University of Alberta

Two Essays on E-commerce and Retailing

by

Jin Li ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of the Doctor of Philosophy

in

Marketing

Faculty of Business

Edmonton, Alberta

Fall 2007

Your file Votre référence
ISBN: 978-0-494-33012-8
Our file Notre référence
ISBN: 978-0-494-33012-8

ABSTRACT

The first essay proposes and tests a model of consumer response to permission emails. The model of consumer response to emails uses a decision-theoretic analysis of consumer information acquisition. Email offers are interpreted as having an information structure similar to financial options. Thus, communications design can be viewed from the perspective of crafting a low option price and a high expected upside. The model motivates a set of hypotheses about communications design features that are tested in the empirical analysis.

The empirical approach constitutes an accessible way for practitioners to measure how email effectiveness is influenced by design features of the subject line, the email body, and the targeting and timing of the email campaign. It can been learnt for the focal retail chain, for example, that it is desirable to use subject lines with three or four words, including the exact date, email body which is short, including a moderate number of links (e.g., 5-10), sent out with short lead times, on weekdays, and where the design varies according to the appeal type.

The second essay uses event-study methodology to examine the short-term response of stock markets to retailers' announcements of online sales channel additions. I find that on average the announcements have had positive effects on retailers' stock price returns. I also find that retailers in different sectors get different abnormal returns from the announcements, which means that investors have different attitudes toward the suitability of different products to the Internet. I then look at the influence of the characteristics of firms and their introduction

strategies on the direction and magnitude of stock market reactions. The study finds that investors are more optimistic about the Internet channel additions of small retailers and retailers with direct channel experience. Channel introduction strategies, however, do not affect the stock market's reaction to these announcements directly. Interactions between a firm's characteristics and its channel introduction strategies suggest that investors expect big retailers to build their own sales websites and retailers with direct channel experience to sell products through rented web space.

Dedication

To my wife Helen and my son Sam.

Acknowledgement

I wish to express sincere appreciation to all individuals and groups who have supported me throughout my doctoral studies. First of all, my profound gratitude goes to my supervisor Dr. Paul Messinger for his constant guidance, encouragement, and support. I am really blessed to have such a wonderful advisor.

In addition, I thank the other committee members. Dr. Peter Popkowski Leszczyc and Dr. Yu Ma have been very helpful with the completion of this thesis. Dr. Eleni Stroulia and Dr. Yonghua Ji have provided insightful comments in shaping my research. Dr. Mengze Shi and Dr. Adam Finn have suggested good directions for further research. I also thank the other faculty members of the Department of Marketing and the Business PhD Program Office for their help.

The support and encouragement of my family enable me to successfully complete the doctoral program. Helen, I thank you for being there, for your patience, for bearing with my devotion to the research, and for taking care of our little son. Sam, I thank you for being patient with my absence from home most of time. I will spend more time with you from now on.

Finally, I am thankful to Canada, a great country of liberty, for her generosity in offering me the chance to pursue the academic career. I shall never forget the warm greetings, friendly smiles, and heart touching encouragements that I have received from countless Canadians in the past five years.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The two essays of my dissertation concern e-commerce and retailing. Specifically, I am looking at how retailers use the Internet to better communicate with and serve their customers.

Since the emergence of the Internet, the Internet and e-commerce have revolutionized both the way people shop and the way the retail industry operates. Although much research has been done on consumer behavior in the Internet environment, many questions about the impact of the Internet on the retail industry remain unanswered.

For the retail industry, the Internet is not only a new sales channel but also a new marketing communication tool. As a marketing channel, the Internet has unique characteristics. These characteristics include the ability to inexpensively store, search, and organize vast amount of product and service information, the ability to customize information on demand immediately, the ability to serve customers around the clock, and the ability to distribute certain products (e.g., CDs) quickly. As a marketing communication tool, the Internet is fast, interactive, economical, and capable of reaching consumers in a broad geographic area. Although the Internet provides a great opportunity for the retail industry, it comes with a cost. To better use the Internet, retailers must answer the following questions:

(1) How do retailers use the Internet as a new marketing communication tool to better communicate with their customers?

(2) Should established bricks-and-mortar retailers add an Internet sales channel to their channel portfolio, and if so, how do they evaluate the Internet channel additions?

My dissertation research includes two essays that address these questions. The first essay answers the first question – how to use the Internet to better communicate with consumers. I use one of the most popular Internet applications – permission email – to study how retailers design effective and efficient Internet communication with their customers. Permission emails differ from standard broadcast email communications in that customers must sign up to receive the communications. I propose and test a model of consumer response to permission emails.

My model of consumer response to emails uses a decision-theoretic analysis of consumer information acquisition. Email offers are interpreted as having an information structure similar to that of financial options. Thus, communications design can be viewed from the perspective of crafting a low option price and a high expected upside. The consumer incurs the option price to gain the possibility, but not the obligation, of taking an offer once its value is revealed. In the current study's communications context, the option price consists of the consumer's costs of attending to, processing, retaining, and following up on an email message. The expected upside is determined largely by the perceived mean and variance of the underlying email offer. My model produces a set of hypotheses about communications design features that are empirically tested.

The empirical approach constitutes an accessible way for practitioners to measure how email effectiveness is influenced by the design features of the subject line, the email body, and the targeting and timing of the email campaign. Based on this study, a focal retail chain can determine that it is desirable to use a subject line with three or four words, including the exact date, a short email body that includes a 5-10 links, sent out with short lead times, on weekdays, and with a design that varies according the appeal type.

The second essay answers the second question – how retailers should evaluate Internet channel additions. In this essay I conduct an event study to examine how the addition of online stores by brick-and-mortar retailers influences their stock price returns. Thanks to the development of Internet and e-commerce techniques, many bricks-and-mortar retailers have a new online store as part of their channel portfolios. For retailers, this dual-channel strategy is now becoming the rule rather than the exception. For example, out of 488 publicly traded retailers on the three major U.S. stock exchanges, 280 had opened online stores by 2004. This form of dual distribution may improve retail sales and enable retailers reach more market niches; it may also cannibalize retailers' offline businesses. In this study I use event-study methodology to examine the short-term response of stock markets to retailers' announcements of online sales channel additions. Event searching through Factiva finds that 78 public retailers have clearly announced online channel additions. On average, the announcements have had positive effects on the stock price returns of these retailers. I also find that retailers in different sectors get different abnormal returns from the announcements, which means that investors have different attitudes toward the suitability of selling products on the Internet.

I then look at the influence of the characteristics of firms and their introduction strategies on the direction and magnitude of stock market reactions. The study finds that investors are more optimistic about the Internet channel additions of small retailers and retailers with direct channel experience. Channel introduction strategies, however, do not affect the stock market's reaction to these announcements directly. Interactions between a firm's characteristics and its channel introduction strategies suggest that investors expect big retailers to build their own sales websites and retailers with direct channel experience to sell products through rented web space.

I believe that my dissertation research will extend our understanding of the Internet's impact on the retail industry. The findings provide retailers with guidelines for using the Internet to better communicate with their customers, serve their customers, build relationships with their customers, and compete with their rivals in the "new economy." In addition, the perspective of financial options developed in the first study may also have applicability for other marketing communications and promotions design problems. And the event study used in the second study may be used to evaluate the influence of other strategic actions on the performance of firms in financial markets

Chapter 2

Essay 1: Permission Email Options

2.1 Introduction

As electronic mail grows to be as widely used as the telephone or postal service, it is incumbent on marketers to understand how to use the medium. The advantages of email for marketing communications include rapid deployment and testing, real time feedback, high potential for personalization, good response rates, richness of the multi-media experience, and low cost (Hoffman 2000), but a key disadvantage is email proliferation (*i.e.*, consumers are receiving more and more emails, many of which are unsolicited from unknown senders). In response to a similar problem with traditional, less-targeted, marketing communications, "permission marketing" emerged as a practice that critically includes the element of permission from current or potential customers before sending out "personalized, anticipated, and relevant messages" (Godin 1999). Permission marketing has been fruitfully applied to direct mail campaigns, but it is also a natural approach for email campaigns.[1]

---

[1] As background, (a) penetration of email in the U.S. by 2004 reached an all-time high of 91 percent of Internet users between the ages of 18 and 64, with 88% of adult Internet users having personal e-mail accounts and 46% having e-mail access at work (eMarketer 2004), (b) approximately 147 million people in the U.S. use e-mail almost every day (eMarketer 2004); and (c) 90% of U.S. consumers report sending and receiving email multiple times daily (DoubleClick 2005). These trends lead Hoffman, Novak, and Venkatesh (2004) to identify email as the most used in-home Internet application for consumers. At the same time, the extent of email proliferation has grown to the level where the average U.S. consumer receives 361 emails per week (DoubleClick 2005), and consumers are indicating reluctance to open email from unknown senders (eMarketer 2006). Anti-spam laws have been a legal response to email proliferation in several countries, and the use of permission marketing has been a voluntary response by practitioners. As compared with spam emails, permission emails (1) are read more frequently, (2) are seen as more interesting, (3) generate higher click-through, and (4) result in more purchases (Kent and Brandal 2003). Permission emails also comply with anti-spam laws.

To date, however, most studies about the effectiveness of permission email campaigns have been limited to survey-based industry reports. These reports summarize consumers' perceptions and stated usage patterns for email communications. These reports, however, lack theoretical grounding and systematic consideration of how consumers actually respond to particular design features of permission email campaigns. The current study addresses these limitations.

The current study develops two intended contributions. First, I propose a multi-staged model that (a) identifies the types of variables that influence consumer response in each stage, (b) analyzes consumer response using a decision theoretic framework, (c) can be interpreted in terms of option design, (d) motivates a set of hypotheses, and (e) is empirically examined using a three-equation recursive system with binomially distributed dependent variables. Second, I illustrate a readily accessible approach to measuring the effectiveness of campaign design features. The approach analyzes aggregate data about a permission email campaign of a specialty retail chain with over 90 stores. The data consist of 645 email communications distributed over a three-year period to subscribers on this chain's email list. I examine the impact on opening, click-through, and opt-out rates of design characteristics that describe the subject-line, body, timing, and targeting of these email communications.

There has been some previous academic research on permission email, but that research has mostly focused on factors that influence consumers' willingness to be included in permission email lists. For example, factors affecting alumni willingness to be added to a school's alumni email list have been examined (Tezinde, Smith, and Murphy 2002). Similarly, the effect of promotional email on previous customers of a hotel has been examined in a field experiment to measure whether email recipients opt-out of the email list or, instead, visit the hotel's website (Marinova, Murphy, and Massey 2002). The present study, by contrast, considers logically subsequent questions, assuming a firm already has permission from a large number of consumers and is sending out emails.

I examine a set of relationships between email campaign design characteristics and consumer response that builds on a conceptual model for direct mail by Vriens *et al.* (1998) involving four campaign factors (quality of the list, characteristics of the offer, design of the mailing, and timing of the mailing) and three stages of consumer response (opening the envelope, taking notice of the elements of the mail package, and responding to the offer).

In the theoretical development, I consider the trade-off, in the spirit of cost-benefit models of sequential information seeking behavior, between fixed costs and expected benefits of opening an email and clicking on a link. According to that stream, when the marginal cost of acquiring product information is lower than the marginal benefit of using the information, consumers should continue searching for product information (Ratchford 1982).[2]

I also note that this email response problem is isomorphic in structure to the choice of purchasing a (compound) financial option. Thus, intuition from the

[2] Subsequent research in this stream reveals that when search costs of price information are low, then prices are close to competitive prices (Bakos 1997), that lowering search costs for quality information decreases prices sensitivity (Lynch and Ariely 2000), and that firms can soften price competition by differentiating themselves on the basis of consumer search cost (Zettlemeyer 2000).

options literature (dating back to Black and Scholes 1973)[3] provides a useful perspective on how communications design parameters (*e.g.*, message, layout, reach, and timing) can be interpreted in terms of the implied option costs and upside potential.[4]

In particular, the options interpretation can be summarized as follows. (a) An email confers an offer, which the receiver has the right, but not an obligation, to pursue. (b) As with all options, the consumer faces a choice of whether to incur an immediate fixed cost (the option price) in exchange for possible future upside benefits (and limited downside risk). (c) The instantiation of the option prices includes cognitive costs of attending to a communication, processing the message, and retaining the information, as well as physical costs of effort and time of following up on the message. And (d) the assessment of expected upside is positively influenced by the perceived mean and variance of the underlying asset as framed by the environment and the design features of the email communication.

My empirical analysis is closest to the interesting work of Ansari and Mela (2003), who also analyze the determinants of click-through rates.[5] There are important differences, however. First, my goal is to analyze the effectiveness of

---

[3] See Cox *et al.* (1979) for a simplified interpretation and Sharpe *et al.* (1998) for an overview. In this paper, I apply the intuitive properties and principles of options, without bringing into consideration the complexities of diversification, hedging, the Black-Scholes pricing formula (1973), or stochastic calculus.

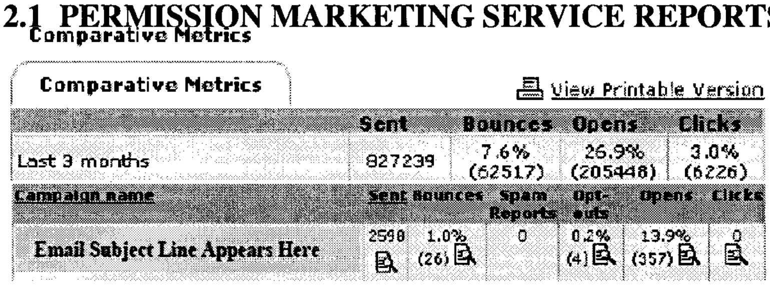
[4] This options perspective also differs from the view of traditional advertising as "interruption marketing", which involves the annoyance of less-targeted, marketing communications that urge a course of action while the prospective consumer is engaging in some unrelated activity. According to this view, consumer response is motivated by interruption avoidance when viewing television (Krugman 1983), browsing content on the web (Cho and Cheon 2004), or responding to emails (Hoffman 2000).

[5] They find, for example, negative effects associated with the order of the link and with the duration since the consumer last clicked on a link in a previous email.

broadcast emails sent out to targeted segments of customers (typically in the thousands),[6] whereas their goal is to provide automated individually-customized emails. Second, my study is based on aggregate data, rather than individual level data. (I think there is value for both kinds of analysis.) Third, I also consider the determinants of the opening and opt-out rates.

Overall, although Ansari and Mela's approach is more ambitious in several ways, mine is more immediately relevant to small and medium-sized practitioners. Only an elite group of firms (*e.g.*, Amazon, Google, Yahoo, and Dell) currently customize recommendations or email campaigns at the level of the individual user, because it requires individual level databases, advanced software, and a tech/marketing group skilled at statistics and computing science. Yet there is a small cottage industry emerging around targeted, but not individually-customized, permission email campaigns of the type I analyze. Some venders sell specialized software for particular industries (*e.g.*, the industry of residential real estate agents), while other companies provide complete fee-based permission marketing services (*e.g.*, the service used by the retail chain that I analyze). My approach basically allows users to harness the data that such

Figure 2.1 PERMISSION MARKETING SERVICE REPORTS*



*Aggregate and Individual Email Level Reports

---

[6] In my sample, there are 54 geographic areas, each ranging from 17 to 7,533 subscribers on the list and the average email is sent to 8.8 areas consisting of about 14,206 subscribers.

services already provide. Figure 2.1 provides an example of such data. The first row describes all emails sent in the last three months. The second row describes a recent email send-out, information about which is tabulated cumulatively (after three months, the send-out information is finalized in a database). This send-out (I keep the subject line anonymous) was sent to 2,598 consumers, and in the first few hours of the first day, 1.0% of the emails had bounced back and 13.9% of consumers had opened the email (none had yet clicked on a link). This paper considers a simple approach to analyzing such data.[7]

More generally, the current paper also has ties with email research in the literatures of management information systems and organizational studies. In an impassioned plea, the editor of *MIS Quarterly* (Weber 2004) calls for research on (1) the impact of e-mail on our lives and (2) better ways of managing e-mail and assisting users to deal better with the problems it poses.[8] Work that comes close to the current paper, and which happens to fall in the second category, consists of an analysis of incentive-based email advertising systems (Gopal *et al.* 2001). My work, instead, addresses measurement and management issues for permission email campaigns.

