2.30 | 1.94 | 2.72 | **1.7** |
| 9200 | 3.69 | 2.84 | 1.48 | 2.28 | 1.64 | 3.35 | 5.37 | 2.28 | 1.96 | 2.77 | **1.7** |
| 9300 | 3.63 | 2.89 | 1.51 | 2.23 | 1.70 | 3.32 | 5.41 | 2.26 | 2.00 | 2.77 | **1.7** |
| 9400 | 3.62 | 3.00 | 1.55 | 2.18 | 1.68 | 3.28 | 5.50 | 2.25 | 2.04 | 2.79 | **1.7** |
| 9500 | 3.60 | 3.16 | 1.60 | 2.15 | 1.65 | 3.27 | 5.62 | 2.24 | 2.14 | 2.83 | **1.7** |
| 9600 | 3.60 | 3.41 | 1.64 | 2.09 | 1.61 | 3.26 | 5.81 | 2.22 | 2.26 | 2.88 | **1.7** |
| 9700 | 3.63 | 3.72 | 1.69 | 2.07 | 1.58 | 3.24 | 5.97 | 2.21 | 2.38 | 2.94 | **1.6** |
| 9800 | 3.61 | 3.84 | 1.75 | 2.08 | 1.60 | 3.25 | 6.09 | 2.19 | 2.61 | 3.00 | **1.6** |
| 9900 | 3.61 | 3.82 | 1.74 | 2.22 | 1.56 | 3.25 | 6.07 | 2.10 | 2.86 | 3.03 | **1.6** |
| 10000 | 3.64 | 3.80 | 1.72 | 2.30 | 1.52 | 3.22 | 6.15 | 2.02 | 2.99 | 3.04 | **1.6** |

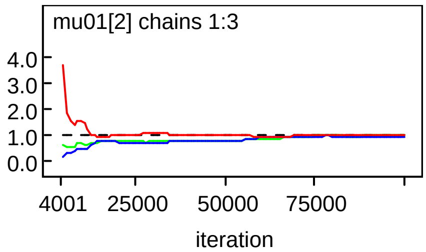
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**.




**History Graph: comparison of three different chains**

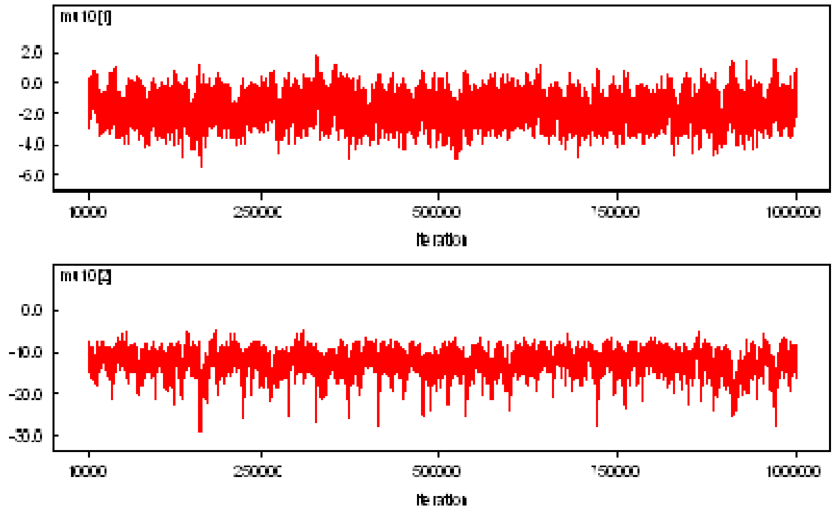


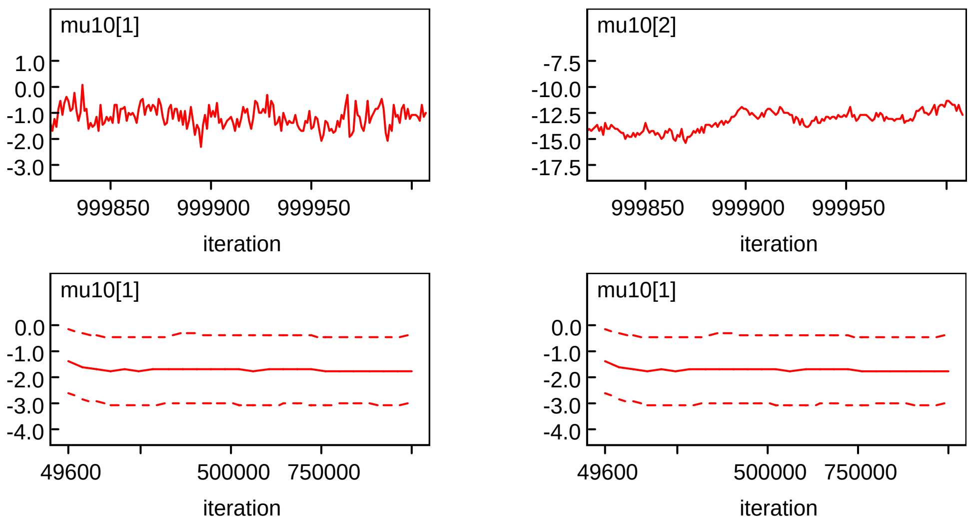

**Brooks-Gelman-Rubin plot**




## Catalogs Sent Trial

History, Trace and Quantiles and with 1 million iterations



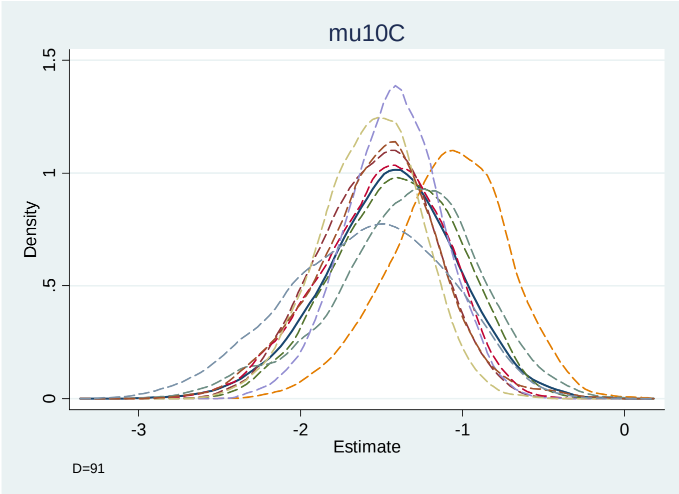

**Comparison of the Mean Results with 10.000 iterations and 100.000 iterations**

| node | mean | sd | MC error | 2.50% | median | 97.50% | start | sample |
|---|---|---|---|---|---|---|---|---|
| mu10[1] | -1.60 | 0.35 | 0.03 | -2.27 | -1.60 | -0.92 | 4000 | 10001 |
| mu10[1] | -1.59 | 0.50 | 0.04 | -2.61 | -1.57 | -0.72 | 90000 | 10001 |
| mu10[2] | -7.55 | 1.45 | 0.14 | -10.03 | -7.45 | -5.17 | 4000 | 10001 |
| mu10[2] | -6.59 | 0.95 | 0.09 | -8.27 | -6.62 | -4.56 | 90000 | 10001 |

**Formal Tests: COMPARISON OF DIFFERENT CHAINS**

Catalog

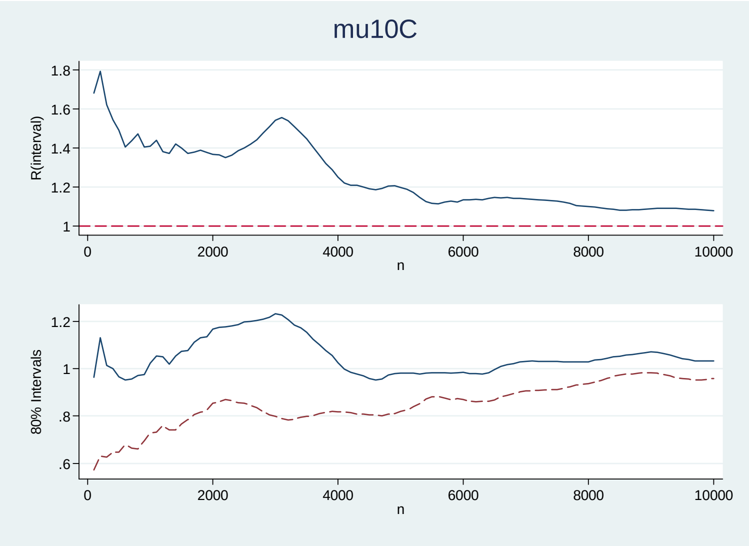
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



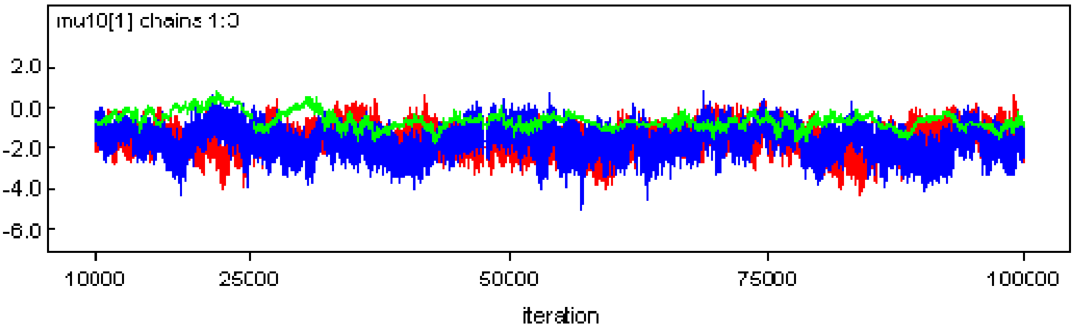

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.39 | 0.57 | 0.61 | 0.59 | 0.56 | 0.49 | 0.63 | 0.66 | 0.64 | 0.57 | **1.7** |
| 200 | 0.52 | 0.62 | 0.81 | 0.87 | 0.58 | 0.51 | 0.67 | 0.44 | 0.67 | 0.63 | **1.8** |
| 300 | 0.61 | 0.59 | 0.62 | 0.72 | 0.60 | 0.60 | 0.68 | 0.46 | 0.74 | 0.63 | **1.6** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9500 | 0.95 | 0.99 | 0.96 | 1.35 | 1.09 | 0.75 | 0.74 | 0.77 | 1.02 | 0.96 | **1.1** |
| 9600 | 0.94 | 1.01 | 0.97 | 1.32 | 1.09 | 0.75 | 0.74 | 0.77 | 1.01 | 0.96 | **1.1** |
| 9700 | 0.94 | 1.01 | 0.97 | 1.30 | 1.09 | 0.75 | 0.73 | 0.78 | 0.99 | 0.95 | **1.1** |
| 9800 | 0.94 | 1.02 | 0.97 | 1.31 | 1.09 | 0.75 | 0.73 | 0.79 | 0.98 | 0.95 | **1.1** |
| 9900 | 0.94 | 1.03 | 0.98 | 1.32 | 1.09 | 0.74 | 0.73 | 0.80 | 0.97 | 0.96 | **1.1** |
| 10000 | 0.94 | 1.04 | 0.98 | 1.31 | 1.08 | 0.75 | 0.73 | 0.83 | 0.96 | 0.96 | **1.1** |

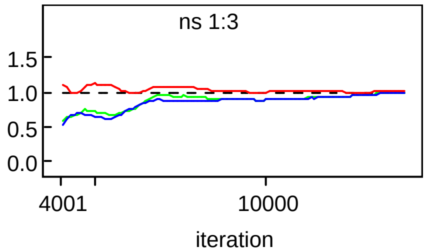
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**.




**History Graph: comparison of three different chains**

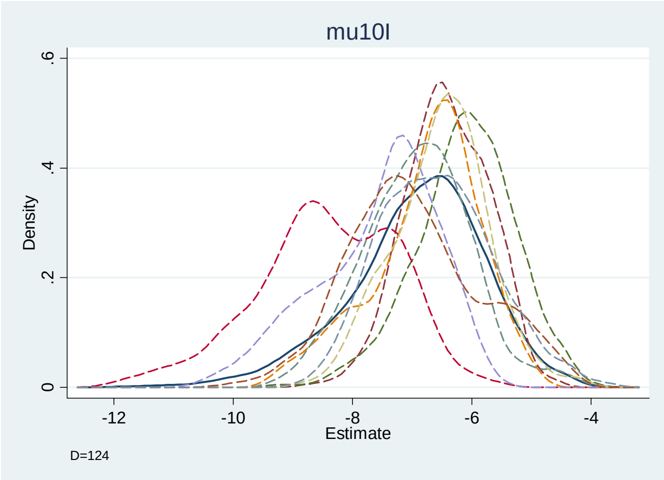



**Brooks-Gelman-Rubin plot**




Internet

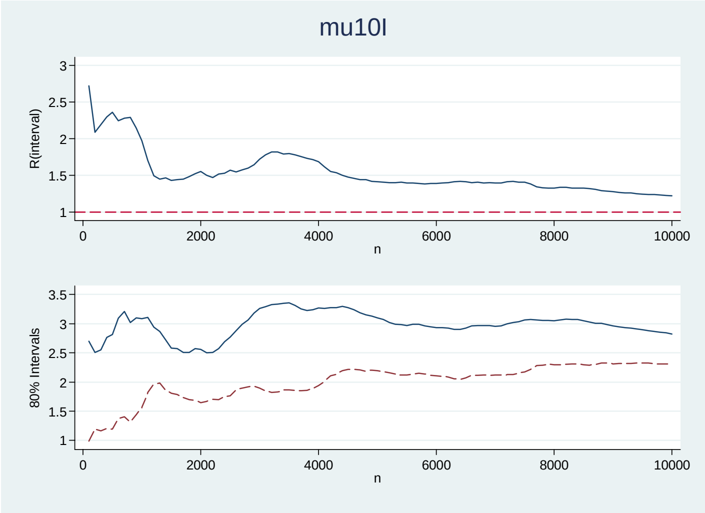
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



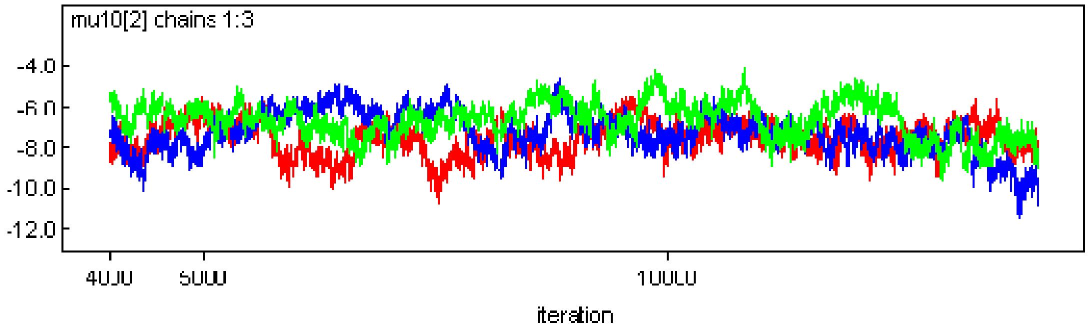

**B.G.R. statistic (R)**

| n | chain 1 | chain 2 | chain 3 | chain 4 | chain 5 | chain 6 | chain 7 | chain 8 | chain 9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.78 | 0.77 | 1.02 | 1.00 | 1.14 | 0.82 | 1.25 | 1.16 | 0.98 | 0.99 | **2.7** |
| 200 | 1.49 | 0.96 | 1.00 | 1.54 | 1.42 | 0.84 | 0.79 | 1.33 | 1.41 | 1.20 | **2.1** |
| 300 | 1.64 | 0.92 | 0.93 | 0.90 | 1.16 | 1.28 | 1.42 | 1.16 | 1.06 | 1.16 | **2.2** |
| 400 | 1.10 | 1.48 | 1.14 | 1.00 | 1.28 | 1.23 | 1.22 | 0.97 | 1.40 | 1.20 | **2.3** |
| 500 | 1.07 | 1.47 | 1.13 | 0.95 | 1.68 | 1.14 | 1.00 | 1.12 | 1.19 | 1.19 | **2.4** |
| 600 | 1.56 | 1.41 | 1.11 | 0.94 | 1.69 | 1.50 | 1.33 | 1.72 | 1.13 | 1.38 | **2.2** |
| 700 | 1.41 | 1.68 | 1.34 | 0.97 | 1.77 | 1.52 | 1.27 | 1.67 | 1.05 | 1.41 | **2.3** |
| 800 | 1.29 | 1.19 | 1.36 | 1.20 | 1.90 | 1.13 | 1.32 | 1.31 | 1.16 | 1.32 | **2.3** |
| 900 | 1.22 | 1.12 | 1.59 | 1.56 | 1.78 | 0.93 | 1.74 | 1.09 | 1.97 | 1.44 | **2.1** |
| 1000 | 1.08 | 1.40 | 1.53 | 1.53 | 1.56 | 0.95 | 1.94 | 1.20 | 2.88 | 1.56 | **2.0** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 1.70 | 1.80 | 3.11 | 2.30 | 3.26 | 1.48 | 1.80 | 3.16 | 2.22 | 2.31 | **1.3** |
| 9100 | 1.68 | 1.79 | 3.12 | 2.29 | 3.34 | 1.47 | 1.82 | 3.16 | 2.21 | 2.32 | **1.3** |
| 9200 | 1.66 | 1.78 | 3.11 | 2.29 | 3.43 | 1.48 | 1.82 | 3.12 | 2.21 | 2.32 | **1.3** |
| 9300 | 1.65 | 1.78 | 3.08 | 2.24 | 3.53 | 1.51 | 1.81 | 3.07 | 2.21 | 2.32 | **1.3** |
| 9400 | 1.65 | 1.78 | 3.08 | 2.19 | 3.67 | 1.50 | 1.80 | 3.01 | 2.25 | 2.32 | **1.3** |
| 9500 | 1.66 | 1.77 | 3.07 | 2.13 | 3.76 | 1.52 | 1.79 | 2.97 | 2.28 | 2.33 | **1.2** |
| 9600 | 1.67 | 1.77 | 3.06 | 2.08 | 3.77 | 1.54 | 1.81 | 2.96 | 2.27 | 2.32 | **1.2** |
| 9700 | 1.67 | 1.77 | 3.03 | 2.04 | 3.76 | 1.54 | 1.80 | 2.97 | 2.24 | 2.31 | **1.2** |
| 9800 | 1.66 | 1.79 | 3.02 | 2.03 | 3.75 | 1.53 | 1.80 | 2.97 | 2.23 | 2.31 | **1.2** |
| 9900 | 1.65 | 1.81 | 3.01 | 2.03 | 3.73 | 1.53 | 1.79 | 3.05 | 2.23 | 2.31 | **1.2** |
| 10000 | 1.62 | 1.84 | 3.00 | 2.02 | 3.71 | 1.53 | 1.80 | 3.07 | 2.22 | 2.31 | **1.2** |

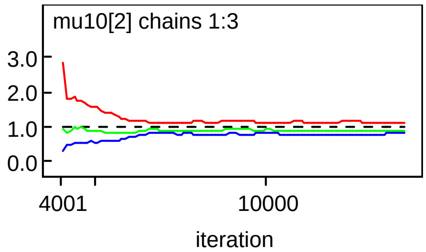
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**.




**History Graph: comparison of three different chains**

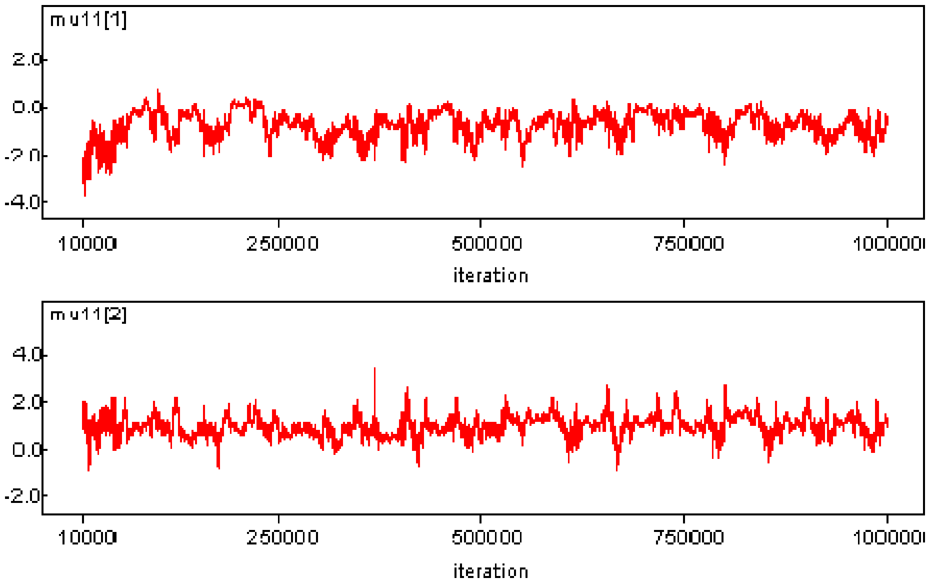


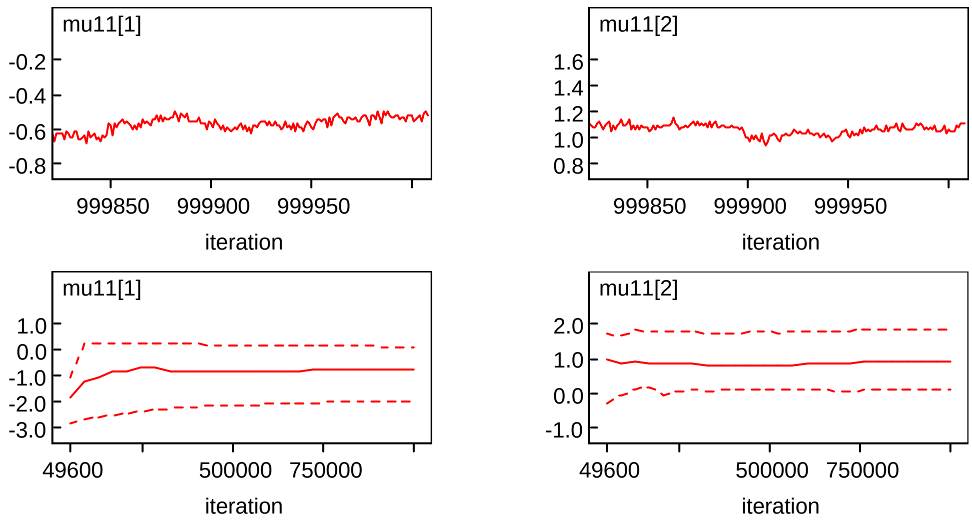

**Brooks-Gelman-Rubin plot**




## Catalogs Sent Post-trial

History, Trace and Quantiles and with 1 million iterations



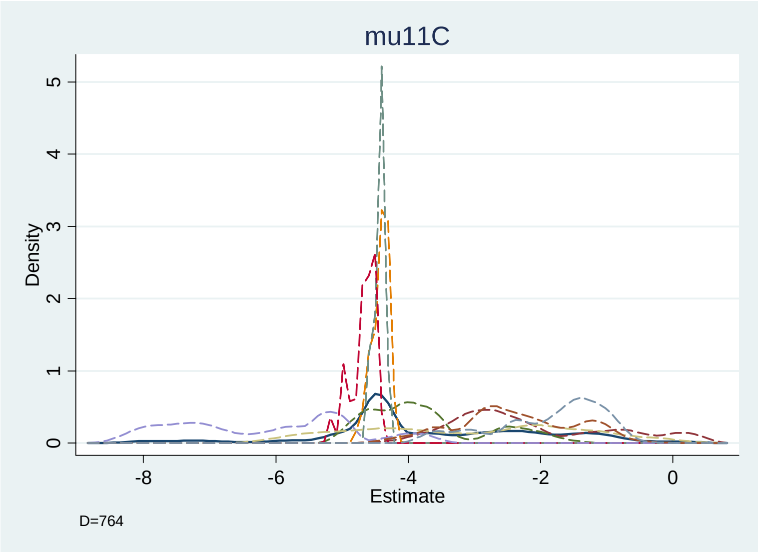

**Comparison of the Mean Results with 10.000 iterations and 100.000 iterations**

| node | mean | sd | MC error | 2.50% | median | 97.50% | start | sample |
|---|---|---|---|---|---|---|---|---|
| mu11[1] | -1.01 | 0.77 | 0.08 | -2.60 | -0.92 | 0.30 | 4000 | 10001 |
| mu11[1] | -1.88 | 0.84 | 0.08 | -3.69 | -1.64 | -0.73 | 90000 | 10001 |
| mu11[2] | 1.09 | 0.58 | 0.06 | -0.30 | 1.17 | 1.92 | 4000 | 10001 |
| mu11[2] | -0.14 | 0.50 | 0.05 | -1.29 | -0.07 | 0.70 | 90000 | 10001 |