---

[7] The only additional data that practitioners need to tabulate for my approach involve descriptions of the design features of emails that are being sent out.

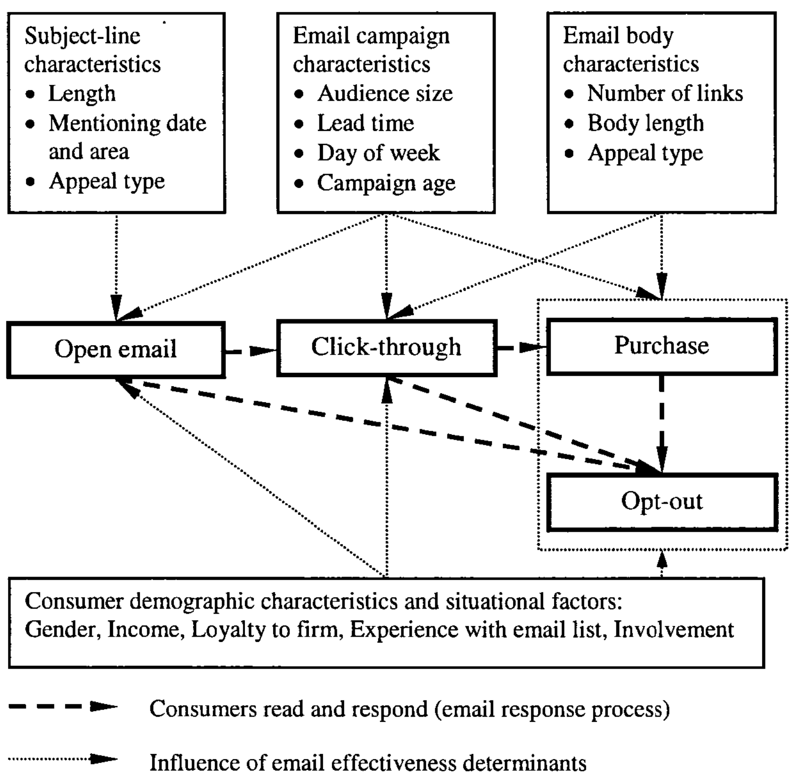
[8] The first part of Weber's call has been answered by studies of Internet usage patterns that argue that (a) email drives people's use of the Internet (Kraut *et al.* 1999), (b) important differences exist across user segments (Tassabehji and Vakola 2005 and Gefen and Straub 1997; the latter apply the Technology Acceptance Model), (c) many organizations use infrastructure to control the impact of such factors as spam and viruses (Tassabehji and Vakola 2005), (d) rich communication (*i.e.*, that is quickly understood) works through the interaction of the email medium and the organizational context (Lee 1994), and (e) privacy should not be taken for granted (Weisband and Reinig 1995). The second part of his call has been answered by such work as alterative proposed workflow mechanisms and algorithms to improve information distribution through e-mail (Zhao *et al.* 2001) and analysis of economic incentives to curb spam (Loder *et al.* 2004).

In the rest of this paper, I propose a model of consumer email response, describe the data, estimate the model, present the results, and conclude with a discussion of needed future research.

2.2 Determinants of Consumer Response to Emails

Along lines similar to Vriens *et al.* (1998), I conceptualize consumer email response as shown in Figure 2.2. The boxes connected by dashed lines in the central horizontal portion of the figure (in bold) represent the consumer response process, described as follows:

Figure 2.2 CONSUMER EMAIL RESPONSE



(1) After receiving an email, the consumer first makes a decision about whether or not to open the email.

(2) If the consumer opens the email, the consumer can (a) click on a link to navigate to the firm's web pages, (b) decide to opt out of the email list, or (c) do nothing.

(3) If the consumer clicks on the links, the consumer can (a) purchase from the retailer, (b) opt out, or (c) do nothing.

The dashed arrows in the consumer email response process indicate that the consumer can stop at any stage of the process. In principle, the consumer can also later purchase directly from the bricks and mortar stores (although I do not have data on this).

The consumer email response process is influenced by four sets of factors, shown at the top and bottom of the figure. In particular, I have the following:

(1) *Email subject-line characteristics may affect the email opening decision.*

(2) *Email body characteristics may affect the click-through, purchase, and opt-out decisions.*[9]

(3) *Email campaign characteristics may affect the opening, click-through, purchase, and opt-out decisions.*

(4) *Consumer demographic characteristics and situational factors may also influence every stage of the consumer response process.*

Due to data limitations, I focus in this paper on the first three sets of relationships (involving the email opening rate, click-through rate, and opt-out rate). I now turn to more detailed development of our research hypotheses.

2.3 Model for Email Response

[9] The email body and campaign characteristics may also affect the purchase conversion rate, if I had that data.

Figure 2.3 MULTI-STAGE MODEL OF EMAIL RESPONSE

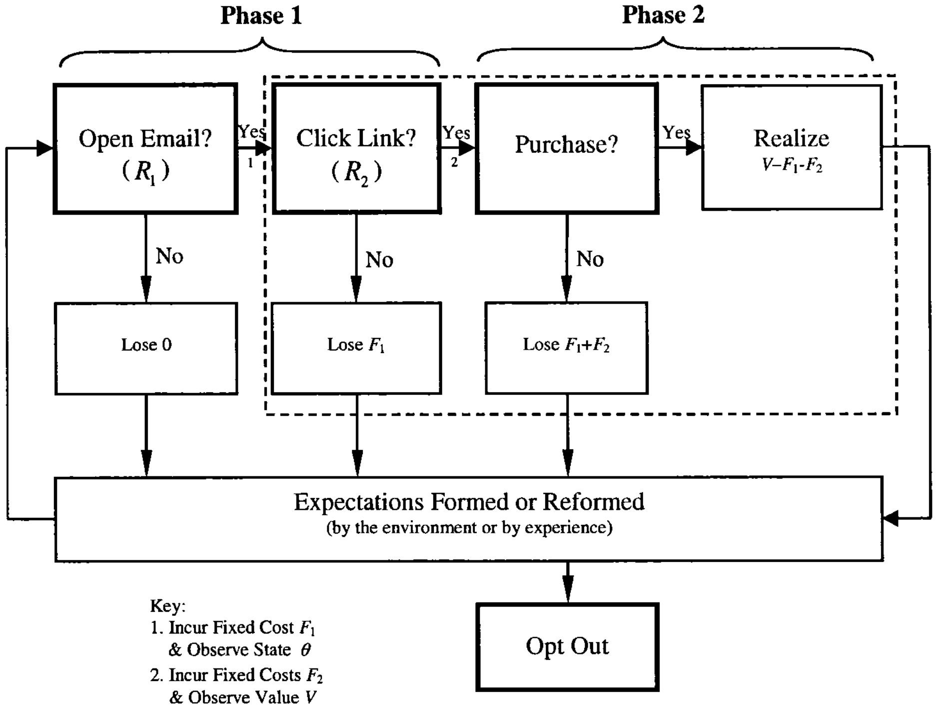


Figure 2.3 replicates the consumer response process of Figure 2.2, adding costs and values of the decisions involved. I suppose that the email contains an offer of some type (*e.g.*, a new product, price promotion, or hobby event), and the potential value of the offer, V, is initially uncertain to the consumer. The environment, past experience, and features of the current email help establish the consumer's general expectations concerning the distribution of V. I start my analysis assuming that the consumer has already agreed to receive permission emails from the company in question. By agreeing to receive permission emails, the consumer has already committed to incurring costs, including (a) "inbox" clutter and associated use of memory space, and (b) some effort required to organize and sort through the "inbox." I specifically distinguish between the minimum amount of unavoidable effort or cost required to handle an additional

email and any effort or cost in excess of that minimum that is discretionary. The former effort includes the effort involved to make an initial cursory cognitive scan of the email. The latter effort includes paying additional attention to and engaging cognitive processes to consider in greater depth whether a certain email offer is worth opening, considering, and pursuing. Since the former costs are sunk relative to the decision of whether to open an email, these costs need not be considered in analyzing the opening or subsequent decisions.[10] For simplicity, I assume that the consumer is risk-neutral.

Starting from this point, the consumer now faces four decisions (shown in the bold outlined boxes in Figures 2.3).

1. Decision to open email. After the aforementioned minimal effort to handle the email, the consumer may incur discretionary effort to attend to, weigh the relative expected benefits and costs, open the email, and begin processing the information contained in the email. If the consumer chooses not to open the email, then there is no loss or gain associated with this decision. If the consumer chooses to open the email, two things happen. (1) A fixed cost of F_1 is incurred for the time and cognitive effort of attending to the title and sender information, opening, and processing the email. (2) The consumer learns something about the prevalent information environment by opening the email, modeled as the realization of a (possibly multidimensional) parameter θ. (In control theory, this would be called a state parameter.) The consumer can condition his or her

[10] The one exception to this is that, I will see that similar future cost are relevant to the decision of whether to stay in a permission marketing relationship with the firm or whether to "opt out."

expected distribution of V on the state parameter θ. The consumer then moves to the next decision.

2. *Decision to click on a link.* The consumer can then choose to click on a link in the email (using the information the consumer realized about θ). If the consumer chooses not to click on the link, the consumer incurs no further discretionary effort and relinquishes the possibility of learning and realizing the value of V. The consumer faces an accumulated loss of F_1 and moves to the fourth "opt-out" decision. If the consumer chooses to click on the link, two things happen. (1) Cognitive costs of F_2 are incurred. (2) The exact value of V is revealed. The consumer then moves to the next decision.

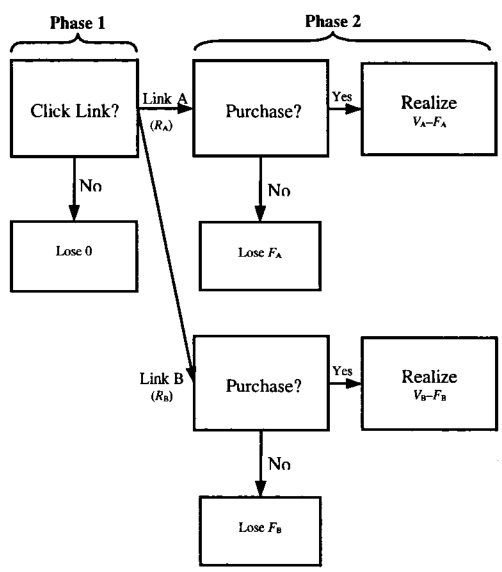
3. *Decision to purchase.* At this point, the consumer has to decide whether or not to take the offer. If no, the consumer does not realize the value of V. If yes, the consumer will realize the value of V. In the former case, the consumer faces an accumulated loss of $F_1 + F_2$. In the latter case, the consumer will realize the net value $V - F_1 - F_2$. The consumer then moves to the last decision.

4. *Decision to opt-out.* After Decision 2 or Decision 3 is completed, the consumer can choose whether to opt-out of the email list. By deciding to maintain the permission marketing relationship (not "opting out"), the consumer is making a decision to incur future fixed costs associated with the required effort to handle and process subsequent permission email. The benefit of maintaining this relationship is the expected upside of being exposed to future offers. If the consumer finds, over time, that his or her expectations are overly optimistic (*i.e.*,

he or she is repeatedly disappointed), the consumer will adjust his or her expectations downward. Eventually, the consumers may wish to discontinue the relationship with the vendor.

As a final note, the dashed box in Figure 2.3 describes the consumer choice of whether to click on a link when the email contains a single link. If there are two links, the dotted box will take the modified form shown in Figure 2.4. Here the consumer has a choice of whether to click on link A or B. Each of these links has different fixed costs F_A and F_B, and different value distributions V_A and V_B. If there are more than two links, a similar model with multiple links in parallel will apply.

Figure 2.4 CHOICE OF MULTIPLE LINKS IN PARALLEL



Analysis and Empirical Specification. Now I analyze the process working backward in Figure 2.3. I start from the purchase decision. At this point, the costs $F_1 + F_2$ are sunk, so these do not influence the purchase decision. The consumer will, thus, only purchase (*i.e.*, take the offer) if $V \geq 0$.

Moving one decision back, I now consider the decision of whether to click on the link. For the purpose of this decision, the cost F_1 associated with opening the email is sunk, but the cost F_2 associated with clicking on the link is not sunk. At this point in the process, the consumer has already learned something about the information environment by opening the email, as represented by the state parameter θ. Applying the logic of previous paragraph, the consumer knows that after clicking on a link he or she will realize (a) $V - F_2$ if the realization of V turns out to be non-negative, or (b) $-F_2$ if the realization of V turns out to be negative. Which of these outcomes is likely to occur is dependent on $p(V;\theta)$, the probability density function for V, given θ. The expected return from clicking on a link is therefore

$$R_2(\theta) = \int_{V<0}(-F_2)p(V;\theta)dV + \int_{V\geq 0}(V - F_2)p(V;\theta)dV \equiv \mathrm{E}_{V\geq 0}(V \mid \theta) - F_2. \qquad (1)$$

Stated in other words, the consumer will only click on the link when this expected return is non-negative, $R_2(\theta) \geq 0$.

For the decision to open the email, I write $f(\theta)$ as the perceived probability (prior to opening the email) of obtaining a particular information environment θ (after open the email). Analogous reasoning indicates that the expected return for the consumer from opening is

$$R_1 = \int_{\{\theta | R_2(\theta)\geq 0\}} \mathrm{R}_2(\theta) f(\theta) d\theta - F_1, \qquad (2)$$

and the consumer will only open this email if this expected return is non-negative, $R_1 \geq 0$. This concludes the first pass on the sequential process involving decisions to open an email, click on a link, and purchase the offer.

I observe that I am successively building up structures inherently similar to Equation 1 that involve two terms: (1) an expectation over the positive portion of the distribution of returns and (2) various fixed costs.

Equation (1) assumes that there is one link contained in a particular email. When there are two links (as shown in Figure 2.4), the return from choosing the more desirable link is given by

$$R_2'(\theta) = \mathrm{Max}(R_{2A}(\theta), R_{2B}(\theta)), \tag{3}$$

where $R_{2i}(\theta) = \mathrm{E}_{V_i \geq 0}(V_i \mid \theta) - F_i$ (from Equation (1)) and $i = \mathrm{A\ or\ B}$. The consumer will only choose this link if the expected return is non-negative, $R_2'(\theta) \geq 0$. When there are more than two links, a similar equation applies.

2.4 Options Interpretation

I could further analyze the above consumer information-seeking problem from first principles, but I simplify the discussion by noting that this information problem shares a common structure with simple and compound financial options. Once I demonstrate this commonality, I can draw on results and intuition from the theory of financial options to describe email communication design.

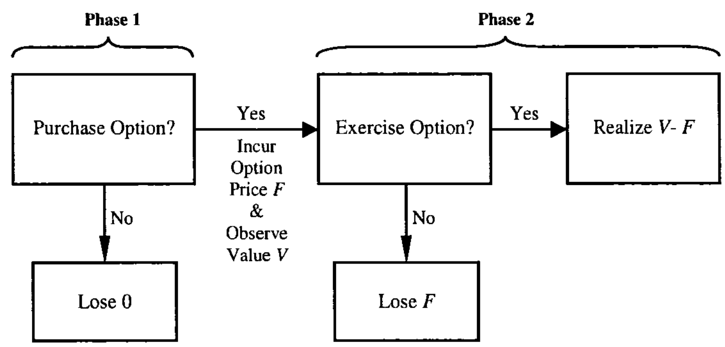
Many finance scholars believe options theory applications are ubiquitous in business. For example, applications abound in financial markets, contract law, real estate development, and production scheduling (see Sharpe *et al.* 1998). An application in the academic world involves hiring faculty to a tenure track, where the option price is the pre-tenure salary and, at tenure time, the university can choose whether to exercise the option (*i.e.*, grant tenure). Less recognized is the application of options to marketing communications. In this section, I point out that the decision-theoretic structure described earlier can be interpreted from the perspective of financial options theory.

Simple Option Structure. The basic structure of a simple option in a two-period context is illustrated in Figure 2.5. The decision-maker, first, faces a choice of whether to pay a fixed fee F (the option price) for the right, but not the obligation to purchase an asset at a preset price (the strike price). At the time of purchasing the option, the net value (gross value less the strike price) of the underlying asset V is a random variable, but by the time of exercising the option, V takes on a known non-random value. If the option is exercised, the economic consequence to the consumer is $V- F$; otherwise, the consumer loses F. [11]

[11] Formally, what is described in this paper is referred to as a "call" option because it involves the right to "call" the underlying asset before or on the expiration date. By contrast, a "put" option is the right to sell an underlying asset at a particular price before or on the expiration date. Many other variations exist. The economic function of call options is to allow a buyer to "lock-in" a needed resource in advance for a predetermined price without being subject to the vagaries of a volatile resource market (which, again, limits the downside risk). Put options similarly can "lock in" a needed sale before investing in production. As background, our consideration of a two period context is closer to a European option which can only be exercised on the expiration date (whereas an American option can be exercised on or before the expiration date).

To maintain the desired focus, we do not bring into consideration issues of diversification. Nor do we apply the Black-Scholes option pricing model (1973) or propose a variation thereof.

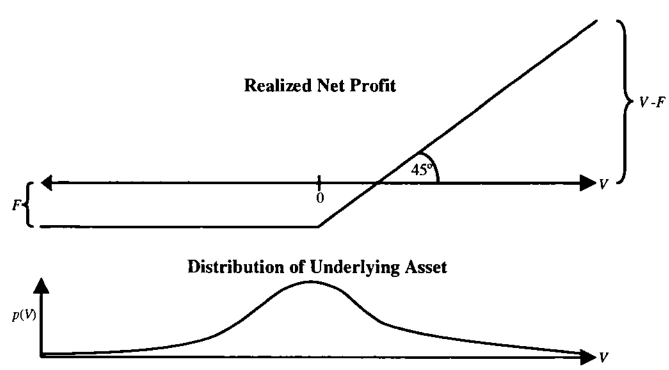
Figure 2.5 SIMPLE OPTION STRUCTURE



My main point of this section is that the simple two-period options structure of Figure 2.5 is isomorphic to (takes the same form as) the consumer's choice of whether to click on a link in an email, as represented within the dashed rectangle in Figure 2.3. Therefore, I can view the decision to click on a link as a decision with the same inherent structure as purchasing a simple (call) option.

As background, the risk profile of such an option is shown in Figure 2.6.

Figure 2.6 RISK PROFILE OF AN OPTION



The upper graph shows how the possible realized values in the underlying asset V, ranging on the horizontal axis from $-\infty$ to ∞, translate to the realized values of the option, ranging on the vertical axis from -F to ∞. On the "upside" (when the realized value of the asset V is positive), the realized value of the option is V-F; on the "downside," the realized value of the option is $-F$. The bottom graph shows a possible density function for the underlying asset V.

These considerations give rise to three well-known properties that describe how options differ from simple gambles:

Property 1. The downside exposure is limited to the option price (F) and the upside potential is unlimited (V-F). For simple gambles, both the downside exposure and the upside potential are unlimited (*i.e.*, both are V).

Property 2. Option value is only influenced by the upside potential of the underlying asset, not the downside risk. The value of a simple gamble is influenced by both.[12]

Property 3. High variance of the underlying asset is often a desirable property of an option, but not a desirable property of a simple gamble.[13]

These properties describe simple options. I now consider compound options. Note that Figure 2.4 describes a compound option in parallel (choosing between links A and B), and the upper portion of Figure 2.3 describes a compound option in series (choosing to (1) open an email, (2) click on a link, and (3) purchase an offer).

---

[12] In other words, an option is more desirable only if the upper portion of the distribution (over the range V>0) has higher expected value (*i.e.*, if $E_{V>0}(V)$ is increased). A simple gamble is more desirable the higher the expected value, E(V) (over the entire support of V). Thus, an option is not more desirable if the upper portion of the distribution is unchanged, even the lower portion of the distribution has higher expected value and E(V) is increased.

[13] For example, an option is more desirable the greater the variance in the underlying asset, if E(V) = 0, the distribution of V is symmetric, and consumers are risk neutral.

As I stated earlier, compound options in parallel are assessed by maximizing over the different possibilities (A and B). Compound options in series are assessed sequentially. Intuitively, for compound options in series, the stages of consumer choice in Phase 1 of Figure 2.3 are similar to stages of a poker hand. At each stage, it is necessary to invest more to maintain the option of seeing whether the hand wins, and the additional investment is balanced against the updated expectation of winning. In Phase 2, the consumer learns the underlying value of the hand. The analogy ends there, however. Unlike poker, options usually do not involve players competing against one another, with bidding and the possibility of bluffing.

In principle, I can construct more complicated compound options consisting of various combinations of simple options, compound options in series, and compound options in parallel. I can assess the expected return of such compound options by repeated application of the associated equations (1), (2), or (3), above. Examination of these equations indicates that the result retains a form similar to a simple option, which I summarize in the following:

> *Proposition.* All options that consist of combined sequences of (1) simple options, (2) options in series, and (3) options in parallel give rise to an expected return that involves a tradeoff between a conditional expectation evaluated over the positive possible realizations (i.e., the upside potential) and various (possibly weighted) fixed costs (the options price). [This is shown by recursion.]

Thus, many compound options ultimately involve balancing the upside potential against the option price, like a simple option. This is a useful result because my characterization of the email decision process in Figure 2.3 is stylized and simplified. This proposition indicates that more complex combinations of decisions in Phase 1 in Figure 2.3 would still give rise to a similar tradeoff between expected upside and the options price components.

To summarize the above discussion, the family of compound options (constructed from combinations of simple options, options in series, and options in parallel) possesses the following additional three properties:[14]

Property 4. Compound options balance fixed costs against the expected upside. The downside is limited and variance in the underlying asset can be desirable.

Property 5. Compound options in series provide opportunities to drop out as more information is learned before incurring all the fixed cost.

Property 6. Compound options in parallel have increased net expected upside, often with diminishing returns.

Putting the above six properties together, I see that there are several ways to increase the likelihood that potential buyers will purchase an option: (1) lower the option price (*i.e.*, the fixed costs), (2) raise the expectation of the underlying

[14] Property 4 arises by combining the above Proposition with the earlier Properties 1 – 3. Property 5 is a restatement of the intuition. Property 6 arises by noting that adding another argument to a maximum always increases its value. The diminishing returns can be shown if the additional options in parallel are drawn randomly from a set of possible options. Then the expected contribution of each additional option to the compound value will have diminishing returns.

distribution, (3) raise the variance of the underlying distribution, (4) avoid frontloading sequential options with the costs before conveying sufficient information to raise the expected upside, and (5) add to the number and diversity of possible choices included in compound options in parallel. My analysis of emails or other communications as options will thus involve choosing the communications parameters to achieve these various objectives.

2.5 Empirical Specification

I observe that I am successively building up structures inherently similar to Equation 1 that involve two terms: (1) an expectation over the positive portion of the distribution of returns and (2) various fixed costs. In particular, I note that Equations (2) and (1) take similar forms:

$$R_1 = \int_{\{\theta | R_2(\theta) \geq 0\}} \mathrm{R}_2(\theta) f(\theta) d\theta - F_1, \quad R_2(\theta) = \mathrm{E}_{V \geq 0}(V \mid \theta) - F_2. \tag{4}$$

$$\eta_1 = up_1 - f_1, \qquad \eta_2(\theta) = up_2(\theta) - f_2(\theta) \tag{5}$$

where (a) the terms η_1 and η_2 describe the attractiveness (expected return) of deciding to continue the process, (b) the positive terms on the right-hand side are applicable conditional expectations, and (c) the negative terms describe applicable costs. The empirical section of this paper will estimate equations of this form.

As a final note, the final decision to opt-out can also be viewed as a choice to acquire an option. In particular, I can describe the opt-in decision by "upside potential" minus the fixed costs associated with staying in the permission marketing relationship. These costs consist of the unavoidable effort required to handle an additional email, discussed earlier. I can thus describe the opt-out decision by an equation

$$-\eta_3 = -up_3 + f_3, \tag{6}$$

where η_3 is the "attraction" of opting in ($-\eta_3$ is the attraction of opting out), $-up_3$ is the foregone upside potential arising from opting out, and f_3 are the fixed costs avoided by opting out. This equation will be estimated in the empirical section.

Thus far, because I have not yet discussed operational determinants of up_j or f_j in Equations (4), (5), and (6), the theoretical development is abstract and not particularly content-rich. The applied content enters by identifying a relevant set of email design attributes (summarized by a vector $\mathbf{x}$) and by trying to understand and measure how each email attribute either embodies or influences (1) the cognitive and other fixed costs of the option decision, f_j, or (2) the expected upside of the option, up_j. If a given design attribute x influences both up_j and f_j, there may be an internal optimal solution (provided $up_j(x)$ and $f_j(x)$ have suitable shapes). Other design attributes influence either up_j or f_j, subject to possible constraints that limit the value of an attribute or where increasing one attribute may cause another design attribute to change. Such constraints may be due to short-term technical restrictions, or there may be inter-temporal relationships with future attributes or states that influence future utility.

The management problem is to optimize $\mathbf{x}$.[15] The empirical problem, to which I now turn, is to identify relevant measurable design attributes and understand how they influence the upside potential and fixed costs of each consumer's email decision.

2.6 Hypotheses

I can now motivate my hypothesis by considering how available design attributes influence the upside potential and fixed costs of each email option. Concerning the impact of subject line variables on the opening rate, I have:

H1 *Moderate **subject line length** will generate higher opening rates than short or long subject lines.* The rationale is that cognitive costs increase proportionately with email subject line length, but upside potential increases with diminishing returns.

H2 *The **appeal type** has a main effect on opening rate and an interaction with the subject line length.* The main effect arises because of variation in cognitive costs and upside potential across different appeal types. The interaction arises because it may take longer to explain the benefit for certain appeal types than for others.

---

15 I leave the more complex aspects of this optimization problem to future research. For example, if up_j and f_j are continuous increasing functions of a single variable x (and there are no short-term constraints on x or inter-temporal relationships with future outcomes), the optional x^* is such that $d(up_j)/dx = df_j/dx$ (I will argue that this is the case for the subject-line length variable). Alternatively, if x takes on two discrete values, I find the one that yields, higher value of $(up_j - f_j)$. If there are two attributes related by some constraint, $\mathrm{T}(x_1, x_2) = 0$, then I have a simple two variable optimization related by the Lagrange multiplier. More generally, I have a complex dynamic programming problem where increasing expected return from a current email may increase the opt-out rate and therefore reduce future returns.

It is important to note that mentioning the exact date or area name of an event in a subject line may have an ambiguous effect. On the one hand, cognitive costs are lower because the benefit is more clearly articulated. On the other hand, the upside potential may be better or worse. If the items mentioned fit the desires of a given respondent, the expected upside is greater. If they do not fit, the expected upside potential is less.[16] Because the influence of this factor could go either way, I do not propose any hypothesis for mentioning the exact date or area name of an event in a subject line. But I do examine this effect in the empirical test part.

Concerning the impact of email body variables on click-through and opt-out, I have:

H3 *The greater the* ***number of links*** *in the email, the higher the click-through and the lower the opt-out rate (with diminishing returns).* This is because offering a menu of diverse offers that appeal to different kinds of people increases the variance of the composite offer (and therefore the upside potential). This occurs with diminishing returns (because of overlapping benefits). In principle, more links could also add cognitive costs, which suggests that a moderate number is generally optimal, but the marginal cognitive cost of processing one additional link is relatively small. In addition, I make an effort to control for cognitive costs by including two other variables (email body length and appeal type).

[16] From a broader perspective, it may nevertheless be desirable for all concerned to let customers know early on whether the product is a good fit or not, even if the impact on the opening rate is negative.

H4 ***Longer email body*** *reduces the click-through rate and increases the opt-out rate.* The reason for this hypothesis is that cognitive costs increase linearly with email body length. In principle, the upside potential also increases (with diminishing returns) if more benefits are communicated, but I make an effort to control for upside potential by including other variables (number of links and appeal type).[17]

H5 *The* ***email-body appeal type*** *has a main effect on the click-through and opt-out rates, and interaction effects with the number of links and email body length.* This is again due to the variation in cognitive costs and upside potential for different appeal types. In particular, price promotion email may generate a higher click-through and lower opt-out rate than those with other appeals.

For the email campaign characteristics, I have the following hypotheses:

H6 *The longer the* **lead time**, *the lower the click-through rate and the higher the opt-out rate.* Intuitively, longer lead times raise the cognitive costs (to store, plan, and remember or retrieve the information) and hence lower the click-through rate and raise the opt-out rate. In addition, I think the appeal type may have an interaction with the lead time because different appeals may require different lead times.

H7 *The smaller the* **audience size**, *the higher the opening rate, the higher the click-through rate, and the lower the opt-out rate.* The rationale for this is

[17] If the latter variables do not completely control for the upside potential, however, the impact of email body may be ambiguous (as I will discuss later, I think this may be the case).

that emails to smaller audiences can be more targeted and therefore they may have higher upside potential for those receiving the email.

H8 *The opening rate and click-through rate for emails sent out on **workdays** will be lower (or higher) than for those sent out on **weekends** and **holidays** if the valuation of time is higher (lower) on workdays than on weekends and holidays.* The effects on openings and click-throughs provide an indication of consumers' relative perceptions of the value of time on workdays and week-ends. Opt-out is more ambiguous. If consumers feel their time is wasted after opening an email or clicking on a link, they will be more annoyed on the type of day when the value of their time is greater, so they may be more likely to opt-out. On the other hand, it takes time to opt out, so I might think consumers would be less likely to opt-out.

2.7 Data

I collected the data from a North American retail specialty chain of more than 90 bricks-and-mortar stores with substantial online presence. The chain, which I refer to as *hobby-retailer*, sells products and services associated with a leisure activity for which consumers develop skills, take lessons, enter tournaments, and learn about and discuss the professional and amateur circuits.[18] In June 2002, to manage its broadcast email campaign to customers, this company started to use a standard, commercially available permission email system called the Constant Contact® Do-It-Yourself Email Marketing® system

[18] In this paper, the company name and the exact products and services offered are disguised or presented in an anonymous fashion. Sufficient information, however, is included to provide a complete understanding of the communications issues present.

(www.constantcontact.com). The retailer built up its email list in two ways: (1) when customers purchased products and services in the retailer's brick-and-mortar stores, customers were asked whether they wished to subscribe to receive email communications from the retailer and to leave their email addresses and (2) customers could also subscribe to the email list online when they shopped at or browsed the retailer's website. In these two ways, permission was obtained.[19] The email subscribers can unsubscribe from the email list at any time because an opt-out hyperlink is part of each email.

My data cover all email advertisements and promotions sent out over a three-year period starting when this communication program began. The unit of analysis is an email. A total of 650 emails were sent out from June 17, 2002 to May 27, 2005, inclusive. The first five went to less than 50 customers for internal testing, and I deleted these from the dataset. For each email, I have information about such things as the number of recipients, subject-line design, email body design, timing, opening rate, click-through rate, and opt-out rate. The specific variables for which I have data are summarized in Table 2.1.[20] Since several of the variables have wide ranges from very small to very large values, I used the natural log transformation to rescale these variables in terms of proportional changes.