**Tests: COMPARISON OF DIFFERENT CHAINS**

Catalog

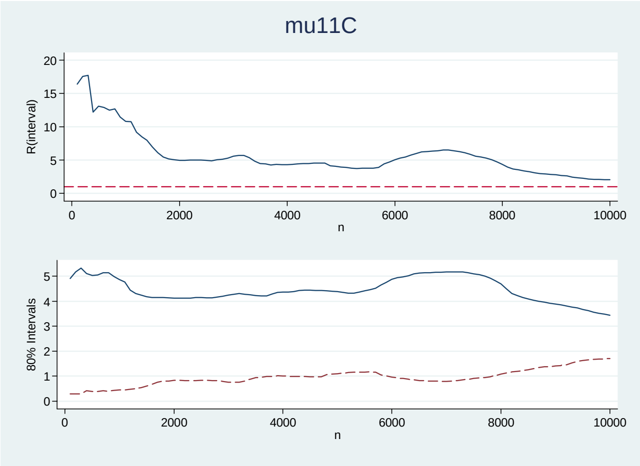
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



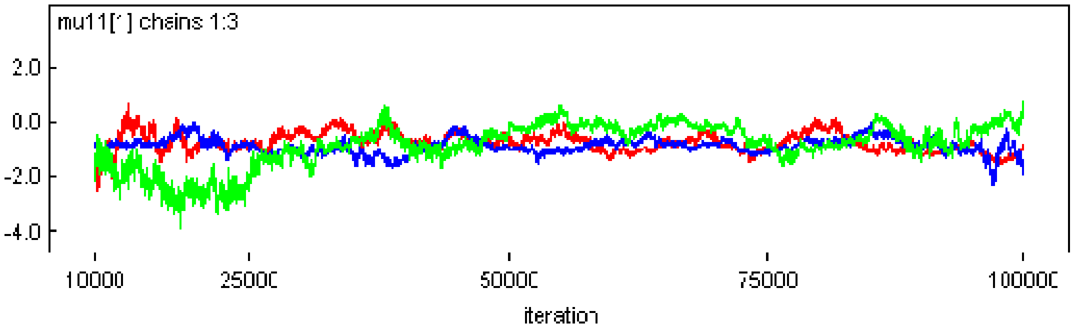

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.30 | 0.28 | 0.11 | 0.04 | 0.05 | 0.04 | 0.94 | 0.37 | 0.55 | 0.30 | **16.4** |
| 200 | 0.29 | 0.33 | 0.09 | 0.04 | 0.04 | 0.04 | 0.82 | 0.33 | 0.67 | 0.29 | **17.6** |
| 300 | 0.31 | 0.25 | 0.13 | 0.05 | 0.09 | 0.06 | 0.80 | 0.26 | 0.77 | 0.30 | **17.7** |
| … | …. | … | … | … | … | … | … | … | … | … | … |
| 9500 | 1.45 | 1.31 | 0.21 | 0.12 | 0.48 | 4.06 | 3.57 | 1.71 | 1.78 | 1.63 | **2.2** |
| 9600 | 1.55 | 1.31 | 0.21 | 0.12 | 0.48 | 4.15 | 3.56 | 1.72 | 1.86 | 1.66 | **2.2** |
| 9700 | 1.59 | 1.30 | 0.20 | 0.13 | 0.48 | 4.20 | 3.55 | 1.74 | 1.91 | 1.68 | **2.1** |
| 9800 | 1.63 | 1.27 | 0.20 | 0.13 | 0.49 | 4.20 | 3.55 | 1.77 | 1.90 | 1.68 | **2.1** |
| 9900 | 1.66 | 1.26 | 0.20 | 0.14 | 0.50 | 4.21 | 3.54 | 1.80 | 1.92 | 1.69 | **2.1** |
| 10000 | 1.67 | 1.26 | 0.20 | 0.14 | 0.51 | 4.21 | 3.53 | 1.84 | 1.94 | 1.70 | **2.0** |

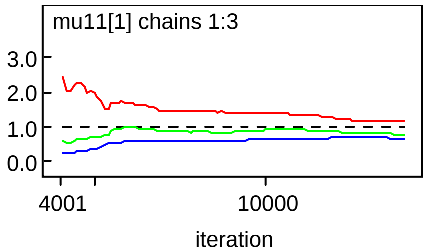
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

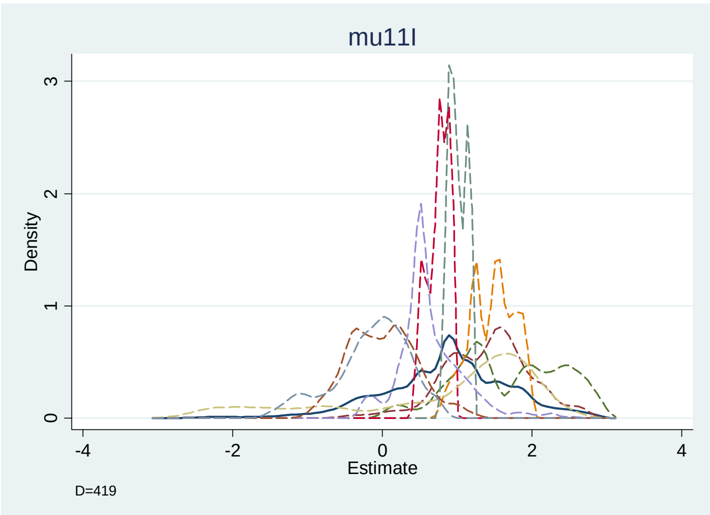



**Brooks-Gelman-Rubin plot**




Internet

**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**

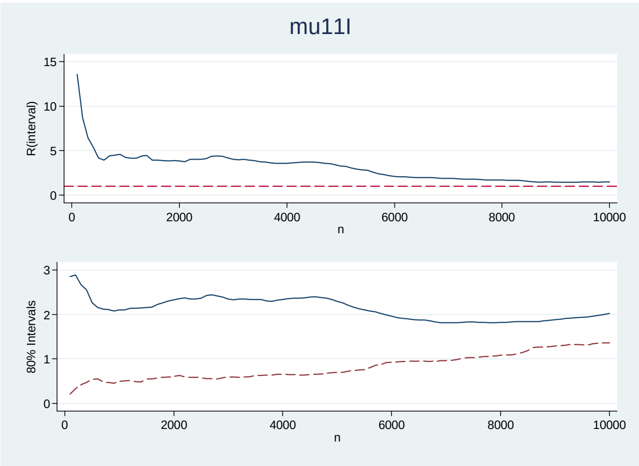



**B.G.R. statistic (R)**

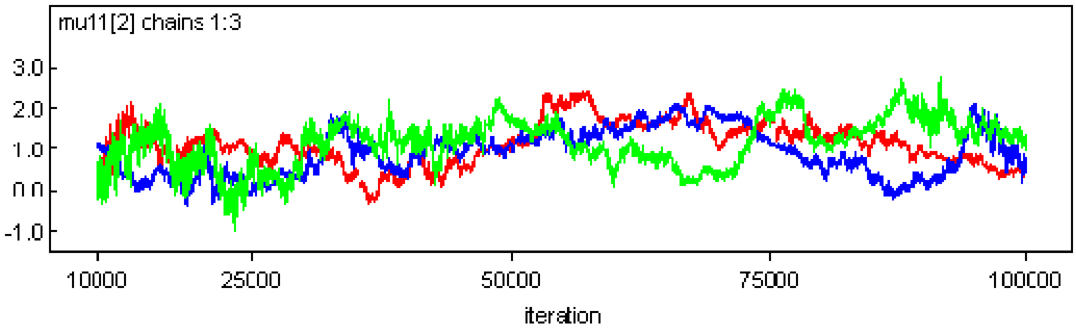
| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.22 | 0.19 | 0.10 | 0.03 | 0.04 | 0.07 | 0.34 | 0.28 | 0.61 | 0.21 | **13.6** |
| 200 | 0.43 | 0.31 | 0.11 | 0.03 | 0.06 | 0.06 | 1.25 | 0.35 | 0.38 | 0.33 | **8.7** |
| 300 | 0.39 | 0.37 | 0.15 | 0.06 | 0.12 | 0.08 | 1.05 | 0.70 | 0.84 | 0.42 | **6.4** |
| 400 | 0.40 | 0.51 | 0.24 | 0.08 | 0.06 | 0.15 | 1.43 | 0.58 | 0.78 | 0.47 | **5.4** |
| 500 | 0.79 | 0.67 | 0.19 | 0.09 | 0.05 | 0.11 | 1.57 | 0.35 | 1.03 | 0.54 | **4.2** |
| 600 | 0.94 | 0.73 | 0.28 | 0.06 | 0.06 | 0.10 | 1.78 | 0.34 | 0.63 | 0.55 | **3.9** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 700 | 0.82 | 0.40 | 0.30 | 0.06 | 0.08 | 0.15 | 1.74 | 0.25 | 0.55 | 0.48 | **4.4** |
| 800 | 0.85 | 0.34 | 0.27 | 0.07 | 0.07 | 0.14 | 1.65 | 0.33 | 0.51 | 0.47 | **4.5** |
| 900 | 0.87 | 0.32 | 0.19 | 0.06 | 0.10 | 0.14 | 1.48 | 0.42 | 0.50 | 0.45 | **4.6** |
| 1000 | 1.06 | 0.30 | 0.16 | 0.06 | 0.12 | 0.13 | 1.67 | 0.45 | 0.49 | 0.49 | **4.2** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 1.59 | 1.50 | 0.69 | 0.28 | 0.33 | 1.19 | 3.90 | 1.21 | 0.91 | 1.29 | **1.5** |
| 9100 | 1.65 | 1.47 | 0.69 | 0.28 | 0.33 | 1.20 | 3.94 | 1.22 | 0.89 | 1.30 | **1.5** |
| 9200 | 1.72 | 1.43 | 0.69 | 0.28 | 0.33 | 1.20 | 4.03 | 1.23 | 0.90 | 1.31 | **1.5** |
| 9300 | 1.83 | 1.40 | 0.69 | 0.28 | 0.33 | 1.22 | 4.10 | 1.22 | 0.92 | 1.33 | **1.4** |
| 9400 | 1.84 | 1.31 | 0.69 | 0.28 | 0.33 | 1.22 | 4.09 | 1.22 | 0.94 | 1.32 | **1.5** |
| 9500 | 1.85 | 1.16 | 0.69 | 0.28 | 0.33 | 1.21 | 4.08 | 1.27 | 0.98 | 1.32 | **1.5** |
| 9600 | 1.87 | 1.06 | 0.68 | 0.28 | 0.33 | 1.21 | 4.07 | 1.31 | 1.02 | 1.31 | **1.5** |
| 9700 | 1.86 | 1.01 | 0.67 | 0.28 | 0.33 | 1.21 | 4.06 | 1.31 | 1.31 | 1.34 | **1.5** |
| 9800 | 1.86 | 0.98 | 0.67 | 0.26 | 0.33 | 1.22 | 4.04 | 1.30 | 1.55 | 1.36 | **1.5** |
| 9900 | 1.85 | 0.95 | 0.66 | 0.25 | 0.33 | 1.21 | 4.02 | 1.30 | 1.64 | 1.36 | **1.5** |
| 10000 | 1.88 | 0.93 | 0.66 | 0.24 | 0.33 | 1.21 | 4.02 | 1.27 | 1.70 | 1.36 | **1.5** |