---

[19] This is a standard opting-in strategy, which avoids unintentional subscribers being included in this study. Johnson *et al.* (2002) find that opting-out generates higher participation in an online survey than opting-in. They also suggest that opting-out strategies may lead to more people who unintentionally subscribe.

[20] If suitable data were available, it would also be interesting to study the impact of email communications on in-store sales and online sales.

Table 2.1 DESCRIPTION OF VARIABLES

	Dependent Variables	Mean [a]	Min	Max	Trans
OpeningRate	Percentage of recipients who open a given email.	.430	.183	.787	
Click-through Rate	Percentage of opens that result in at least one click-through from a link.	.038	0	.305	
Opt-out Rate	Percentage of opens for which the "unsubscribe" button is clicked on.	.004	0	.031	
	Independent Variables: Subject Line(SL)				
Length of SL	The number of letters in a subject line (including spaces and punctuation).	33.9	5	80	Ln(x)
Area Name Mentioned	Dummy variable. "1" indicates an applicable city, community, or street is mentioned. "0" otherwise (including mention of a state, province, or country).	419			
Exact Date Mentioned	Dummy variable. "1" indicates an exact day of an event is mentioned. "0" otherwise (but it may mention a month or year).	273			
Type of SL	Five dummy variables equal to "1" under the following conditions ("0" otherwise).	[b]			
	Independent Variables: Email Body (EB)				
Number of Links	Number of hyperlinks in an email. (Each linked to a web page of the retailer.)	3.2	1	36	Ln(x)
Length of Email Body	"Number of scrolls" measured by displaying the email full-screen using 100% zoom size on a 17-inch LCD monitor.	2.1	1	5	
Type of EB	Dummy variables used the same coding scheme as for subject-line type.	[b]			
	Independent Variables: Campaign Characteristics				
Number of Areas	Each email is sent to one or more areas, defined by the company, and checked off by the consumer when subscribing. Each area consists of a city or town and the surrounding area. By May 2005, there are 54 geographical areas, four of which were in the U.S. and the rest in Canada. The number of subscribers in each area ranges from to 17 to 7,533 with a mean of 1,367. Each of the emails was sent to one or more areas.	8.8	1	54	Ln(x)

Table 2.1 DESCRIPTION OF VARIABLES (CONTINUED)

Lead Time	Number of days between the date the email was sent out and the date the announced event happens or starts (if the event lasts more than one day). If an email is sent on or after the event day, I code it as zero (and the transformed variable also has a value of 0).	10.5	0	369	Ln(x+1)
Time Since the First Email (Age of the Email System)	The number of days the retailer had been using the Constant Contact® email system by the date a given email was sent out. The very first email was sent out on June 17, 2002, and the variable was coded as 1. The very first email in my dataset (after deleting the first five test emails) was sent out on day 11.	626.9	10	1075	Ln(x)
Workday	Dummy variable. Measure whether an email is sent out on a non-holiday weekday. In particular, "1" describes any non-holiday weekday and "0" describes any holiday weekday or weekend day.	607			
Both Countries	Dummy variable. Indicates an email went to consumers in both the U.S. and Canada. This case only applies for 16 emails.	16			
U.S. Only	Dummy variable. Indicates that the e-mail only went to consumers in the U.S. (in which case this variable equals "1", otherwise, it equals "0"). This and the previous variable are mutually exclusive, but not collectively exhaustive; if both are zero, the email went only to customers in Canada, which is captured in the intercept.	91			

[a] For the dummy variables, this is not a mean, but a sum or count (of 645 total emails).

[b] Subject line type and email body type are each described by five dummy variables: A given subject line or email body may be classified into more than one (or none) of these categories. The variables for subject line type are distinct from the variables for email body type because the most applicable categorization of the subject line was not always the same as for the email body. Often, for example, the subject line is described by only one type, but the email body is described by more than one type.

Store openings indicates the opening of a new bricks-and-mortar store. *Price Promotions* indicates the email contains information about price discounts, seasonal sales, or coupons in bricks-and-mortar stores or in the online store. *New Products* email provides new product information. *Training programs* indicates the email describes online or offline training programs or courses. *Tournaments* indicates the email provides information about upcoming tournaments, games, or matches. In total, 37 subject line and 41 email bodies were classified as Store Openings; 18 subject line and 112 email bodies were classified as Price Promotions; 114 subject line and 219 email bodies were classified as New Products; 165 subject line and 200 email bodies were classified as Training Programs; and 149 subject line and 235 email bodies were classified as Tournaments.

There are no missing values for any variables in the analysis of the opening rate, so the sample size for the analysis of opening rate is 645 emails sent out from June 27, 2002 to May 27, 2005, inclusive. One email contained missing values for several key variables for the analyses of click-through and opt-out rates, so for these analyses the sample size is 644. The average number of email addresses to which a typical email was sent was 14,206 (the range was from 53 to 83,100). Of those sent out, the average number of broadcast messages that did not bounce back was 12,721 (the range was from 42 to 75,309).

2.8 Statistical Model

From the theoretical development, I am interested in whether the attraction, η_{ji}, for opening the email, clicking on a link, or opting out, is influenced by the independent variables as described by the above hypotheses, where i indexes the email ($i = 1,\ldots,645$) and j indexes the dependent variable ($j = 1,2,3$). For the j^{th} dependent variable, I write the associated $k = 1,\ldots,K_j$ independent variables as x_{jki}.[21] Note that the assignment of the independent variables in the (three-dimensional) array [x_{jki}] describes how I encode the relationships associated with Figure 2.3 and the hypotheses H1-H10.[22] I write η_{ji} as follows:

[21] In my notation, for any variable with two or more subscripts, the first subscript (*j*) denotes the associated dependent variable and the last subscript (*i*) indicates the applicable email.

[22] Note that, to test the hypotheses, my coding of [x_{jki}] includes quadratic terms and interaction effects. The exact specifications that I estimate will become clear in the Tables 2 and 3 that report the results.

(*Linear Predictor*) $$\eta_{ji} = \beta_{j0} + \beta_{j1}x_{j1i} + \cdots + \beta_{jK_j}x_{jK_ji} = \sum_{k=0}^{K_j} \beta_{jk}x_{jki} \,. \qquad (7)$$

I have no direct measure of the attraction variables, η_{ji}, however, so these variables are latent in the model. Instead, for the i^{th} email, I have, as dependent variables, the *opening rate*, the *click-through rate*, and the *opt-out rate* (obtained from reports of the type shown in Figure 2.1). Each of these variables is a success rate that ranges 0 to 1, which I write as y_{ji}. Associated with each dependent variable is the number of trials a_{ji} from which the success rate is calculated. That is, for the i^{th} email, y_{1i} is the opening rate ("opens" in Figure 2.1) and a_{1i} is the number of consumers that received that email (emails that bounced back are not included). Similarly, y_{2i} is the click-through rate ("clicks" in Figure 2.1) and $a_{2i} \equiv a_{1i}y_{1i}$ is number of consumers that opened the i^{th} email. Lastly, y_{3i} is the opt-out rate ("opt-outs" in Figure 2.1) and $a_{3i} \equiv a_{2i}$ is again the total number of consumers that opened the i^{th} email.

For these dependent variables, it is natural to model the number of successes $a_{ji}y_{ji}$ as binomially distributed with a_{ji} trials and probability of success p_{ji},

(*Distribution Assumption*) $$a_{ji}y_{ji} \sim \text{binomial}(a_{ji}, p_{ji}),\ i = 1,\ldots,645\,. \qquad (8)$$

To complete the model, I must specify the relationship between the success probability, p_{ji}, and the underlying attraction, η_{ji}, for opening an email, clicking on a link, or opting out. Intuitively, p_{ji} (a latent variable that ranges from 0 to 1) should be an increasing function of η_{ji} (a latent variable that ranges from $-\infty$ to ∞). In principle, many functions could serve, but because of its tractability, I assume a logit function to link p_{ji} and η_{ji} (which is commonly used with binomially dependent variables in this type of model; see McCullagh and Nelder 1989):

(*Link Function*) $$p_{ji} = \frac{e^{\eta_{ji}}}{1+e^{\eta_{ji}}}. \quad (9)$$

For $j=1$, Equations (1), (2), and (3) constitute the statistical model for the opening rate based on the binomial distribution. For $j=2 \text{ or } 3$, I need to modify Equation (1) to include the relationships between the dependent variables shown in Figure 2.2. I accordingly consider a recursive structure of the following form

(*Recursive Structure*)

$$\eta_{1i} = \sum_{k=0}^{K_1} \beta_{1k} x_{1ki},\ \eta_{2i} = \sum_{k=0}^{K_2} \beta_{2k} x_{2ki} + \gamma_{21} y_{1i},\ \eta_{3i} = \sum_{k=0}^{K_3} \beta_{3k} x_{3ki} + \gamma_{31} y_{1i} + \gamma_{32} y_{2i} \quad .(10)$$

In this recursive structure, the click-through and opt-out rates do not exert causal influences on the opening rate, because of the temporal sequencing, but the opening rate may influence click-though and opt-out rates. Similarly, the opt-out rate does not exert a causal influence on the click-through rate, but the click-through rate may influence the opt-out rate.[23]

[23] Although I explicitly consider causal relationships between the three dependent variables, it is conceivable that there is non-causal correlation between the three dependent variables induced by common left-out independent variables in two or three of these equations. One possible left out variable is the basic attractiveness of the event or offer being promoted. Then I might expect correlations (but not causation) between low opt-out rates, high click-through rates, and high opening rates. To try to avoid this problem, I include a number of email campaign variables as determinants of all three dependent variables, and I include several variables describing the email body as determinants of both the click-through and opt-out rates.

Another thing worth mentioning is that the recursive models are not consistent with the theoretical model given in Equation (2), which is a forward-looking model (*i.e.*, opening decision depends on anticipated click-through rate). The explanation of this inconsistency is, in the empirical model of opening rate, when consumers make decision, click-through and opt-out actions haven't occurred yet. So I do not put the click-through or opt-out action in the opening rate model as independent variables. In other words, all that can be included in the opening rate is expectations of the future actions. But these expectations, themselves, are mostly based on variables that can not be observed at the time of opening, so those also enter as error and do not influence the expected value of η_1.

Overall, Equations (8), (9), and (10) describe the statistical model. I estimate each of the three equations in (10) using the binomial case of the generalized linear model (McCullagh and Nelder 1989). In the terminology of the generalized linear model, Equation (7) is the *linear predictor*, Equation (8) is the *distributional assumption*, and Equation (9) is the *link function*. The estimation is done using the GLM procedure in the software package *R*, version 2.4.0 (see Venables and Ripley 2002 for an overview).[24] I now turn to the results.

---

[24] The likelihood function arises as follows. For the i^{th} email, the probability of observing $a_{ji}y_{ji}$ successes from a_{ji} trials, for a binomial distribution, is $f(a_{ji}y_{ji}) = \binom{a_{ji}}{a_{ji}y_{ji}} p_{ji}{}^{a_{ji}y_{ji}}(1-p_{ji})^{a_{ji}-a_{ji}y_{ji}}$, where $a_{ji}y_{ji} = 0, 1, \ldots, a_{ji}$. So the log-likelihood function (dropping the j subscript throughout) is

$$\log\prod_{i=1}^{n} f(a_i y_i) = \sum_{i=1}^{n} a_i y_i \log(p_i/(1-p_i)) + \sum_{i=1}^{n} a_i \log(1-p_i) + \sum_{i=1}^{n} \log\binom{a_i}{a_i y_i},$$

where $p_i = e^{\Sigma\beta_k x_{ki}}/(1+e^{\Sigma\beta_k x_{ki}})$.

2.9 Model Results

2.9.1 Opening Rate Equation

The estimates for this equation are contained in Table 2.2.

Impact of Subject-Line Variables. I find that moderate subject-line length generates higher opening rates than short or long subject-line length (consistent with H1) because the linear term associated with the subject-line-length measure has a significant positive coefficient (.993) and the quadratic term has a significant negative coefficient (-0.174). If the appeal type is generic, I can approximate the subject-line length that maximizes the opening rate as roughly 17.4 letters (*i.e.*, 17.4 = exp(.993/(2*0.174)), recognizing that the subject-line-length variable is the natural log of the number of letters in the subject line.[25] Therefore, the optimal subject-line length consists of perhaps three or four words, which makes intuitive sense.

[25] If I factor in the interaction terms, the optimum subject lines range from 14.1 letters (for tournaments) to 29.4 letters (for new products), but this is a very approximate range since four of the interactions are not significant. For details of the optimizations, see Appendix 1. Note that the optimizations use the fact that the dependent variable (*e.g.*, opening rate) is an increasing function of the linear predictor, $\eta = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$.

Table 2.2 PARAMETER ESTIMATES OF MODEL FOR OPENING RATE [a]

Independent Variables:	*Parameter Estimates:* [b]
Intercept	-4.257***
Subject-line Variables	
Number of Letters (Subject-line Length)	.993***
Quadratic Term of Number of Letters	-.174***
Area Name Mentioned	-.027
Exact Date Mentioned	.030*
Appeal Type: Store Opening	-.438
New Product	-.713
Price Promotion	-.159
Training Program	-.532*
Tournament	.217
Interactions: Number of Letters × Store Opening	.097
Number of Letters × New Product	.183
Number of Letters × Price Promotion	.045
Number of Letters × Training Program	.154*
Number of Letters × Tournament	-.071
Email Campaign Variables	
Number of Areas (Audience Size)	-.014*
U.S. Only	-.586***
Both Countries	.067**
Workday	.048
Time Since the First Email (Age of the System)	3.167***
Quadratic Term of Time Since the First Email	-.796***
Null Deviance (644 degrees of freedom)[c]	235,414
Residual Deviance (624 degrees of freedom)[c]	53,266

[a] The expected value of the response (y_i) is given by $y_i = \frac{e^{\Sigma x_i \beta_i}}{1+e^{\Sigma x_i \beta_i}}$.

[b] *** is significant at the .001 level; ** is .01 level; * is .05 level.

[c] The deviance is twice the log-likelihood ratio of testing a model (either the null model with only an intercept or the full model) against an unrestricted model. By analogy to R-squared, I could define quasi-R-squared as 1 – (residual deviance/null deviance); it would be .774 for this model. I hasten to add that I provide this only for heuristic evaluation of the model, since the formal properties of R-squared and related test statistics do not hold or only hold as approximations.

This contradicts the principle of advertising copy design that the shorter the headline, the greater the recognition of the advertisement (discussed in Ramond's 1976 exhaustive review of advertising copy research). I think this is so because, email is unlike traditional print advertisements for which the headline, graphics, and copy layout create a unified visual impression that be scanned virtually simultaneously. When deciding whether to open an email the subject line is usually one of two main things upon which the consumer must rely.[26] Consequently, the subject line should provide more of a self-contained inducement to go on (balancing the option cost and expected upside) than the headline of a traditional print advertisement would require.[27]

I note, next, that the coefficient on "area name mentioned" is not significant, while the coefficient on "exact date mentioned" is significantly positive at 0.05 level. The lack of the former effect may arise because the area mentioned is desirable for some customers but undesirable for others and this replicates a similar finding of Marinova, Murphy and Massey (2002).[28] On the other hand, mentioning the exact date of an event leads to a more clear-cut response from consumers, consistent with Godin's (1999) argument that permission marketing needs to be "anticipated, personal, and relevant." From the perspective of options theory, the value of time specificity is that the cognitive cost of processing the message is smaller and the expected upside benefit (to some consumers) is larger.

---

[26] Some email systems provide a preview or thumbnail of the email content, but the majority of users do not use systems that provide thumbnails. According to a recent poll (Lyris Technologies 2005), 64 percent of (141) consumers surveyed use a web-based email account like Hotmail or Yahoo to check email, and these systems do not provide previews or thumbnails. This number has grown from previous years. About 36 percent use software such as Outlook or Eudora, which does provide previews or thumbnails. For these, the user must select the email to display the preview or thumbnail, but not all users configure the system to show the thumbnails or previews, although most probably do.

[27] The other bit of information upon which users rely to decide whether to open emails is the sender's name/email address; this would be an important design factor to analyze, but there is no variation in that variable in this data set.

[28] In an experimental setting, they find that subject-line relevance does not significantly affect email opening rates or opt-out rates. They attribute this outcome to the experimental stimuli and to the confounding effect of seasons.

Table 2.2 also shows that some subject-line appeal types generate different opening rates than others (supporting H2), which is consistent with industry reports and related research.[29] In particular, "Training Programs" have a significant negative direct effect on the opening rate (at the .05 level). There is also a largely countervailing positive interaction of "Training Programs" with the subject-line length (significant at the .05 level).[30] Overall, variation in opening rate and optimal subject line length appears to arise because different subject-line appeals convey different expected upside potential.

Impact of Email Campaign Variables. Table 2.2 also shows significant coefficients for several of the email campaign variables. I defer this discussion to a later section.

Model Fit. A measure of model fit is provided by noting that the model reduces the null deviance of 235,414 (which arises when $\eta_{1i} = \beta_{10}$ is estimated as a single parameter) to a residual deviance of 53,266 (which arises when the model of Table 2.2 is estimated). The deviance from a model is twice the log-likelihood ratio of testing a model against an unrestricted model. The associated F-test is very significant suggesting that the full model is superior to the null model.

---

[29]For example, when DoubleClick (2005) asked respondents in a survey "what type of subject lines compel you to open a permission email," 65% chose discount offer, 42% chose free shipping, 39% chose new product announcement, and 34% chose compelling info/news. EmailLabs (2006) reports that female consumers are more likely to open retail email with subject lines that emphasize saving money: 72.5% of respondents say they are motivated by "discounted price" and 60.1% are motivated by a "free shipping offer." In related work, Lam *et al.* (2001) find that price promotions are more effective than new product promotions in increasing store traffic. Chittenden and Rettie (2003) also find that a strong incentive described in the subject line and a targeted recipient list can improve the response rate (measured as the click-through rate) and reduce respondents' chances of "unsubscribing."

[30]For training program, the optimal subject line can be calculated to be 27.0 letters (using the formula from Appendix 1). Thus, communicating about a training program requires a longer subject line.

2.9.2 Click-through and Opt-out Equations

The estimates for the click-through and opt-out equations are contained in Table 2.3.

Recursive Structure. Table 2.3 shows that the opening rate does not significantly influence the click-through or opt-out rates, and the click-through rate does not significantly influence the opt-out rate either. I also estimated this model without the recursive terms, and the other estimates were nearly the same. (This suggests that any possible simultaneity bias from other left-out variables is not too large a concern.) While the recursive terms are not significant, I find this model conceptually closer to the framework in Figure 2.2, so I report this version.

Impact of Email Body Variables. I find the linear term associated with the effect of email-length on the click-through rate has a significant negative coefficient (-1.117), and the quadratic term has a significant positive coefficient (.322). For a generic appeal type, I can calculate that the email-length that *minimizes* the click-through rate is roughly 1.7 scrolls or screens (*i.e.*, 1.73 = (1.117/(2*.322)). However, since all emails in the database were at least one scroll long and very few were larger than 3.0 scrolls, this means that the ideal email length was bimodal. When I consider the interaction terms with appeal type, I indeed find that four of the five appeal types reach a minimum click-through rate somewhere near 2.5 scrolls, so for these four appeal types, from the range of 0 to 2.5 scrolls, longer email body reduces the click-through rate (partially consistent with H4). However, for new product announcements, email body longer than one scroll does generate greater click-through (not consistent with H4).[31] I also find that longer email body raises the opt-out rate in a diminishing fashion (consistent with H4). Here, the linear term associated with the effect of email-length on the opt-out rate has a significant positive coefficient (.654) and the quadratic term has a significant negative coefficient (-.056). In this case, for a generic appeal type the email-length that maximizes the opt-out rate is 5.84 (*i.e.*, 5.84 = (.654/(2*.056)), which is greater than all emails actually sent out. Thus, over the relevant range, larger email size leads to larger opt-out rates, but with a diminishing effect.

The findings regarding email length are consistent with established advertising literature on ad clutter. Elliot and Speck (1998) define "perceived ad clutter" as a consumer's conviction that the amount of advertising in a medium is excessive. Cho and Cheon (2004) propose that ad clutter on the Internet could be operationalized as the number of banners ads, pop-up ads, advertorials, text links, etc. that appear on a single Web page. They find that people avoid advertising messages on the Internet because of perceived ad clutter. Here, I find that in permission email design, a long email body could also lead to ad clutter for most appeal types (except for new product announcements). Chittenden and Rettie (2003) conducted one of the few studies of this issue for emails. They similarly find that short email body generates higher click-through than long email body. They also find that long email body generates higher opt-out (unsubscribe) rates. My findings are consistent with theirs, and my interpretation treats email length and ad clutter as part of the cognitive fixed costs of the option.

---

[31] Factoring in the interaction effects for appeal types, the minimum point was reached at 2.26 for store openings, 1.01 for new products, 2.46 for price promotions, 2.59 for training programs, and 2.41 for tournaments.

Table 2.3 PARAMETER ESTIMATES OF MODELS FOR CLICK-THROUGH RATE AND OPT-OUT RATE [a]

Independent Variables:		*Parameter Estimates:* [b]	
		Click-through Rate	*Opt-out Rate*
Intercept		-4.185***	-620.9***
Opening Rate		.416	.292
Click-through Rate			-.714
Email Body Variables			
Number of Scrolls (Email Length)		-1.117***	.654***
Quadratic Term of Number of Scrolls		.322***	-.056
Number of Links		1.579***	.415***
Quadratic Term of Number of Links		-.307***	-.051*
Appeal Type:	Store Opening	.640	-.003
	New Product	-2.436*	1.181***
	Price Promotion	.073	.511***
	Training Program	.141	1.334**
	Tournament	.160	.574**
Interactions:	Number of Scrolls × Store Opening	-.336	.403
	Number of Scrolls × New Product	.465	-.600***
	Number of Scrolls × Price Promotion	-.465***	-.257***
	Number of Scrolls × Training Program	-.549*	-.475*
	Number of Scrolls × Tournament	-.432**	-.170*
	Number of Links × Store Opening	-.495	-.241
	Number of Links × New Product	1.070***	.300***
	Number of Links × Price Promotion	.265*	-.197***
	Number of Links × Training Program	.709***	-.261**
	Number of Links × Tournament	.265	-.165*
	Lead Time × Store Opening	.305	-.216
	Lead Time × New Product	-.549*	-.123*
	Lead Time × Price Promotion	.280***	.066*
	Lead Time × Training Program	.339**	-.095
	Lead Time × Tournament	.450***	.036
Email Campaign Variables			
Number of Areas (Audience Size)		.093*	.023
U.S. Only		.161	.374*
Both Countries		.463**	.471***
Lead Time		-.501***	-.061
Workday		1.213***	-.406***
Time Since the First Email (Age of the System)		.656	418.1***
Quadratic Term of Time Since the First Email		-.313	-71.05***
Null Deviance (643 degrees of freedom) [c]		237,810	17,806
Residual Deviance (611 and 610 d.o.f.) [c]		65,868	2,006.7

[a] The expected value of the response (y_i) is given by $y_i = \dfrac{e^{\Sigma x_i \beta_i}}{1+e^{\Sigma x_i \beta_i}}$.

[b] *** is significant at the .001 level; ** is .01 level; and * is .05 level.

[c] If I define quasi-R-squared as 1 – (residual deviance/null deviance); it would be .723, and .887 for this table.

Concerning the influence of number of links on consumers' clicking-through and opting-out, I see two effects that work in opposite directions. On one hand, more links increase the click-through rate (consistent with H3). On the other hand, more links increase the opt-out rate (not consistent with H3). In particular, the linear term associated with the effect of number of links on the click-through rate has a significant positive coefficient (1.579) and the quadratic term has a significant negative coefficient (-.307). From these numbers, I can calculate that the optimal number of links for a generic appeal type was 13.1 (*i.e.*, 13.09 = exp(1.579/(2*.307)). Similarly, the linear term associated with the effect of number of links on the opt-out rate has a significant positive coefficient (.415) and the quadratic term has a significant negative coefficient (-.051). From these numbers, the number of links that generates the maximum opt-out rate for a generic appeal type was 58.5 (*i.e.*, 58.48 = exp(.415/(2*.051)). The data indicate that the actual number of links was always at least one and rarely more than eleven. So, overall, I find that that more email links yield both higher click-through and opt-out rates, with diminishing effects. On balance, adding links leads to a short term benefit of more click-through, but too many links can negative long-term repercussions in the form of a smaller subscriber list.

These findings differ from the hypotheses proposed by Ansari and Mela (2003). They propose that increases in the number of links in an email exacerbate clutter and therefore increase the cognitive cost of perusing the email. In their study, however, they fail to find a significant negative effect of number of links on click-through rates. They attribute the failure to the small number of links (never more than eight). Chittenden and Rettie (2003) also do not find a significant difference in click-through rates for emails with a greater number of links compared with emails with fewer links. In the present study, adding links while controlling for a fixed length of email appears to be a desirable thing, which can be thought of in the options context as enhancing the upside potential.[32]

Table 2.3 also shows that different appeal types generate significant main effects on click-through rates and opt-out rates (particularly the latter) and a variety of significant interactions with email length (number of scrolls) and number of links. These effects are somewhat time-consuming to interpret, but very interesting for the retailer involved. For example, store openings do not have significant main or interaction effects; and, generally, the appeal is benign—mostly moderately positive, but small. New products have a large negative main effect, but this can be cancelled out by the significant interaction effect for number of links.[33] Thus, if the retailer is going to announce new products, it should list several examples concretely, each with a separate link. Listing six or seven products is probably fine: the upside expectation rises because there is a greater chance of a match with the customer's preference. But if one goes too much above this, the opt-out rates begin to rise. Price promotions, training programs, and tournaments all have no significant main effect for appeal type, significant negative interactions with number of scrolls and positive interactions with lead time. For these activities, it pays to give some advance notice and to

---

[32]The past studies have not looked at the impact of the number of links on the opt-out rate. I find a positive relationship in this data set.

[33] If the firm uses about 10 links, the positive impact is 1.07*ln(10), which outweighs the -2.436 main effect.

avoid making the consumer scroll too many times. For price promotions and training programs, there are also significant positive effects for adding more links. Again providing links to six or seven promoted items or to particular listed courses raises the upside because there is a greater chance of a customer match, and these can fit comfortably into emails without much raising the email length. I asked management about this, and they indicated that, when they first rolled out this campaign, they did not include very many links (often each email included only one link). With experience, they later realized that for several email types, it is advantageous to include more links (without making the length of the emails themselves much longer).

Model Fit. Overall, fit of the model is indicated by noting that the model reduces the null deviance of the click-through and opt out equations, respectively, from 237,810 and 17,806.4, to a residual deviance, respectively, of 65,868 and 2,006.7. The associated F-tests are very significant.

Impact of Email Campaign Variables (all three equations). As shown in the conceptual model (Figure 2.2), email campaign variables influence decisions to open email, to click-through on links, or to opt-out – all three of the dependent variables. I now turn to the results for the opening-rate model in Table 2.2 and for the click-through-rate and opt-out-rate models in Table 2.3.

First, I find that longer lead time has a significant negative effect on the click-through rate (consistent with H6). Intuitively, longer lead times may raise the cognitive costs (to store, plan, and remember or retrieve the information). However, lead time does not significantly affect the opt-out rate. In the click-through and opt-out models, I also checked the quadratic term of lead time, but it was not significant. So I conclude that the relationship between lead time and the click-through rate is linear.

Second, I find that the number of areas being targeted has a small, significant negative coefficient (-.014) in the opening rate equation, but a larger (but still small) significant positive coefficient (.093) in the click-through rate equation. Being more targeted, thus, only appears beneficial for increasing the opening rate (partially consistent with H7).

Third, I find that U.S. subscribers are significantly less likely to open email than Canadians (the "U.S. Only" coefficient is -.586 and very significant) and somewhat more likely to opt-out of the email list (the "U.S. Only" coefficient is .374 and significant). I suspect that the reason for this outcome is that the focal retailer has more brand equity in Canada than in the U.S., but it is possible that U.S. subscribers have less patience with permission email for various other reasons.

Fourth, I find that email opening behavior is not significantly different for weekdays or week-ends. An explanation of this is that, for modern consumers, increased integration of personal and professional lives tends to blur the boundary between work time and rest time. DoubleClick's (2005) survey finds that email is fully integrated into the personal and professional lives of most consumers. This integration has facilitated merging of personal and professional time with most consumers checking work email from home and checking personal email while at the office. In the survey, over 40% consumers report checking their work email account at home both in the evenings and on weekends and they do this "all the time" or "frequently". And over 50% of consumers check their personal email accounts from work at least "occasionally." This constant checking means that there is no perfect day or time for deploying email campaigns.

However, the study finds that consumers are significantly more likely to click-through and significantly less likely to opt-out on workdays. On workdays, consumers seem to have kinder inclinations toward emails, suggesting that they less like having email encroach on the week-end leisure time. It is worth mentioning that in a separate analysis, I but did not find any significant differences across the five weekdays.

Lastly, I find that the effectiveness of permission email appears to decrease with the aging of the system. For all three models, the linear term is positive and the quadratic term is negative (and significant for the first and third models). I calculated that these coefficients are associated with curves that attain a maximum in 7.3 days, 2.9 days, and 19.0 days, respectively. Since the earliest email in the database was sent out after 11 days (I excluded the first five test emails from the sample), the model finds decreasing rates for the first two equations, and decreasing rates after 19 days into the system for the third equation. One important thing that I do not know with certainty is whether the decline in the opening rate was due to falling interest in this retailer's emails or due to a general reaction to world-wide increases in spam during this period. I do note that the opt-out rate also decreased, and if spam were the primary influencer, I would have observed the opposite result. On balance, it seems that those subscribers who do not quickly remove themselves from the list stay on the list for a long time. They may still open and click-through these emails, but once the novelty wears off, they become more discriminating when they do so.

These findings are consistent with previous research and industry reports. eMarketer (2006) argues that the dynamic of the Internet and its population means that any information collection mode about Internet users ages extremely quickly. DuFrene *et al.* (2005) even argue that the growth of permission email may eventually cause consumers to perceive it as spam. As the feeling of novelty goes away and subscribers get bored with these emails, the "engagement" deteriorates, and the likelihood of response may diminish. Industry reports also observe this phenomenon: DoubleClick (2005) reports that in the U.S. the average opening rate for permission email declined from 38.2% in the first quarter of 2004 to 30.2% in the first quarter of 2005, and the average click-through rate declined from 8.4% to 7.9% during the same period.

2.10 Conclusions, Limitations, and Future Research

Overall, the present study is intended to contribute two things. First, I illustrate a relatively accessible approach to measuring how email effectiveness is influenced by design and campaign variables. Second, I illustrate how a decision-theoretic analysis, interpretable as an options design problem, can be used to consider permission email design.