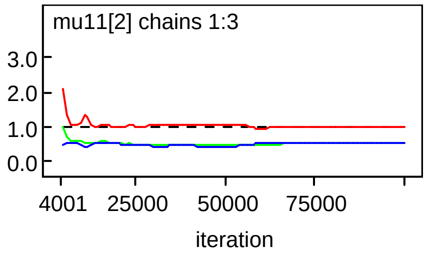
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

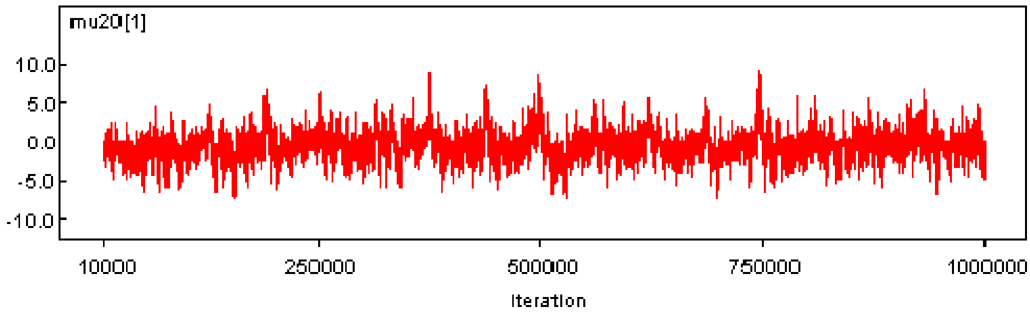


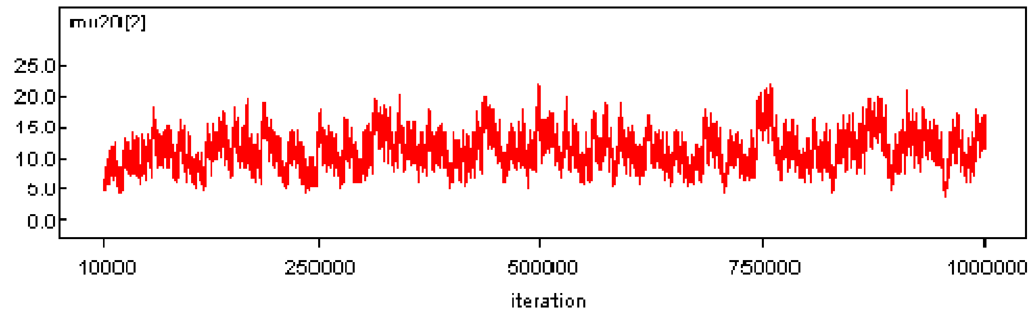

**Brooks-Gelman-Rubin plot**

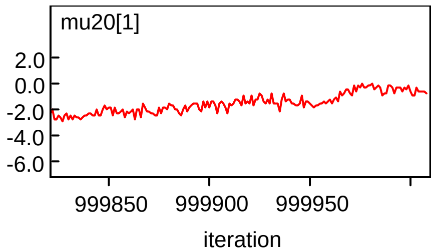


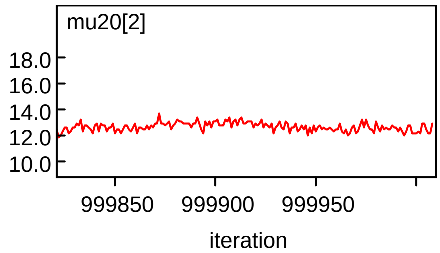

## E-mails Sent Trial

History, Trace and Quantiles and with 1 million iterations



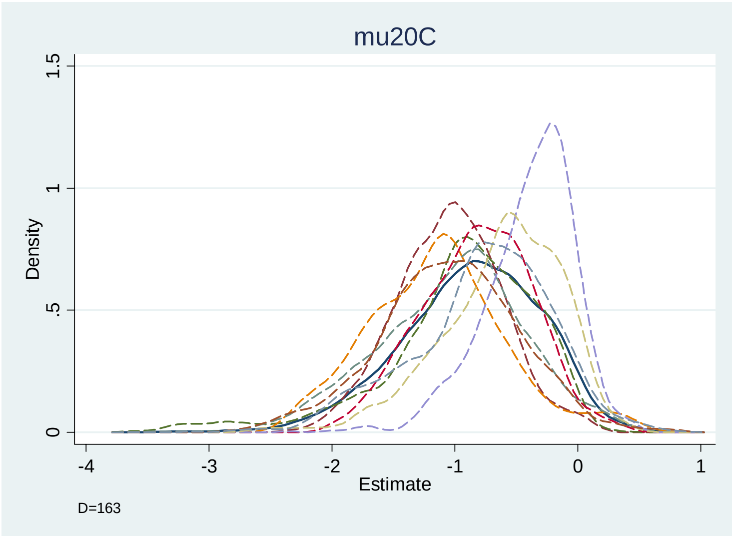

**Comparison of the Mean Results with 10.000 iterations and 100.000 iterations**

| node | mean | sd | MC error | 2.50% | median | 97.50% | start | sample |
|---|---|---|---|---|---|---|---|---|
| mu20[1] | -0.71 | 0.56 | 0.05 | -1.71 | -0.77 | 0.46 | 4000 | 10001 |
| mu20[1] | -0.78 | 0.58 | 0.05 | -2.02 | -0.72 | 0.20 | 90000 | 10001 |
| mu20[2] | 3.74 | 0.67 | 0.06 | 2.16 | 3.78 | 4.92 | 4000 | 10001 |
| mu20[2] | 3.94 | 1.09 | 0.11 | 2.33 | 3.71 | 6.79 | 90000 | 10001 |

**Formal Tests: COMPARISON OF DIFFERENT CHAINS**

Catalog

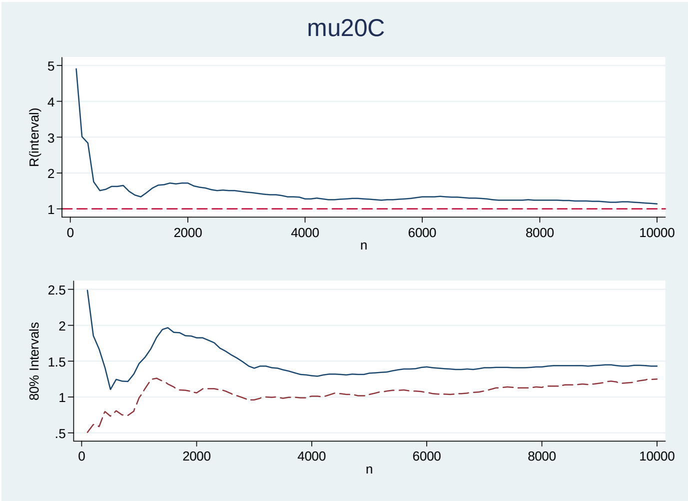
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



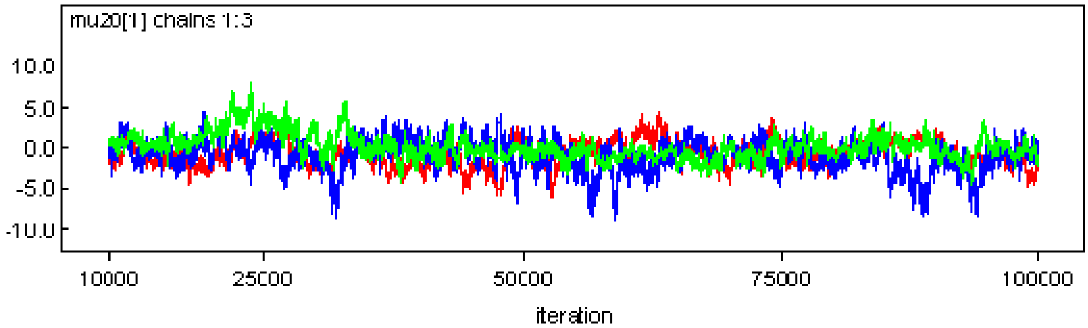

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.31 | 0.34 | 0.79 | 1.22 | 0.30 | 0.38 | 0.14 | 0.63 | 0.45 | 0.51 | **4.9** |
| 200 | 0.79 | 1.14 | 0.58 | 0.53 | 0.45 | 0.29 | 0.13 | 0.47 | 1.16 | 0.62 | **3.0** |
| 300 | 0.67 | 0.42 | 0.61 | 0.44 | 0.49 | 0.63 | 0.18 | 1.22 | 0.63 | 0.59 | **2.8** |
| 400 | 0.47 | 0.66 | 0.77 | 1.07 | 0.58 | 0.72 | 0.24 | 1.48 | 1.18 | 0.80 | **1.8** |
| 500 | 0.44 | 0.70 | 0.73 | 1.13 | 0.83 | 0.74 | 0.50 | 0.80 | 0.70 | 0.73 | **1.5** |
| … | …. | … | … | … | … | … | … | … | … | … | … |
| 9400 | 1.09 | 0.93 | 1.49 | 1.19 | 1.22 | 0.69 | 1.06 | 1.39 | 1.68 | 1.19 | **1.2** |
| 9500 | 1.09 | 0.94 | 1.50 | 1.15 | 1.22 | 0.68 | 1.05 | 1.42 | 1.70 | 1.20 | **1.2** |
| 9600 | 1.11 | 0.94 | 1.58 | 1.15 | 1.21 | 0.71 | 1.05 | 1.42 | 1.73 | 1.21 | **1.2** |
| 9700 | 1.17 | 0.93 | 1.68 | 1.14 | 1.21 | 0.74 | 1.04 | 1.41 | 1.72 | 1.23 | **1.2** |
| 9800 | 1.20 | 0.93 | 1.68 | 1.13 | 1.20 | 0.76 | 1.07 | 1.48 | 1.71 | 1.24 | **1.2** |
| 9900 | 1.19 | 0.95 | 1.69 | 1.12 | 1.20 | 0.76 | 1.10 | 1.47 | 1.71 | 1.24 | **1.2** |
| 10000 | 1.21 | 0.94 | 1.67 | 1.11 | 1.22 | 0.75 | 1.15 | 1.47 | 1.72 | 1.25 | **1.1** |

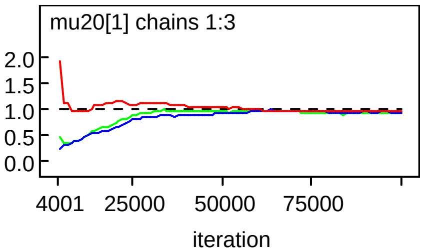
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

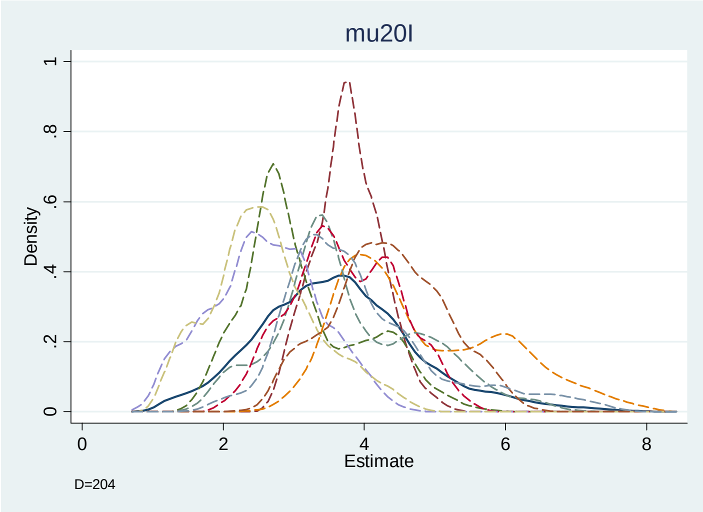



**Brooks-Gelman-Rubin plot**




Internet

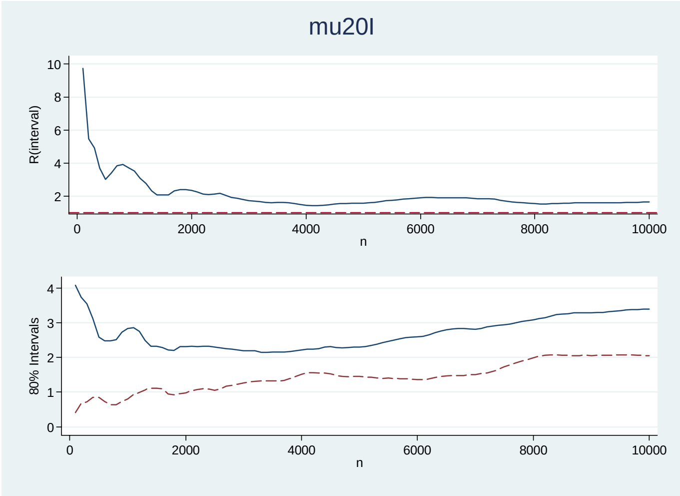
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



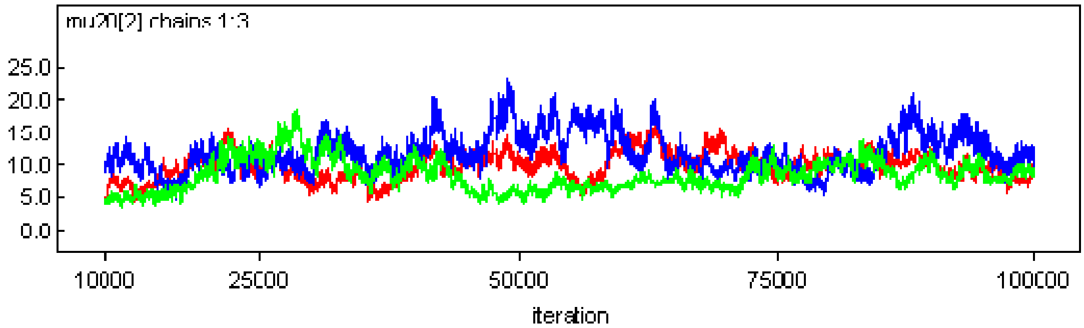

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.33 | 0.26 | 0.38 | 0.71 | 0.31 | 0.62 | 0.20 | 0.44 | 0.53 | 0.42 | **9.7** |
| 200 | 0.81 | 0.63 | 0.53 | 0.72 | 0.35 | 0.44 | 0.26 | 1.74 | 0.68 | 0.68 | **5.5** |
| 300 | 0.65 | 0.44 | 0.61 | 0.92 | 0.31 | 0.60 | 0.27 | 1.95 | 0.74 | 0.72 | **4.9** |
| 400 | 1.08 | 0.77 | 1.06 | 0.91 | 0.29 | 0.77 | 0.31 | 1.18 | 1.27 | 0.85 | **3.7** |
| 500 | 1.17 | 0.62 | 0.98 | 1.04 | 0.51 | 0.75 | 0.55 | 1.03 | 1.01 | 0.85 | **3.0** |
| 600 | 1.20 | 0.66 | 0.50 | 0.73 | 0.59 | 0.90 | 0.43 | 0.68 | 0.87 | 0.73 | **3.4** |
| 700 | 0.90 | 0.62 | 0.47 | 0.80 | 0.51 | 0.57 | 0.49 | 0.65 | 0.78 | 0.64 | **3.9** |
| 800 | 0.63 | 0.57 | 0.49 | 0.76 | 0.47 | 0.54 | 0.65 | 0.76 | 0.89 | 0.64 | **3.9** |
| 900 | 0.83 | 0.75 | 0.54 | 0.71 | 0.66 | 0.55 | 0.54 | 1.05 | 0.96 | 0.73 | **3.7** |
| 1000 | 0.93 | 0.84 | 0.53 | 0.98 | 0.74 | 0.63 | 0.40 | 1.18 | 0.99 | 0.80 | **3.5** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 1.06 | 1.05 | 2.94 | 3.55 | 1.90 | 1.09 | 1.11 | 2.34 | 3.44 | 2.05 | **1.6** |
| 9100 | 1.06 | 1.07 | 2.93 | 3.54 | 1.91 | 1.10 | 1.15 | 2.34 | 3.44 | 2.06 | **1.6** |
| 9200 | 1.05 | 1.07 | 2.91 | 3.53 | 1.95 | 1.10 | 1.18 | 2.33 | 3.43 | 2.06 | **1.6** |
| 9300 | 1.05 | 1.08 | 2.91 | 3.52 | 1.96 | 1.09 | 1.23 | 2.32 | 3.40 | 2.06 | **1.6** |
| 9400 | 1.06 | 1.15 | 2.90 | 3.51 | 1.95 | 1.11 | 1.28 | 2.31 | 3.36 | 2.07 | **1.6** |
| 9500 | 1.05 | 1.18 | 2.84 | 3.49 | 1.95 | 1.14 | 1.37 | 2.31 | 3.32 | 2.07 | **1.6** |
| 9600 | 1.03 | 1.22 | 2.79 | 3.48 | 1.95 | 1.19 | 1.42 | 2.30 | 3.30 | 2.07 | **1.6** |
| 9700 | 0.99 | 1.23 | 2.77 | 3.41 | 1.94 | 1.28 | 1.47 | 2.29 | 3.29 | 2.07 | **1.6** |
| 9800 | 0.97 | 1.24 | 2.74 | 3.30 | 1.94 | 1.32 | 1.52 | 2.28 | 3.28 | 2.07 | **1.6** |
| 9900 | 0.95 | 1.29 | 2.70 | 3.20 | 1.94 | 1.32 | 1.58 | 2.28 | 3.28 | 2.06 | **1.6** |
| 10000 | 0.98 | 1.35 | 2.65 | 3.10 | 1.93 | 1.36 | 1.57 | 2.27 | 3.26 | 2.05 | **1.7** |

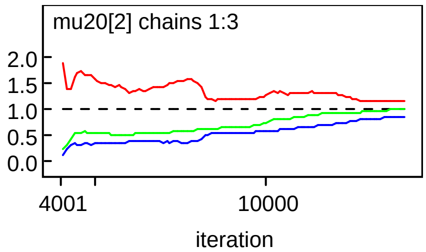
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

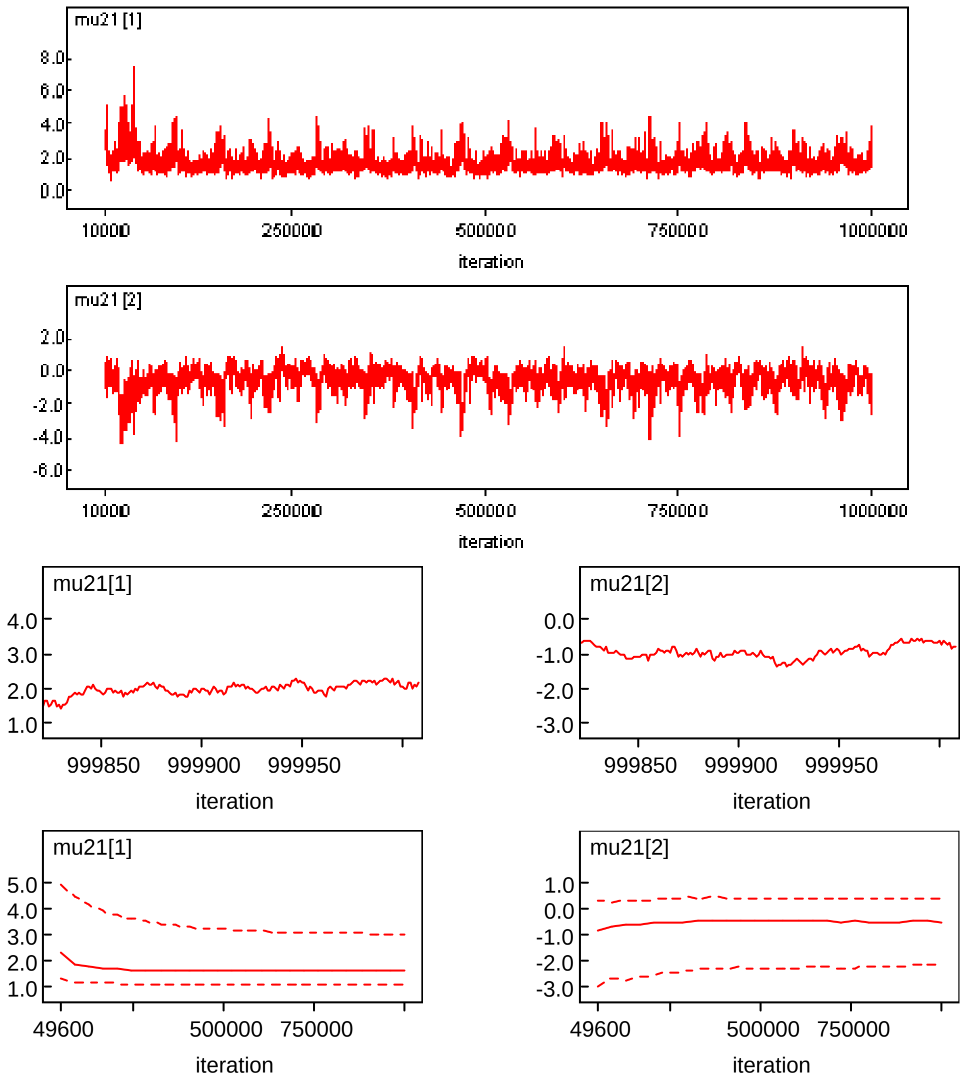



**Brooks-Gelman-Rubin plot**




## E-mails Sent Post-trial

History, Trace and Quantiles and with 1 million iterations



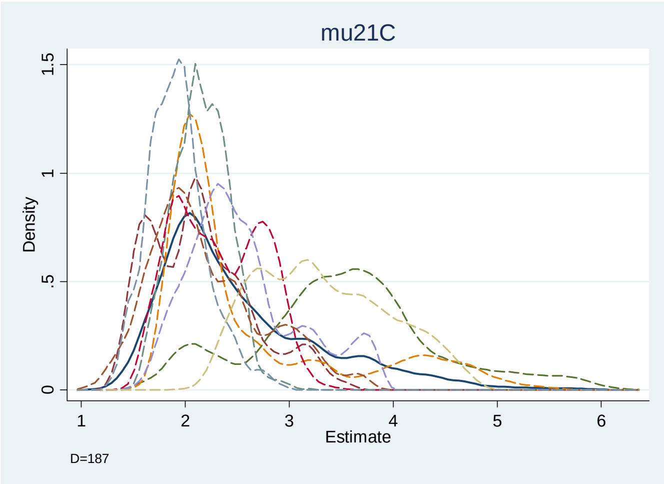

**Comparison of the Mean Results with 10.000 iterations and 100.000 iterations**

| node | mean | sd | MC error | 2.50% | median | 97.50% | start | sample |
|---|---|---|---|---|---|---|---|---|
| mu21[1] | 1.93 | 0.29 | 0.03 | 1.43 | 1.91 | 2.60 | 4000 | 10001 |
| mu21[1] | 1.88 | 0.33 | 0.03 | 1.36 | 1.86 | 2.57 | 90000 | 10001 |
| mu21[2] | -0.66 | 0.36 | 0.04 | -1.29 | -0.69 | 0.02 | 4000 | 10001 |
| mu21[2] | -0.90 | 0.65 | 0.06 | -1.69 | -1.14 | 0.68 | 90000 | 10001 |