The value of the first contribution is that it is relatively easy for small and medium-sized firms to analyze determinants of effectiveness of their permission email campaigns. Firms engaged in permission marketing already have available (or can easy obtain) data on the dependent variables (opening, click-through, and opt-out rates) at the level of the individual broadcast email. They also have copies of the emails they already set out. So the only additional information required can be obtained by measuring the design attributes present in those emails that were sent out (*e.g.*, counting the number of characters, lines, and links; analyzing possible visual dimensions; classifying the appeal type based on objective criteria; and compiling the time, date, and other distribution statistics of those emails sent out). If this is done on an on-going basis, very little marginal effort is required—even going back through a yearlong history of emails involves only a few days of staff time for one person. Also, the software used for estimation of the generalized linear model is globally available freeware (and simple application of regressions in spreadsheet software also yields tolerable approximations). For example, with this approach, the retailer of this study can learn that it is desirable to use emails that have subject lines of about three or four words (depending on the nature of the appeal), mention the exact date, include enough links to increase click-through (but not so many as to lead to high opting-out), make the email body short, use relatively short lead times, distribute on weekdays, and vary the design across different appeal types. Probably the greatest challenge for permission email is aging of the email system and keeping the communications fresh.

One natural question to the findings in the present study is whether these findings can be generalized to all emails (including spam). Although I am confident that some findings could apply to general email, consumers may respond differently to unsolicited email. The extent of email proliferation has grown to the level where consumers receive too many emails to peruse all of them, so consumers are reluctant to open email from unknown senders (eMarketer 2006). A well-designed subject line of spam may be attractive to some consumers, but the fact that the email is from an unknown sender may scare away most consumers. I therefore strongly recommend that retailers should send out email to customers who have already given their permission to the retailers. Although it takes time and money to build a permission email list and create targeted, relevant offers and messages, permission email has significant advantages over spam (see Footnote 1). Besides, anti-spam laws have made it more and more difficult to send spam.

The value of the second intended contribution is that it helps us to think about the design attributes of email campaigns in an intuitive way applying simple options models whereby consumers balance the cognitive costs of attending, processing, remembering, and acting on an email against the expected upside benefits. This helped frame my thinking and hypotheses about the design variables, and most hypotheses were confirmed.

While my model describes sequential search, (and builds on that literature), the options interpretation that I suggest differs in perspective from the sequential search literature in two ways. First, the approach looks at the problem from the perspective of a firm sending out an email (or offering an option), potentially to heterogeneous consumers. Such an approach gives insight into the risk profile associated with particular design attributes of a communication.

Second, the approach explicitly recognizes that the initiative lies with the firm that sends out the email option (rather than supposing that the firm is facilitating an ongoing search process by the consumer).

For future consideration, I also think that the options interpretation has applicability in other domains relating to marketing communications and promotions. I provide some examples in Table 2.4. In each example, there are fixed costs of attending to, processing, remembering, and acting on the communication or promotion, and the consumer chooses to do so based on the expected upside potential. Also for these examples, the purchase likelihood is increased by selecting marketing techniques or design features that lower the options price, raise the expected value of the underlying offer, or raise the expected variance of the underlying offer.

From a behavioral perspective, thinking through these kinds of examples suggests the need for research into the cognitive processing of the consumer. A distinction may be made between the types of cognition required for phase 1 processing (attending sufficiently to decide to acquire the option) and for phase 2 processing (deciding whether to exercise the option).

Phase 1 processing is influenced by a tradeoff between cognitive effort and the upside potential, as determined by the prevailing expectations frame.

Table 2.4 POSSIBLE APPLICATION AREAS FOR THE OPTIONS INTERPRETATION

Consumer applications areas	*Deciding whether to go into a store when passing by; then, upon entering, deciding whether to exercise the option to buy.* *Choosing to examine a coupon; then deciding whether to buy the product with the coupon.* *Attending to a television or print ad; examining the product; then deciding whether to buy.*[34] *Trying a free product for a period of time (or buying a product with a return policy); then deciding whether to keep the product.* *Listening to a sales pitch; then deciding whether to take the deal.*
Marketing techniques for lowering the implicit option price	*Simplify cognitive processing by clearly articulating offer features.* *Teach consumers queues or vocabulary (e.g., "midnight madness sale") that efficiently evokes meaning about the nature of the offer to make future communications more efficient.* *Signal the simplicity of processing (e.g., through white space in print media and non-competing fonts).* *Make the appeal actionable and avoid making people wait for a long time or use much memory.* *Provide a gift contingent on acquiring the option.*[35]
Marketing techniques for raising the expected value of the underlying distribution of benefits or the variance	*Make the upside benefits as clear as possible.* *Teach consumers what brands mean to help set in consumers' minds a high expected value and low variance of the offer.* *Provide a companion benefit (which has no uncertainty) such as a rebate to shift the underlying distribution to the right.* *Offer free samples to help the consumer learn the value of the offer.*
Marketing techniques for compound offers in series.	*Avoid frontloading sequential options with the costs before conveying sufficient information to raise the expected upside.*
Marketing techniques for compound offers in parallel.	*Provide a menu of offers or features that appeal to different kinds of people to increase the expected value.*[36]

[34] This is analogous to opening an email, clicking on a link in the email body, and then deciding whether to buy.

[35] For example, some timeshare condominiums provide a free week-end hotel stay if customers listen to a one-hour pitch.

It would be valuable to understand how expectations are influenced by various signals and queues implied by the design variables, by the general environment, by competing, complementary, or bundled products, or by situational factions facing the consumer at the time of decision. Phase 2 processing leads to more final and consequential outcomes, and in this phase there may be heterogeneity in consumer processing and time required to "pull the trigger." There also may be considerations of sunk costs (e.g., Arkes and Blumer 1985 and Kogut 1990) and escalation of commitment (Brockner 1992).

There are some limitations of this study and possible directions for further work.

First, the data used in the current study are aggregated at the email level and I cannot track individual recipient's responses to the email. Following Ansari and Mela (2003), with individual level data, it would be desirable to study the impact of consumer demographics as they progress through the email response process (described in Figure 2.2 of the current paper). Such data would enable consideration of several hypotheses:

(1) Higher income may lead to greater valuation of time, greater perceived fixed cognitive costs associated with subject lines and long email length, and therefore lower opening and click-through rates (*i.e.*, there is an interaction term between subject-line length and income in the opening rate equation and an interaction between email-body length and income in the click-through rate equations). Thus, it may be the

[36] Vagueness (or bravado) in the offer may raise variance, but this myopic approach may lead to disappointed customers who will lower their expectations and avoid contact with the vender (analogous to opting out of a list).

case that email has some of the same features as coupons as a mechanism for price discrimination (similar to Narasimhan 1984)

(2) There may be gender differences in the perceived fixed costs of opening emails, clicking on links, where, generally, there might be utility gained from the "shopping experience" of browsing through offers for one gender and not the other. This permits examination of questions similar to those of Phillip and Suri's (2004) who examine the impact of gender differences on the evaluation of promotional email.

(3) Recipients' response to permission email may depend on the time the last email was opened by a consumer. Measuring this could help optimize the time interval between sending out emails.

Generally, examination of disaggregate data would help construct more effective email lists and enable greater customization of email. Nevertheless, having noted the desirability of examination of such questions with disaggregated data, this study demonstrates that interesting conclusions can also be obtained from aggregate data of the type available to the retailer associated with the current study. There is a place for both disaggregate and aggregate analysis, depending on whether one needs conclusions at the level of the individual consumer or at the level of the individual broadcast email, and depending on the firm's desired investment in time, effort, and cost.

Second, I do not have experimental control over the independent variables (but, instead, rely on the variation used by the firm in carrying out its business). While this provides for strong external validity, experiments and field studies could be done that build in systematic variation in the subject line, the email body, and the email campaign characteristics (along the lines of Vriens *et al.'s* (1998) conjoint experiment of direct mail response). Controlled experiments also could help us understand the influence of recipients' situational factors[37] and the use of personalized salutations.[38] In addition, it would also be interesting to look at whether email list subscribers' attitudes change over time.[39] Perhaps most importantly, because I do not have experimental control, while most of the hypotheses are confirmed and the associated intuitive rationales are reasonable, it is not possible to claim that the options approach is the only theory that can account for the empirical results. This suggests the need for future comparative empirical and/or experimental studies and subsequent syntheses.[40]

Third, other dependent and independent variables may be of interest. The dependent variables examined in this study are the opening rate, the click-through rate, and the opt-out rate. It would also be valuable to analyze traffic, sales, and profit data, both on-line and in-store, to facilitate examination of the impact of permission email campaigns across channels. It is worth mentioning that sales data, in particular, could help us complete examinations of the option

---

[37] This would follow the lines of experiments about situational factors done by Li and Bukovac (1999), who study the cognitive impact of banner ad characteristics, and by Cho (2003), who studies the influence of consumer involvement on banner ad click-through behavior.

[38] This builds on Heerwegh (2005), who finds that personalized invitation email significantly increases the response rates to web-based surveys, and on Roy and Berger (2005), who examine factors that increase response rates to email and mixed-mode surveys.

[39] DuFrene *et al.* (2005) observe that student subjects' attitudes change over an eight-week period, but a study with industry data could observe attitude changes over a much longer period.

[40] The options approach is already an extension of the cost-benefit approach, but it would seem natural to think about how situational factors influence the perceived option cost and the expected upside potential in a way that embraces the points raised regarding theories of interruption avoidance in the literature on permission marketing (Godin 1999).

process described in Figure 2.5, which illustrates two stages of a simple option structure: purchasing an option and exercising an option. Due to data limitations, the present study only looks at whether consumers will purchase the option. Sales data could give us an opportunity to examine consumers' option exercising behavior. In addition, measures of brand equity, such as traditional measures of aided and unaided brand recall, may be of interest because they describe the extent to which consumers read emails (this goes beyond the information provided by the opening rate used in this study) and because, in the long run, these emails could help increase awareness and remind consumers of the existence of the firm and the brand.[41] To analyze such factors, it would be desirable to analyze longitudinal individual choice data on-line and in-stores to examine the effectiveness of permission email in enhancing brand equity and stimulating purchase. In addition, it might be desirable to consider recipients' willingness to forward permission email as a further measure of email effectiveness.[42] Lastly, it would be desirable to consider other measures of communications message, layout, reach, and timing that go beyond the limitations of my own dataset. An obvious candidate would be to examine the influence of the sender's name/address on the opening rate (for which there was no variation in the data). Because of the importance of the topic, I encourage such extensions.

[41] This point is made clearly by Dreze and Hussherr (2003) in the context of an analysis of the click-through rate for banner ads. In a context closer to the current paper, Merisavo and Raulas (2004) use a survey to examine the impact of email marketing on brand loyalty. They find that regular email marketing has positive effects on brand loyalty, that consumers visit retail stores because of emails, and that consumer exposed to email marketing recommend the brand to their friends. Rettie *et al.* (2005) come to similar conclusions in the context of a similar form of telemarketing – short message service (SMS). They find that SMS advertising is effective in improving both consumers' brand attitude and purchase intentions.
[42] Building on research on electronic word-of-mouth (Phelps *et al.* 2004), one could also examine recipients' motivation to pass along permission email.

In conclusion, as Weber (2004) indicates, email will only become more prominent in our lives, and we need to harness its value and avoid its pitfalls. My hope is that continued study of permission marketing will make it a “win-win” for all concerned – that consumers get desired information that they choose to receive and that firms learn to use the medium in an effective and respectful way to help build relationships with consumers.

Appendix 2.1
Calculations of Optimal Levels of the Independent Variables

In this appendix, I describe my calculations of the optimal levels of independent variables with linear and quadratic terms. Although these calculations are straightforward, I describe them for completeness. I use the opening rate equation (based on the variables in Table 2.2 as an example. Analogous calculations were done for the click-through and opt-out rate equations.

To optimize the opening rate p_1, I need to also optimize the linear predictor η_1 because the link function $p_1 = \frac{e^{\eta_1}}{1+e^{\eta_1}}$ is an increasing function (to simplify the notation, I drop the subscript for the i^{th} email). Table 2.2 describes the 13 independent variables in the opening rate model, which I repeat below:

x_1: ln(Number of Letters)
x_2: Area Name Mentioned
x_3: Exact Date Mentioned
x_4: Store Opening
x_5: New Product
x_6: Price Promotion
x_7: Training Program
x_8: Tournament
x_9: ln(Number of Areas)
[Describes audience size]
x_{10}: U.S. Only
x_{11}: Both Countries
x_{12}: Workday
x_{13}: ln(Time Since the
First Email)

The linear predictor, in particular, is given in the model by,

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \beta_3 x_2 + \beta_4 x_3 + \beta_5 x_4 + \beta_6 x_5 + \beta_7 x_6 + \beta_8 x_7 + \beta_9 x_8 + \beta_{10} x_1 x_4 + \beta_{11} x_1 x_5 + \beta_{12} x_1 x_6 + \beta_{13} x_1 x_7 + \beta_{14} x_1 x_8 + \beta_{15} x_9 + \beta_{16} x_{10} + \beta_{17} x_{11} + \beta_{18} x_{12} + \beta_{19} x_{13} + \beta_{20} x_{13}^2$$

To optimize η based on x_k, k=1, 2, ..., 13, I set $\frac{\partial \eta}{\partial x_k} = 0$. For example to optimize η based on x_1 (ln(Number of Letters)), I set $\frac{\partial \eta}{\partial x_1} = 0$. This implies

$$x_1 = \frac{\beta_1 + \beta_{10} x_4 + \beta_{11} x_5 + \beta_{12} x_6 + \beta_{13} x_7 + \beta_{14} x_8}{-2\beta_2}$$

When $\frac{\partial^2 \eta}{\partial x_1^2} = 2\beta_2$ is negative (positive), this value is a maximum (minimum).

Now, since the particular variable x_1 is log transformed, I have

$$\text{Optimal Number of Letters} = \exp\left(\frac{\beta_1 + \beta_{10} x_4 + \beta_{11} x_5 + \beta_{12} x_6 + \beta_{13} x_7 + \beta_{14} x_8}{-2\beta_2}\right)$$

For a particular appeal type, I can calculate the optimal value for x_1 as follows. For example, if the appeal is a "Store Opening," $x_4 = 1$ and $x_5 = x_6 = x_7 = x_8 = 0$, I have

$$\text{Optimal Number of Letters} = \exp\left(\frac{\beta_1 + \beta_{10}}{-2\beta_2}\right).$$

If the appeal type is generic, I have $x_4 = x_5 = x_6 = x_7 = x_8 = 0$, and

$$\text{Optimal Number of Letters} = \exp\left(\frac{\beta_1}{-2\beta_2}\right).$$

This describes my calculations of the values for the independent variables that optimize the opening rate. Analogous optimizations are done for click-through and opt-out rates.

Bibliography

Ansari, A., & Mela, C. F. (2003). E-customization. *Journal of Marketing Research,* 40 (2), 131-145.

Arkes, Hal R and Catherine Blumer (1985). The Psychology of Sunk Cost, *Organizational Behavior and Human Decision Processes*, 35 (1), 124-140.

Bakos J. (1997). Reducing Buyer Search Cost: Implications for Electronic Marketplaces, *Management Science*, 43 (2), 1676-1692.

Black, F. and Scholes, M. S. (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81 (3), 637-654.

Brockner, Joel (1992). The Escalation of Commitment to a Failing Course of Action: Toward Theoretical Progress. *Academy of Management Review*, 17 (January), 39-61.

Chittenden, L. and Rettie, R. (2003). An Evaluation of E-mail Marketing and Factors Affecting Response. *Journal of Targeting, Measurement and Analysis for Marketing, 11*(3), 203.

Cho, C. and Cheon, H. J. (2004). Why Do People Avoid Advertising on the Internet? *Journal of Advertising, 33*(4), 89.

Cho, C. (2003). The Effectiveness of Banner Advertisements: Involvement and Click-through. *Journalism and Mass Communication Quarterly.* 80 (3), 623-645.

Cox, J. C.; Ross, A.; and Rubenstein, M. (1979). Option Pricing: A Simplified Approach. *Journal of Financial Economics*, 7 (3), 229-263.