**Formal Tests: COMPARISON OF DIFFERENT CHAINS**

Catalog

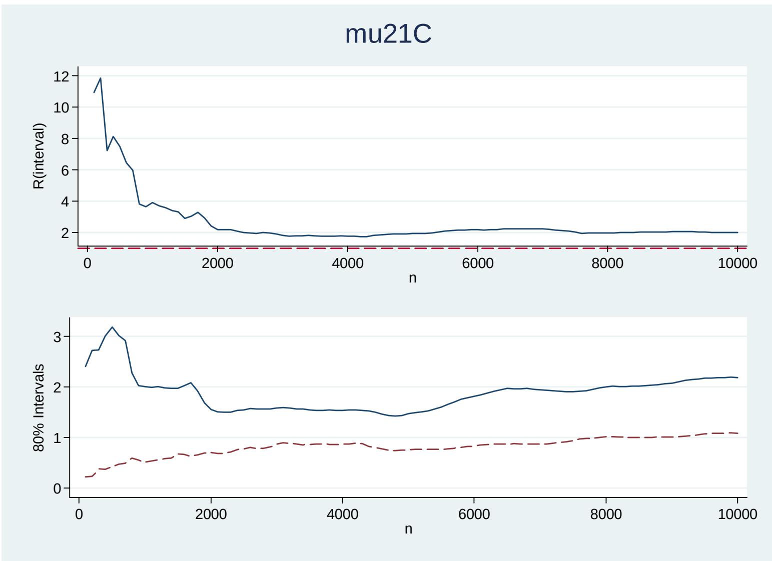
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



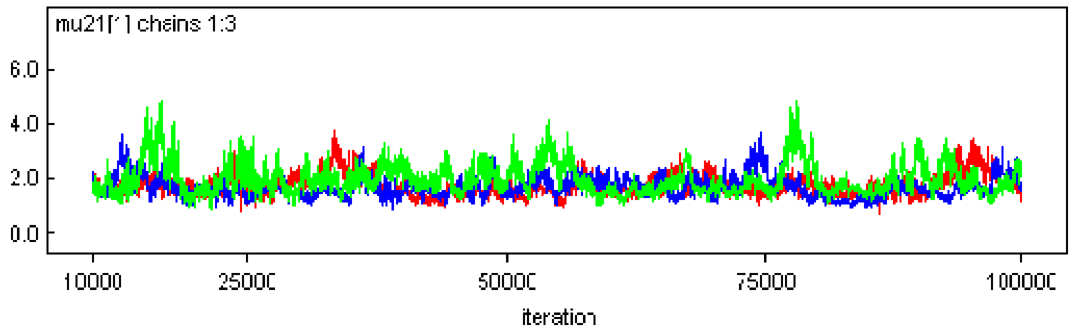

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.16 | 0.16 | 0.34 | 0.13 | 0.21 | 0.29 | 0.35 | 0.23 | 0.12 | 0.22 | **10.9** |
| 200 | 0.11 | 0.22 | 0.42 | 0.18 | 0.27 | 0.22 | 0.31 | 0.16 | 0.18 | 0.23 | **11.9** |
| 300 | 0.18 | 0.23 | 0.72 | 0.31 | 0.35 | 0.35 | 0.30 | 0.69 | 0.27 | 0.38 | **7.2** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9500 | 1.12 | 1.88 | 0.37 | 0.81 | 1.10 | 1.45 | 1.44 | 0.82 | 0.65 | 1.07 | **2.0** |
| 9600 | 1.12 | 1.98 | 0.37 | 0.81 | 1.10 | 1.45 | 1.43 | 0.85 | 0.68 | 1.09 | **2.0** |
| 9700 | 1.12 | 1.94 | 0.37 | 0.81 | 1.10 | 1.45 | 1.43 | 0.87 | 0.69 | 1.09 | **2.0** |
| 9800 | 1.13 | 1.92 | 0.37 | 0.81 | 1.11 | 1.44 | 1.42 | 0.91 | 0.68 | 1.09 | **2.0** |
| 9900 | 1.15 | 1.91 | 0.37 | 0.81 | 1.11 | 1.44 | 1.42 | 0.94 | 0.67 | 1.09 | **2.0** |
| 10000 | 1.18 | 1.89 | 0.37 | 0.80 | 1.09 | 1.44 | 1.41 | 0.94 | 0.65 | 1.09 | **2.0** |

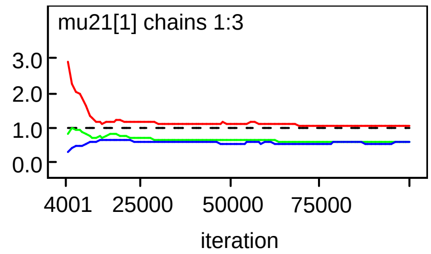
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**



**History Graph: comparison of three different chains**

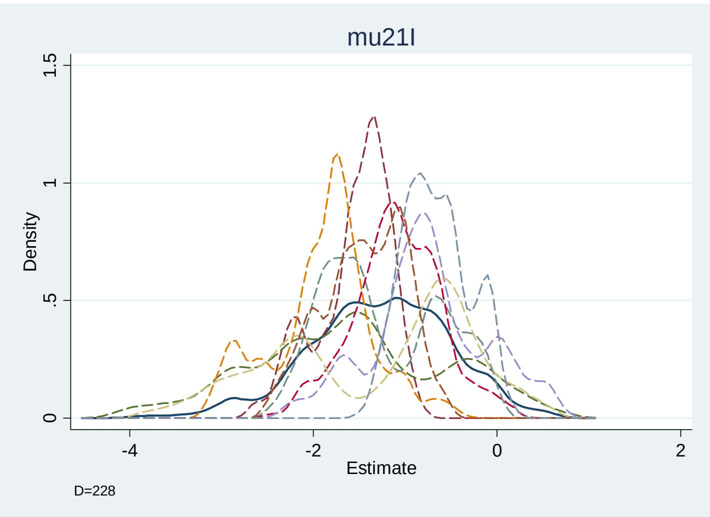



**Brooks-Gelman-Rubin plot**




<u>Internet</u>

**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**

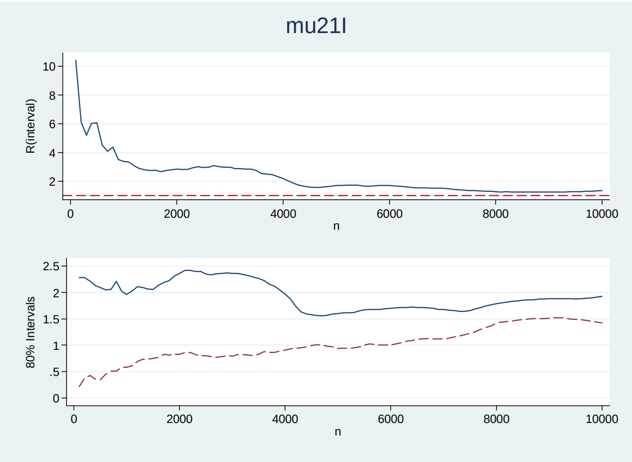



**B.G.R. statistic (R)**

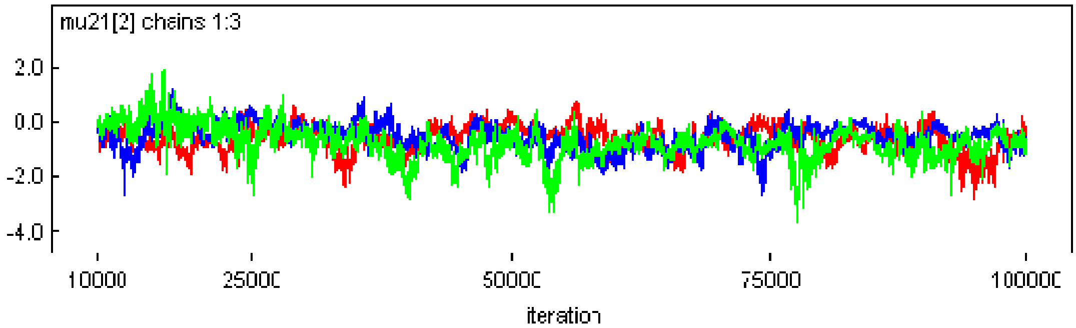
| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.11 | 0.14 | 0.46 | 0.14 | 0.14 | 0.24 | 0.31 | 0.26 | 0.18 | 0.22 | **10.4** |
| 200 | 0.13 | 0.16 | 1.34 | 0.19 | 0.21 | 0.36 | 0.18 | 0.55 | 0.26 | 0.37 | **6.1** |
| 300 | 0.28 | 0.12 | 0.53 | 0.28 | 0.26 | 0.79 | 0.32 | 0.97 | 0.28 | 0.43 | **5.2** |
| 400 | 0.32 | 0.22 | 0.50 | 0.25 | 0.30 | 0.39 | 0.30 | 0.62 | 0.27 | 0.35 | **6.0** |
| 500 | 0.31 | 0.23 | 0.43 | 0.17 | 0.21 | 0.50 | 0.52 | 0.30 | 0.43 | 0.34 | **6.1** |
| 600 | 0.32 | 0.20 | 0.74 | 0.36 | 0.26 | 0.75 | 0.47 | 0.54 | 0.43 | 0.45 | **4.5** |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 700 | 0.25 | 0.18 | 0.91 | 0.43 | 0.45 | 0.82 | 0.49 | 0.67 | 0.35 | 0.51 | **4.1** |
| 800 | 0.20 | 0.31 | 0.99 | 0.41 | 0.50 | 0.67 | 0.36 | 0.79 | 0.31 | 0.51 | **4.4** |
| 900 | 0.35 | 0.43 | 1.10 | 0.35 | 0.56 | 0.62 | 0.59 | 0.69 | 0.50 | 0.58 | **3.5** |
| 1000 | 0.36 | 0.48 | 0.94 | 0.31 | 0.66 | 0.58 | 0.75 | 0.61 | 0.55 | 0.58 | **3.4** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 1.06 | 2.15 | 0.84 | 1.80 | 1.21 | 2.05 | 2.43 | 1.27 | 0.82 | 1.51 | **1.2** |
| 9100 | 1.03 | 2.24 | 0.83 | 1.79 | 1.21 | 2.11 | 2.42 | 1.28 | 0.79 | 1.52 | **1.2** |
| 9200 | 1.02 | 2.22 | 0.83 | 1.79 | 1.20 | 2.17 | 2.41 | 1.27 | 0.77 | 1.52 | **1.2** |
| 9300 | 1.00 | 2.21 | 0.84 | 1.78 | 1.20 | 2.16 | 2.40 | 1.27 | 0.76 | 1.51 | **1.2** |
| 9400 | 1.00 | 2.20 | 0.83 | 1.78 | 1.19 | 2.13 | 2.33 | 1.27 | 0.75 | 1.50 | **1.3** |
| 9500 | 1.00 | 2.22 | 0.84 | 1.76 | 1.18 | 2.08 | 2.34 | 1.26 | 0.73 | 1.49 | **1.3** |
| 9600 | 1.00 | 2.28 | 0.84 | 1.72 | 1.17 | 2.03 | 2.33 | 1.26 | 0.71 | 1.48 | **1.3** |
| 9700 | 1.00 | 2.36 | 0.83 | 1.69 | 1.15 | 1.98 | 2.28 | 1.25 | 0.70 | 1.47 | **1.3** |
| 9800 | 1.01 | 2.44 | 0.84 | 1.67 | 1.15 | 1.89 | 2.19 | 1.25 | 0.69 | 1.46 | **1.3** |
| 9900 | 1.03 | 2.51 | 0.84 | 1.65 | 1.14 | 1.73 | 2.14 | 1.25 | 0.68 | 1.44 | **1.3** |
| 10000 | 1.05 | 2.51 | 0.85 | 1.61 | 1.14 | 1.65 | 2.08 | 1.24 | 0.67 | 1.42 | **1.4** |