DoubleClick (2005). Sixth Annual Consumer Email Study. December, www.doubleclick.com.

Dreze, X., and Hussherr, F. (2003). Internet Advertising: Is Anybody Watching? *Journal of Interactive Marketing, 17*(4), 8.

DuFrene, Debbiw D., Brian T. Engelland, Carol M. Lehman, Rodney A. Pearson (2005). Changes in Consumer Attitudes Resulting from Participation in a Permission E-mail Campaign. *Journal of Current Issues and Research in Advertising*, 27 (1), 65-77.

Elliot, Michael T., and Paul S. Speck (1998). Consumer Perceptions of Advertising Clutter and Its Impact Across Various Media. *Journal of Advertising Research*, 38 (January/February), 29-41.

EmailLabs (2006). Email Marketing Statistics and Metrics. http://www.emaillabs.com/resources_statistics.html#holidayemails, April 3, 2006.

eMarketer (2004). The eMail Marketing Report, October, http://www.emarketer.com/.

eMarketer White Paper (2006). Effective E-Mail: The Seven Golden Rules You Know (But May Forget to Follow). Free download from http://www.emarketer.com/ on March 29, 2006.

Gefen, David and Detmar W. Straub (1997). Gender Differences in the Perception and Use of Email: An Extension of the Technology Acceptance Model. *MIS Quarterly*, 21 (4), 389-400.

Godin, S. (1999). *Permission Marketing: Turning Strangers into Friends, and Friends into Customers*. New York: Simon & Schuster.

Gopal, Ram D., Zhiping Walter, and Arvind K. Tripathi (2001). Admediation: New Horizons in Effective Email Advertising. *Communications of the ACM*, 44 (12), 91-96.

Heerwegh, D. (2005). Effects of Personal Salutations in Email Invitations to Participate in a Web Survey. *Public Opinion Quarterly*, 69 (4), 588-598.

Hoffman, J. (2000). An Anatomy of E-mail Marketing. *Direct Marketing, 63*(3), 38-39.

Hoffman, D. L., Novak, T. P., and Venkatesh, A. (2004). Has the Internet Become Indispensable? *Communications of the ACM, 47*(7), 37-42.

Johnson, E. J., Bellman, S., and Lohse G. L. (2002). Defaults, Framing and Privacy: Why Opting In-Opting Out. *Marketing Letters*, 13 (1), 5-15.

Kent, R. and Brandal H. (2003). Improving Email Response in a Permission Marketing Context. *International Journal of Marketing Research*, 45 (4), 489-503.

Kogut, C. A. (1990). Consumer Search Behavior and Sunk Cost. *Journal of Economic Behavior and Organization*, 14 (3), 381-392.

Kraut; Robert; Tridas Mukhopadhyay; Janusz Szczypula; Sara Kiesler; and Bill Scherlis (1999). Information and communication: Alternative uses of the Internet in households. *Information Systems Research*; 10 (4), 287-303.

Krugman, Herbett E. (1983). Television Program Interest and Commercial Interruption: Are Commercials on Interesting Programs Less Effective? *Journal of Advertising Research*, 23 (February/March), 21-23.

Lam, Shun Yin, Mark Vandenbosch, John Hulland, Michael Pearce (2001). Evaluating Promotions in Shopping Environments: Decomposing Sales Response into Attraction, Conversion, and Spending Effects. *Marketing Science*, 20 (2), 194-215.

Lee, Allen (1994). Electronic Mail as a Medium for Rich Communication: An Empirical Investigation Using Hermeneutic Interpretation. *MIS Quarterly,* 18 (2), 143-157.

Li, H., & Bukovac, J. L. (1999). Cognitive Impact of Banner Ad Characteristics: An Experimental Study. *Journalism & Mass Communication Quarterly, 76*(2), 341-353.

Loder, Thede, Marshall Van Alstyne, and Rick Wash (2004). An Economic Answer to Unsolicited Communication. *Proceedings of the 5th ACM conference on Electronic Commerce*, (May), 40 – 50.

Lynch, J. G. and Ariely, D. (2000). Wine Online: Search Costs Affect Competition on Price, Quality, and Distribution. *Marketing Science*, 19 (1), 83-103.

Lyris Technologies (2005). More Consumers Using Web Based Email. http://www.lyris.com/news/pr/pr-111705.html, Nov. 17, 2005.

Marinova, A., Murphy, J., and Massey, B. L. (2002). Permission E-mail Marketing as a Means of Targeted Promotion. *The Cornell Hotel and Restaurant Administration Quarterly, 43*(1), 61.

McCullagh, Peter and Nelder, John A. (1989). *Generalized Linear Models*. 2nd ed., Chapman and Hall, London.

Merisavo, M. and Raulas M. (2004). The Impact of E-mail Marketing on Brand Loyalty. *Journal of Product and Brand Management*, 13 (7), 498-505.

Narasimhan, C. (1984) "A Price Discrimination Theory of Coupons," *Marketing Science*, 3, 2, (Spring), pp. 128-147.

Phelps, J. E., Lewis, R., Mobilio, L., Perry, D., and Raman, N. (2004). Viral Marketing or Electronic Word-of-mouth Advertising: Examining Consumer Responses and Motivations to Pass Along Email. *Journal of Advertising Research, 44*(4), 333-348.

Phillip, M. V., & Suri, R. (2004). Impact of Gender Differences on the Evaluation of Promotional Emails. *Journal of Advertising Research, 44*(4), 360-368.

R Development Core Team (2006). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Ramond, Charles (1976). *Advertising Research: The State of the Art*, New York: Association of National Advertisers.

Ratchford, B. (1982). Cost-benefit Models for Explaining Consumer Choice and Information Seeking Behavior. *Management Science*, 28 (2), 197-212.

Rettie, R., Grandcolas U. and Deakins B. (2005). Text Message Advertising: Response Rates and Branding Effects. *Journal of Targeting, Measurement and Analysis for Marketing*. 13 (4), 304-312.

Roy, A. and Berger P. D. (2005). Email and Mixed Mode Database Surveys Revisited: Exploratory Analyses of Factors Affecting Response Rates. *Database Marketing & Customer Strategy Management*, 12 (2), 153-171.

Sharpe, W. F.; Alexander, G. J.; Bailey, J. V.; and Sharpe, W. C. (1998). *Investments*, 6th ed. Englewood Cliffs, NJ: Prentice-Hall.

Tassabehji, Rana and Maria Vakola (2005). Business Email: the Killer Impact. *Communications of the ACM*, 48 (11), 64-70.

Tezinde, T., Smith, B., and Murphy, J. (2002). Getting Permission: Exploring Factors Affecting Permission Marketing. *Journal of Interactive Marketing, 16*(4), 28.

Venables W. N. and Ripley B.D. (2002). *Modern Applied Statistics with S*, fourth edition, Springer-Verlag New York, Inc.

Vriens, Marco, Hiek R. van der Scheer, Janny C. Hoekstra, and Jan Roelf Bult (1998). Conjoint Experiments for Direct Mail Response Optimization. *European Journal of Marketing*, 32 (3/4), 323-339.

Weber, Ron (2004). The Grim Reaper: The Curse of E-Mail (editorial by the Editor-in- Chief). *MIS Quarterly*, 28 (3), iii-xii.

Weisband, S. P. and Reinig B. A. (1995). Managing User Perceptions of email Privacy. *Communications of the ACM*, 38 (12), 40-47.

Zettlemeyer, F. (2000). The Strategic Use of Consumer Search Cost. Working paper, University of California at Berkeley Haas School of Business.

Zhao, J.L., A. Kumar, and E. A. Stohr (2001). Workflow-Centric Information Distribution Through Email. *Journal of Management Information Systems*, 17 (3), 45-72.

Chapter 3

Essay 2: The Market Valuation of Internet Channel Additions: An Event Study of Retail industry

3.1 Introduction

As an important element of the marketing mix, distribution channels are not necessarily fixed or permanent, and innovators in the marketplace often adopt more efficient ways to make goods available to buyers (Lilien, Kotler, and Moorthy 1992). One of the successful distribution channel innovations is multiple channels. As Frazier (1999) points out, the use of multiple channels of distribution is now becoming the rule rather than the exception. Despite possible conflicts between channels, Frazier (1999) notes that there are occasions when channels might complement each other. For example, Victoria Secret uses two primary channels to sell its products: retail stores and mail catalogs. The mail catalogs are likely to increase traffic at the retail stores by providing greater exposure for the brand.

With the emergence of the Internet and development of e-commerce, many established companies have added the Internet to their channel portfolio in the hope that it will improve sales and enable them to reach more market niches. But they also have some concerns about the new Internet channel. One of the concerns is that online activities may cannibalize retailers' offline business and hurt their profits (Alba *et al.* 1997). Studies have been done that evaluate the performance of firms after they have added or announced their intension to add Internet channels One example is Geyskens, Gielens, and Dekimpe's (2002) study on the market valuation of the addition of Internet channels by newspaper companies. Their study assesses the impact of adding an Internet channel on stock market returns and concludes that, on average, stock markets react positively to Internet channel additions. The reaction, however, is affected by power of firms, new channel introduction strategies, and marketplace characteristics. One of the limitations of this study is that it considers only the short-term responses of stock markets (*i.e.*, investors) rather than the long-term responses of consumers (as measured by, say, sales and consumer satisfaction) or the financial performance of firms (as measured by profits and stock prices). A follow-up paper (Deleersnyder *et al.* 2002) examines the long-term effects of Internet channel additions on the performance of firms. Using a database of newspaper companies, the authors find that in the newspaper industry the added Internet channel rarely cannibalizes the traditional one. A limitation of their study is that when they look at long-term performance, they use only two variables: circulation and advertising revenues. They do not look at long-term financial market responses to channel additions.

Although the above-mentioned two studies examine the effect of channel additions, it is very hard to generalize their conclusions to other industries. Unlike the retail industry, the newspaper industry provides only information-based services. Lee and Grewal (2004) examine this issue for the retail industry from the perspective of a strategic response to new technologies. They look at the organizational resources and strategic responses of traditional store-based retailers to new technologies (the Internet) and how they affect performance (*i.e.*, market valuation of the firm, operationalized as Tobin's q). Specifically, they find that the adoption of the Internet as a communication channel and e-alliance formation positively influence a firm's performance. The adoption of the Internet as a sales channel, however, seems to matter only to firms that have preexisting catalog operations.

Unlike the research by Geyskens, Gielens, and Dekimpe's (2002) and by Deleersnyder *et al.* (2002), the present study looks at the channel addition issue from the perspective of the retail industry. My research also differs from Lee and Grewal's (2004) by looking at the immediate response of the stock market to retailers' online channel additions. While their study examines three different strategic responses (*i.e.*, communication channel addition, sales channel addition, and e-alliance), I consider only the addition of sales channels or, more specifically, the announcement of such additions.

Although online channels are a special case of direct sales channels, which have been well researched, traditional direct channel research does not always apply to online channels. The online channel is different from traditional direct channels such as catalogs in the following ways: (1) The online channel is a faster and sometime cheaper way than other direct channels to provide customers with vivid product information. It could even immediately deliver some products such as digital music and movies to customers. (2) The online channel provides customers with an interactive multi-media shopping environment, in which customers and retailers can communicate interactively with each other. Because of these two advantages the online channel is becoming a more and more dominant direct sales channel to retailers. And these special aspects of the channel justify more academic research in this area.

The paper is organized as follows: (1) The advantages and disadvantages to retailers of online channel additions are discussed, and some factors that affect their potential to add online channels are analyzed. (2) An event study is conducted to examine the stock market response to retailers' announcements of online channel additions. (3) Factors that affect the variation of abnormal returns across different retailers are explained. (4) Conclusions and managerial applications are discussed.

3.2 Online Channel: Good and Bad for Retailers

Based on my examination of publicly traded retailers in three major U.S. stock markets, 280 of all 488 public retailers had opened online stores by 2004. Although adding a new online channel may be treated as part of the Internet frenzy during the Net boom era, it does bring significant benefits to retailers.

3.2.1 Advantages of the Online Channel

The advantages of adding an online channel include the following:

(1) Increased sales. Geyskens, Gielens, and Dekimpe's (2002) conclude that the Internet can increase sales in three ways: market expansion, brand switching, and relationship deepening. Market expansion means that a retailer, especially a small, local retailer, can use the Internet channel to reach consumers who cannot buy from bricks-and-mortar stores. Brand switching means that the new Internet channel can attract consumers from competitors. Relationship

deepening means the Internet channel can sell more to retailers' current customers.

(2) A reduction of transaction costs. For retailers, the distribution channel has three functions: distribution, transaction, and communication. Compared to the traditional channel, the Internet channel can reduce the cost of all three, at least for some products. The Internet can be used to distribute some products (e.g., online music or movies), which saves costs incurred in traditional distribution channels. Many products can be sold through the Internet with lower transaction costs. Peterson *et al.* (1997) argue that whether a product is suitable to sell online depends on the characteristics of the product. They categorize products and services along three dimensions: cost and frequency of purchase, value proposition (tangible or intangible), and degree of differentiation. They propose that low outlay, frequently purchased, and intangible products are especially suitable to sell on the Internet. Finally, for the communication function, the Internet channel is a fast, cheap, and broad media to distribute information to buyers.

3.2.2 Disadvantages of the Online Channel

Despite these advantages, the Internet is not entirely free of costs for retailers. The disadvantages of adding a new online channel include the following:

(1) An online channel can increase conflict between channels. Channel cannibalization between traditional channels and an Internet channel is always a serious concern among practitioners and academic scholars. Deleersnyder *et al.*

(2002) summarize four reasons for this concern. First, sales shift from the entrenched channels to the Internet channel. Second, the consumer can compare prices across firms quickly and easily online; therefore, the Internet channel is likely to increase the power of the consumer. Third, the consumer may be less inclined to make impulse purchases online and will buy less through the new channel. Fourth, the existing channel may not welcome the new channel. Although their study of the newspaper industries of the UK and the Netherlands does not find any proof of channel cannibalization, without further empirical evidence we cannot rule out this concern in the retail industry.

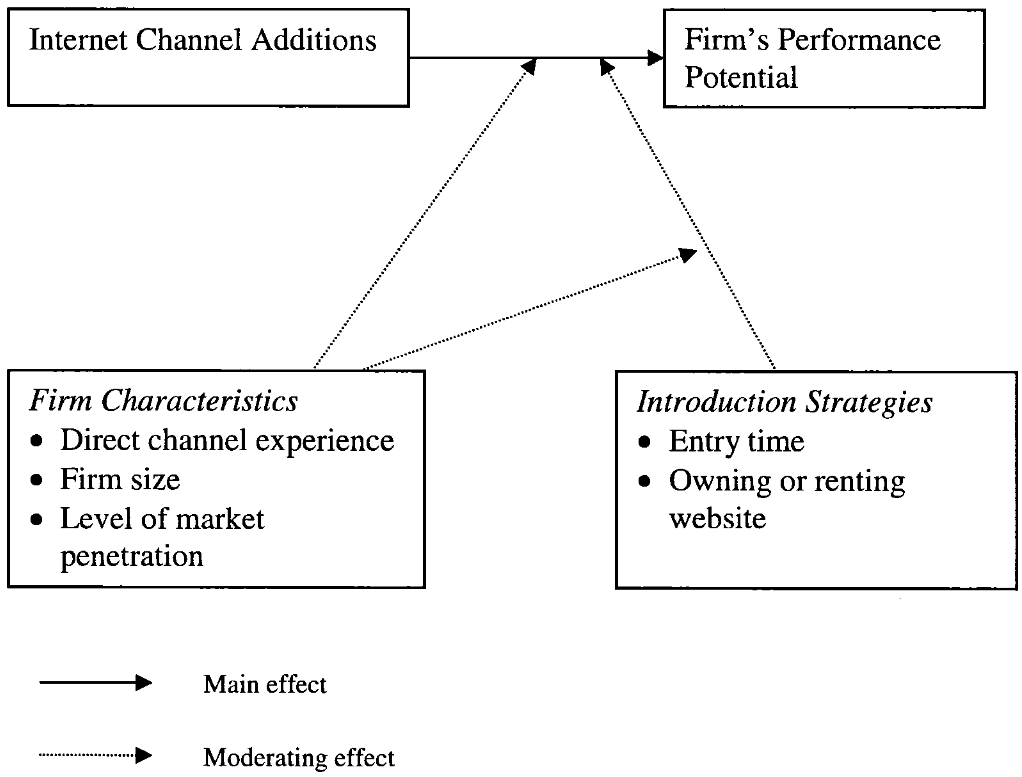
(2) An online channel can increase transaction cost. As the framework of Peterson *et al.* (1997) suggests, not all products are suitable to sell on the Internet. These low-outlay, frequently purchased, and tangible products are especially suitable to sell in bricks-and-mortar stores, and selling them online will increase transaction cost.