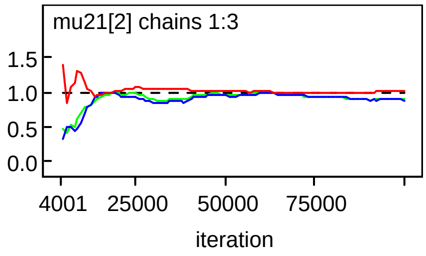
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

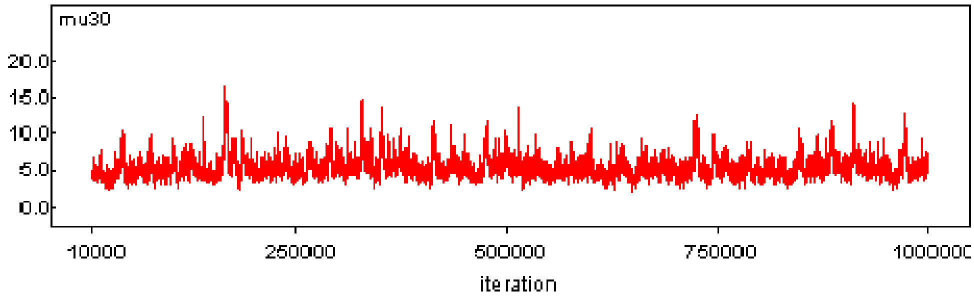


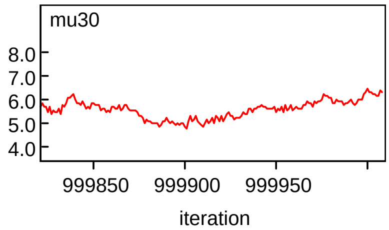

**Brooks-Gelman-Rubin plot**

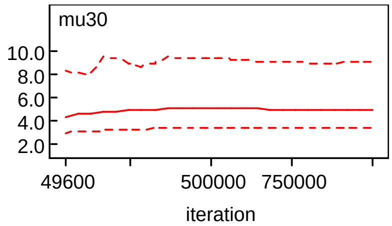



## State Dependence Trial

Graphs: History, Trace, Quantiles and Density Graphs with 1 million iterations



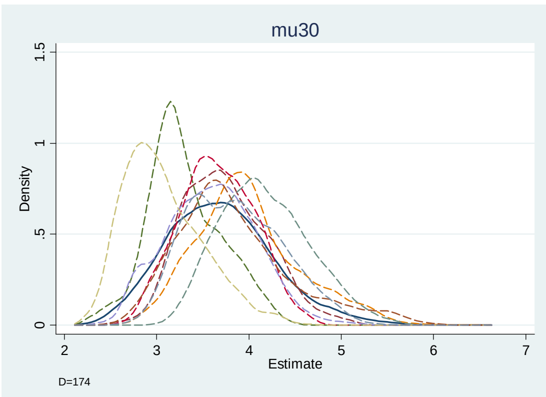




**Comparison of the Mean Results with 10.000 iterations and 100.000 iterations**

| node | mean | sd | MC error | 2.50% | median | 97.50% | start | sample |
|---|---|---|---|---|---|---|---|---|
| mu30 | 3.75 | 0.55 | 0.05 | 2.86 | 3.70 | 4.87 | 4000 | 10001 |
| mu30 | 3.85 | 0.51 | 0.05 | 3.00 | 3.81 | 4.90 | 90000 | 10001 |

**Tests: COMPARISON OF DIFFERENT CHAINS**

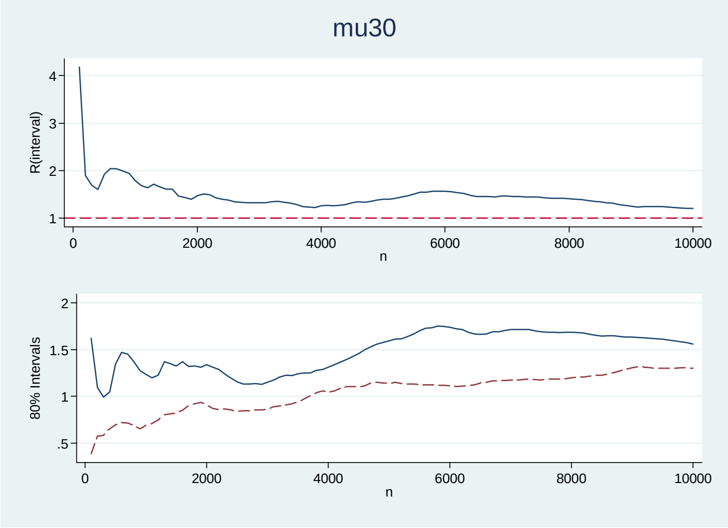
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line).**



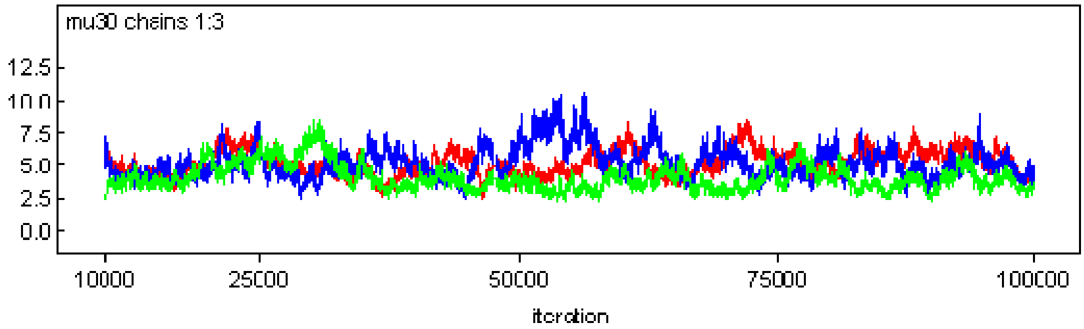
**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.28 | 0.35 | 0.66 | 0.36 | 0.37 | 0.37 | 0.33 | 0.40 | 0.38 | 0.39 | **4.2** |
| 200 | 0.55 | 0.72 | 0.82 | 0.43 | 0.41 | 0.94 | 0.34 | 0.50 | 0.48 | 0.58 | **1.9** |
| 300 | 0.56 | 0.37 | 0.58 | 0.56 | 0.48 | 1.15 | 0.29 | 0.63 | 0.64 | 0.58 | **1.7** |
| 400 | 0.83 | 0.43 | 0.79 | 1.31 | 0.51 | 0.68 | 0.45 | 0.44 | 0.47 | 0.65 | **1.6** |
| 500 | 0.80 | 0.42 | 0.90 | 1.02 | 0.59 | 0.55 | 0.69 | 0.48 | 0.84 | 0.70 | **1.9** |
| 600 | 0.89 | 0.57 | 0.66 | 0.70 | 0.61 | 0.78 | 0.68 | 0.70 | 0.90 | 0.72 | **2.0** |
| 700 | 0.84 | 0.57 | 0.87 | 0.65 | 0.64 | 0.79 | 0.41 | 0.83 | 0.83 | 0.71 | **2.0** |
| 800 | 0.58 | 0.51 | 0.83 | 0.77 | 0.71 | 0.69 | 0.46 | 0.87 | 0.78 | 0.69 | **2.0** |
| 900 | 0.66 | 0.43 | 0.69 | 0.76 | 0.65 | 0.58 | 0.47 | 0.75 | 0.93 | 0.66 | **1.9** |
| 1000 | 0.87 | 0.48 | 0.69 | 0.93 | 0.61 | 0.50 | 0.51 | 0.73 | 0.93 | 0.69 | **1.8** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 1.25 | 0.88 | 1.50 | 1.29 | 1.15 | 1.19 | 1.07 | 2.03 | 1.41 | 1.31 | **1.3** |
| 9100 | 1.26 | 0.93 | 1.52 | 1.29 | 1.15 | 1.23 | 1.09 | 2.00 | 1.40 | 1.32 | **1.2** |
| 9200 | 1.26 | 0.97 | 1.51 | 1.25 | 1.16 | 1.26 | 1.08 | 1.92 | 1.40 | 1.31 | **1.2** |
| 9300 | 1.25 | 1.02 | 1.51 | 1.23 | 1.16 | 1.28 | 1.08 | 1.85 | 1.40 | 1.31 | **1.2** |
| 9400 | 1.25 | 1.03 | 1.48 | 1.22 | 1.18 | 1.27 | 1.09 | 1.79 | 1.38 | 1.30 | **1.2** |
| 9500 | 1.24 | 1.05 | 1.50 | 1.21 | 1.21 | 1.27 | 1.11 | 1.75 | 1.36 | 1.30 | **1.2** |
| 9600 | 1.24 | 1.07 | 1.49 | 1.23 | 1.24 | 1.26 | 1.15 | 1.71 | 1.35 | 1.30 | **1.2** |
| 9700 | 1.24 | 1.06 | 1.49 | 1.25 | 1.25 | 1.25 | 1.16 | 1.71 | 1.33 | 1.30 | **1.2** |
| 9800 | 1.23 | 1.05 | 1.48 | 1.26 | 1.25 | 1.25 | 1.21 | 1.71 | 1.32 | 1.31 | **1.2** |
| 9900 | 1.23 | 1.06 | 1.48 | 1.27 | 1.25 | 1.25 | 1.20 | 1.69 | 1.33 | 1.31 | **1.2** |
| 10000 | 1.23 | 1.08 | 1.48 | 1.26 | 1.25 | 1.25 | 1.20 | 1.63 | 1.34 | 1.30 | **1.2** |

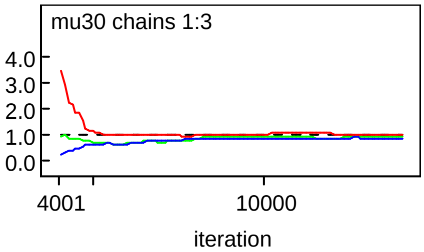
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

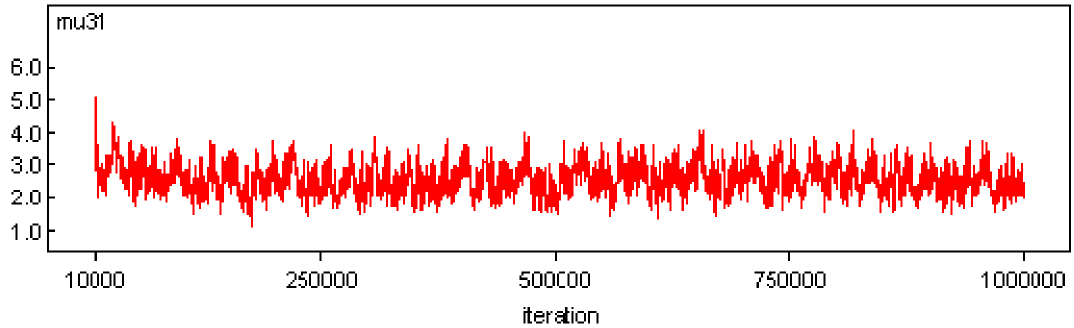


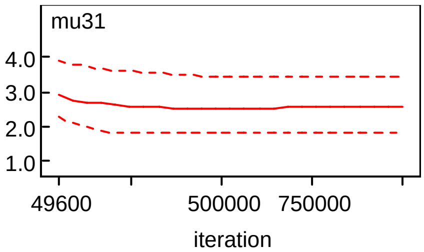

**Brooks-Gelman-Rubin plot**

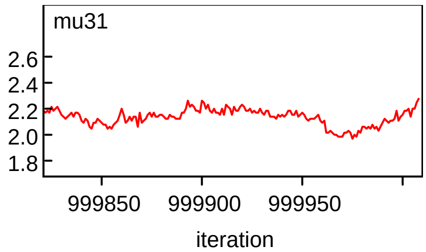



## State Dependence Post-trial

<u>Graphs:</u> History, Trace, Quantiles and Density Graphs with 1 million iterations



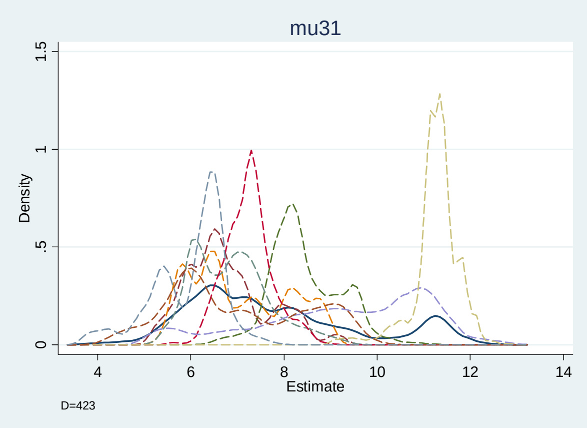




**Comparison of the Mean Results with 10.000 iterations and 100.000 iterations**

| node | mean | sd | MC error | 2.50% | median | 97.50% | start | sample |
|---|---|---|---|---|---|---|---|---|
| mu31 | 6.05 | 0.60 | 0.06 | 4.87 | 6.02 | 7.25 | 4000 | 10001 |
| mu31 | 5.97 | 0.78 | 0.08 | 4.02 | 6.22 | 7.15 | 90000 | 10001 |

**<u>Tests: COMPARISON OF DIFFERENT CHAINS</u>**

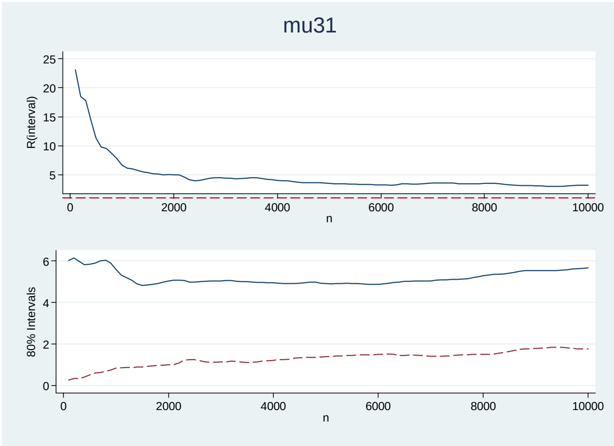
**Density Graph Intercept Trial (internet): comparison of nine density graphs (mean graph – bold line)**



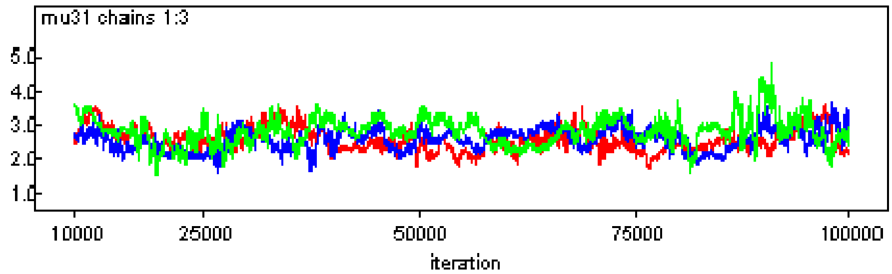

**B.G.R. statistic (R)**

| n | chain1 | chain2 | chain3 | chain4 | chain5 | chain6 | chain7 | chain8 | chain9 | average | R Brooks-Gelman-Rubin |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 100 | 0.44 | 0.22 | 0.28 | 0.14 | 0.43 | 0.30 | 0.07 | 0.39 | 0.08 | 0.26 | **23.1** |
| 200 | 0.26 | 0.27 | 0.53 | 0.15 | 0.39 | 0.49 | 0.11 | 0.67 | 0.12 | 0.33 | **18.5** |
| 300 | 0.41 | 0.30 | 0.18 | 0.25 | 0.47 | 0.55 | 0.14 | 0.51 | 0.22 | 0.34 | **17.8** |
| 400 | 0.46 | 0.22 | 0.41 | 0.47 | 0.37 | 0.55 | 0.20 | 0.73 | 0.18 | 0.40 | **14.5** |
| 500 | 0.64 | 0.58 | 0.48 | 0.73 | 0.35 | 0.43 | 0.21 | 1.02 | 0.17 | 0.51 | **11.4** |
| 600 | 0.68 | 0.58 | 0.58 | 0.89 | 0.38 | 0.45 | 0.30 | 1.38 | 0.18 | 0.60 | **9.8** |
| 700 | 0.65 | 0.75 | 0.70 | 0.49 | 0.47 | 0.45 | 0.42 | 1.52 | 0.22 | 0.63 | **9.5** |
| 800 | 0.34 | 1.11 | 0.64 | 0.41 | 0.98 | 0.75 | 0.30 | 1.49 | 0.19 | 0.69 | **8.7** |
| 900 | 0.62 | 1.24 | 0.52 | 0.47 | 1.03 | 1.40 | 0.40 | 0.92 | 0.15 | 0.75 | **7.8** |
| 1000 | 1.16 | 1.18 | 0.43 | 0.58 | 1.05 | 1.37 | 0.63 | 0.94 | 0.16 | 0.83 | **6.7** |
| ... | .... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 9000 | 2.00 | 1.45 | 1.09 | 1.79 | 1.31 | 2.21 | 1.16 | 2.48 | 2.59 | 1.79 | **3.1** |
| 9100 | 2.02 | 1.51 | 1.09 | 1.79 | 1.30 | 2.19 | 1.29 | 2.47 | 2.59 | 1.80 | **3.1** |
| 9200 | 2.01 | 1.59 | 1.08 | 1.78 | 1.29 | 2.17 | 1.38 | 2.46 | 2.58 | 1.81 | **3.0** |
| 9300 | 2.00 | 1.71 | 1.04 | 1.77 | 1.29 | 2.16 | 1.47 | 2.44 | 2.57 | 1.83 | **3.0** |
| 9400 | 2.00 | 1.76 | 1.01 | 1.79 | 1.29 | 2.13 | 1.56 | 2.43 | 2.55 | 1.84 | **3.0** |
| 9500 | 1.99 | 1.76 | 0.99 | 1.78 | 1.29 | 2.06 | 1.64 | 2.42 | 2.54 | 1.83 | **3.0** |
| 9600 | 1.97 | 1.78 | 0.96 | 1.77 | 1.28 | 1.95 | 1.62 | 2.41 | 2.54 | 1.81 | **3.1** |
| 9700 | 1.96 | 1.76 | 0.95 | 1.78 | 1.27 | 1.82 | 1.61 | 2.40 | 2.52 | 1.79 | **3.1** |
| 9800 | 1.96 | 1.74 | 0.94 | 1.79 | 1.27 | 1.72 | 1.60 | 2.38 | 2.51 | 1.77 | **3.2** |
| 9900 | 1.97 | 1.72 | 0.93 | 1.81 | 1.26 | 1.66 | 1.59 | 2.37 | 2.50 | 1.76 | **3.2** |
| 10000 | 1.99 | 1.70 | 0.92 | 1.85 | 1.28 | 1.62 | 1.66 | 2.36 | 2.48 | 1.76 | **3.2** |

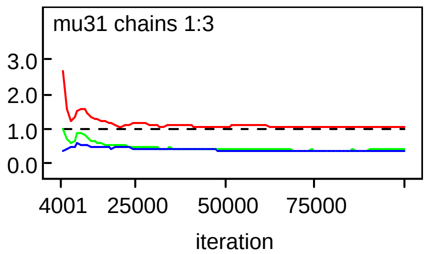
**Brooks-Gelman-Rubin plots for assessing the convergence of 9 chains**




**History Graph: comparison of three different chains**

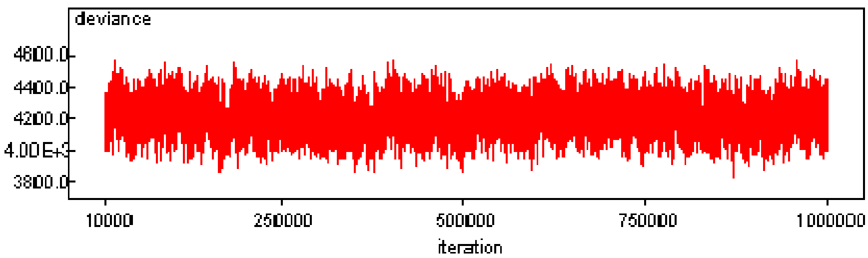


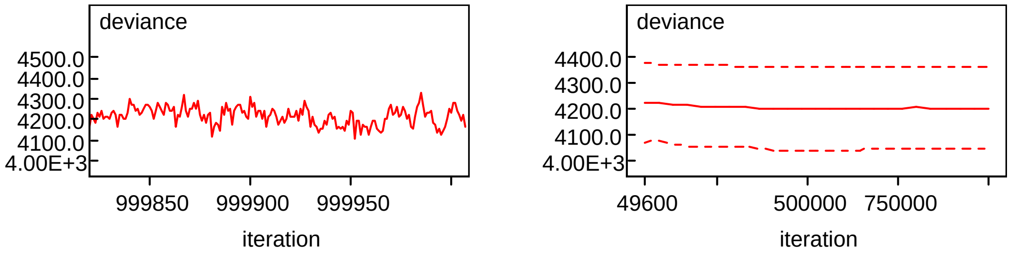

**Brooks-Gelman-Rubin plot**




## Deviance

History, Trace and Quantiles and with 1 million iterations






*Related references*

Brooks S P and Gelman A (1998) Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. **7**, 434-45

Brooks, S. and A. Gelman. 1998. General methods for monitoring convergence of iterative simulations. Journal of Computational and Graphical Statistics 7: 434-55.

Gelman, A. and D.B. Rubin. 1992. Inference from iterative simulation using multiple sequences (with discussion). Statistical Science 7: 457-511.

Gelman, A., Carlin, B.P., Stern, H.S. and D.B. Rubin. 1995. Bayesian Data Analysis. Chapman and Hall, London.

Gelman, A. "Applied Bayesian Modelling" 2003. Wiley.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. pp. 169-93. In: Bayesian Statistics 4 (eds J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith.) Oxford University Press, Oxford