Because a retailer faces so many concerns in deciding to add an online channel, it is interesting to look at the net effect of online channel additions. Before measuring the net effect, I will discuss some factors that may moderate the impact of the pros and cons. The present study examines characteristics of retail firms and Internet channel introduction strategies.

3.3 Hypotheses

In this section I examine the moderating role of a retail firm's characteristics and Internet channel introduction strategies on the potential performance of Internet sales channel additions. The discussion framework is illustrated in Figure 3.1.

Figure 3.1 THE EFFECT OF INTERNET CHANNEL ADDITIONS ON PERFORMANCE AND MODERATING EFFECTS



3.3.1 Firm Characteristics

Geyskens et al. (2002) consider three dimensions of firm resources and capabilities: channel power, direct channel experience, and firm size. Because retailers are the final firms serving consumers, I do not consider distributors or channel power. On the other hand, because retailers differ from each other in their market penetration level, I look at the influence of this factor on the success of Internet channel additions. In the present study, I look at three characteristics of firms: direct channel experience, firm size, and level of market penetration.

Direct Channel Experience. The direct channel experience of a retailer may affect the success of Internet channel additions. A retailer who has an established direct channel can easily apply that experience to the new Internet channel. I propose a positive relation between a retail firm's direct channel experience and its online channel addition success.

H_1: The performance potential of an Internet channel addition is higher for retailers with direct channel experience than those without such experience.

Firm Size. In the traditional bricks-and-mortar economy, large retailers have many competitive advantages over small ones (*e.g.*, more resources to serve more consumers). In addition, large retailers obtain economies of scale. In the "new economy," Internet stores can theoretically reach any consumer with Internet access in any corner of the world. Because the advantage large retailers traditionally have is weakened, I propose that small retailers benefit more from the Internet store addition than their larger peers.

H_2: The performance potential of an Internet channel addition is higher for small retailers than for big retailers.

Level of Market Penetration. Similarly, in the old economy national and international retailers could use more stores to penetrate deeply into the market while local retailers could serve consumers with limited number of stores. This penetration advantage is also weakened by the appearance of the new Internet channel, and I propose that retailers with fewer stores benefit more from the Internet store addition.

H_3: The performance potential of an Internet channel addition is higher for retailers with low level of market penetration than for those with high level of market penetration.

3.3.2 Introduction Strategy Characteristics

When retailers decide to go online, they have to decide how to introduce the new channel. This study considers two introduction strategies: time of entry and owning or renting a sales website.

Time of Entry. Kalyanaram *et al.* (1995) conclude that there is enough evidence for the advantage of early entry. In contrast, some research suggests that by waiting and learning from the experience of early entrants, late entrants can save on the cost on learning of new technologies (Geyskens *et al.* 2002). Because the effect of entry time in the retail industry is unclear, I do not propose any hypothesis about it although I do examine this effect in the analysis.

Owning or Renting a Website. Some retailers build their own websites to sell products, while others sell through famous portal websites (*e.g.*, Yahoo.com) or through the online marketplace (*e.g.*, eBay.com). I argue that, despite the cost, building a website can increase the creditability of the Internet channels and build a better relationship with customers. I therefore propose that selling products through retailers' own online stores works better than selling them through a rented website space.

H_4: The performance potential of an Internet channel addition is higher for retailers who sell through their own websites than for those who sell through a rented website space.

3.3.3 Interaction between Firm Characteristics and Introduction Strategies

The success of different introduction strategies is also moderated by firm characteristics. First I look at the influences on the entry time decision of different retailers, *i.e.*, how retailers choose entry time based on their direct channel experience, firm size, and level of market penetration. I argue that for the retailers with direct channel experience, an early Internet channel addition works better. If they do not move online on time, they will lose the advantage of direct channel experience. For firm size, I think that big retailers will benefit more than small retailers from an early Internet channel addition because big retailers have more resources and can afford the costs of a failure. Later entry gives small retailers time to imitate and learn from the pioneers. For level of market penetration, the Internet's geographic reach allows local retailers to reach a much broader area than they otherwise reach, but if they delay moving online, they will miss the opportunity. I therefore argue that retailers who have low market penetration level can benefit more from an early Internet channel addition than national and international retailers do. The above arguments are summarized by the following hypotheses.

H_5: Retailers with direct channel experience benefit more from an early Internet channel addition than retailers without such experience do.

H_6: Big retailers benefit more from an early Internet channel addition than small retailers do.

H_7: Retailers that have low market penetration benefits more from an early Internet channel addition than retailers that have high level of market penetration.

I now look at influences on the decisions of different retailers to own or rent a website, *i.e.*, based on their direct channel experience, firm size, and level of market penetration, should retailers build their own sales websites or sell through other websites? I argue that retailers without direct channel experience should have their own websites to build up a reputation among consumers of their direct sales business. However, having their own website is not crucial for the retailers with direct channel experience. In terms of size, big retailers have enough resources to build their own websites while small retailers can begin by renting. For level of market penetration, because national and international retailers can better combine their online activities with their traditional stores, they benefit more from their own websites. I summarize the above arguments with the following hypotheses.

H_8: Retailers without direct channel experience benefit more from building their own websites than retailers with such experience do.

H_9: Big retailers benefit more from building their own websites than small retailers do.

H_{10}: Retailers that have high market penetration level benefit more from owning their websites than retailers that operate in fewer areas do.

The hypotheses (H_1 to H_{10}) are summarized in Table 3.4.

3.4 The Event Study

In this section, I will conduct an event study to examine the main effect illustrated in Figure 3.1 — the influence of Internet channel additions on the performance potential of retailers. Event study methodology was developed more than thirty years ago and appears frequently in financial services studies to measure the effect of an economic event on the value of firms. Event studies attempt to measure abnormal changes in the stock prices of publicly traded companies that occur in conjunction with an economic "event" such as the announcement of a new regulatory initiative or a new marketing strategy. The logic of conducting event studies is based on the economic concept of the "perfect market" — the price of publicly traded stocks should reflect the reaction of financial markets to the introduction of new information (MacKinlay 1997).

The event study has become popular because it obviates the need to analyze accounting-based measures of profit, which have been criticized because they are often not very good indicators of the true performance of firms (McWilliams and Siegel 1997). Event studies have already been used broadly in marketing research. Examples include Peltzman's (1981) study on regulations and rulings on false advertising, Jarrell and Peltzman's (1985) study on product recalls, Horsky and Swyngedouw's study (1987) on a company's name change, Chaney *et al.*'s (1991) study on new product introduction, Agrawal and Kamakura's (1995) study on celebrity endorsement, Lane and Jacobson's (1995) study on brand extension, and Geyskens et al's (2002) study on Internet channel additions.

3.4.1 Event Study Methodology

As a well-developed research method in finance research and other related fields, the event study has a general flow of analysis. Mackinlay (1997) summarizes it as follows:

1. *To define the event of interest and identify the event window.*

The event of interest is the action of releasing information to the public (*i.e.*, announcements). The event window is the period over which the stock prices of the firms involved in this event will be examined. Obviously, an event window should include the day of the announcement. In practice, it is customary to expand the event window to cover multiple days to permit examination of periods surrounding the event.

2. *To determine the selection criteria for the inclusion of a given firm in the study.*

The selection criteria may define publicly traded companies on certain stock exchange(s) and/or companies in a certain industry or some industries.

3. *To measure abnormal returns.*

In order to measure the impact of an event, we measure the abnormal returns that are due to the happening of the event. We first define the concept of stock return.

The stock return is defined as the percentage change in the stock prices.

$$R_{it} = \frac{P_{it} - P_{i(t-1)}}{P_{i(t-1)}} \qquad (1)$$

where R_{it} is the return of stock i at time t, P_{it} is the stock price of stock i at time t, and $P_{i(t-1)}$ is the stock price of stock i at time t-1.

The abnormal return for stock i on event date t is defined as:

$$AR_{it} = R_{it} - E(R_{it}|X_t) \qquad (2)$$

where AR_{it}, R_{it}, and $E(R_{it}|X_t)$ are the abnormal, actual, and expected normal returns respectively for time t. X_t is the conditioning information for the normal return model. Although there are several ways to model the normal return, Mackinlay (1997) concludes that two of them are used more commonly. The first one is *the Constant Mean Return Model.* This model assumes that the mean return of a given security is constant through time. The second one is *the Market Model.* This model assumes a stable linear relation between the market return and the security return. For security i, the market model is

$$R_{it} = \alpha_i + \beta_i R_{mt} + \varepsilon_{it} \qquad (3)$$

$$E(\varepsilon_{it}) = 0 \qquad \mathrm{var}(\varepsilon_{it}) = \sigma^2_{\varepsilon_i}$$

where R_{it} and R_{mt} are the return for stock i and the return for the market portfolio, respectively, on time t. ε_{it} is a zero mean disturbance term. α_i, β_i, and $\sigma^2_{\varepsilon_i}$ are the parameters of the market model. In application some stock index is used for the market portfolio, *e.g.*, S&P 500 Index, the CRSP Value Weighted Index, or the CRSP Equal Weighted Index. To estimate the values of α_i, β_i, and $\sigma^2_{\varepsilon_i}$, stock prices and market portfolio prices for a period of time are needed. This period of time is called the estimation window. The estimation window comes before the event window, and the two do not overlap.

As to the estimation method of the market model, an ordinary least squares analysis (OLS) is an efficient and consistent procedure under general conditions. Given the market model parameter estimates, the abnormal returns can be measured and analyzed. For any day in the event window, the abnormal return is given as follows:

$$AR_{it} = R_{it} - \hat{\alpha}_i - \hat{\beta}_i R_{mt} \qquad (4)$$

where $\hat{\alpha}_i$ and $\hat{\beta}_i$ are the estimates for the market model parameters α_i and β_i. Equation (4) gives the abnormal return for one announcement of one stock. When we conduct an event study across several announcements or across several stocks, we test the average effect of an event by computing the average of the abnormal returns over all announcements:

$$AAR_t = \sum_{i=1}^{N} AR_{it} / N \qquad (5)$$

where N is the number of announcements being studied. So far I have discussed the abnormal return for a specific day. When the event window covers more than one day, we need to aggregate returns in different days to draw overall inferences for the event of interest. To aggregate the abnormal returns through days, a measure of cumulative abnormal return (CAR) is introduced. $CAR_i(t_1, t_2)$ is defined as the cumulative abnormal return for stock i from time t_1 to t_2 where t_1 and t_2 are time points within the event window.

$$CAR_i(t_1,t_2) = \sum_{t=t_1}^{t_2} AR_{it} \qquad (6)$$

Then the cumulative abnormal returns are aggregated through stocks or announcements. Equation (5) gives the aggregated abnormal return for a specific day. Similarly, for any interval in the event window, the average cumulative abnormal return (*CAAR*) for all stocks or announcements is given as follows:

$$CAAR\ (t_1,t_2) = \sum_{t=t_1}^{t_2} AAR_t \qquad (7)$$

4. *To test the significance of abnormal returns*

Under the null hypothesis that the event has no impact on the behavior of returns (mean or variance), the distributional properties of the abnormal returns can be used to draw inferences over any period within the event window. The distribution of the abnormal returns in the event window is:

$$AR_{it} \sim N(0,\sigma^2(AR_{it}))$$

When the length of the estimation window becomes large, the variance of AR_{it} approaches the variance of the error term in equation (3), $\sigma^2_{\varepsilon_i}$.

5. *Nonparametric test-the sign test*

So far I have discussed the parametric test, which is based on specific assumptions about the distribution of abnormal returns. Sometimes we can also use nonparametric tests, which do not require the above mentioned assumptions. The sign test is one of the nonparametric tests, and it requires that the abnormal returns (or more generally cumulative abnormal returns) are independent across securities. The null hypothesis is that the expected proportion of positive abnormal returns is 0.5. The test statistic is calculated as follows:

$$Z = \left(\frac{N^{+}}{N} - 0.5 \right) \frac{\sqrt{N}}{0.5} \sim N(0,1) \qquad (8)$$

Where N^{+} is the number of cases in which the abnormal return is positive and N is the total number of cases.

3.4.2 Data and Results

3.4.2.1 The Event of Interest and the Sample

The industry of interest in this study is the retail industry. According to COMPUSTAT North America database, there were 488 publicly traded retailers on three major stock exchanges in the U.S. by 2004 (*i.e.*, New York Stock Exchange, the American Stock Exchange, and the NASDAQ). I checked all these 488 retailers when I looked for the event of interest. The event of interest in this study is the first announcement of a retailer's Internet sales channel additions. One announcement counts as one event. A retailer who makes no announcement is not included in the sample. Therefore there is only one event date for each the retailer in the sample. The event date in this study is not the opening day of an Internet store, but the day a retailer first announced the Internet channel addition. Some retailers open websites not to sell products, but to provide information. Lee and Grewal (2004) classify the online channels of retailers into two categories: communications channels and sales channels. In the present study, I examine the only additions of sales channel.

I used Factiva to find the event for each of the 488 retailers. Factiva is an online resource that includes the full text of newspapers from around the world in addition to Dow Jones & Reuters newswires, business journals, market research reports, analyst reports, and web sites. Factiva's forerunner is "Dow-Jones Interactive", which was replaced by Factiva on July 1, 2003. "Dow Jones Publications" are used in many event studies (*e.g.*, Horsky and Swyngedouw 1987 and Geyskens et al. 2002). The key words I searched in Factiva were the company names of the retailers and "online or Internet". For example, I searched for events about Wal-Mart using the key words "Wal-Mart and "online or Internet". "Searching Dates" were set to "all dates" in the database. For the "Sources" option, I chose "Dow Jones Newswires" and "Major US News and Business Publications". The "Region" option was set to "North American Countries".

I found that 94 retailers clearly announced that they would add an Internet channel to their current bricks-and-mortar channels. Of these 94 retailers, 16 also announced some other news on the same day (*e.g.*, an appointment of a new chairman). I deleted these 16 retailers to avoid confounding effects. The names and event dates for the 78 retailers are given in the Appendix. Of the 78 remaining retailers, the first online channel addition announcement happened on November 30, 1994, and the last one on October 13, 2004. These retailers belong to eight SIC retailing sectors (see Table 3.1).

Table 3.1 THE DISTRIBUTION OF 78 RETAILERS IN EIGHT RETAILING SECTORS

SIC Code	SIC Description	No. of retailers
52	Building materials, hardware, garden supply, & mobile	5
53	General merchandise stores	6
54	Food stores	6
55	Automotive dealers and gasoline service stations	7
56	Apparel and accessory stores	16
57	Furniture, home furnishings and equipment stores	9
58	Eating and drinking places	2
59	Miscellaneous retail	27
Total		78

3.4.2.2 The Event Window

The event window used in the present study is a 31-day event window (-15 to +15), which is quite short. Although many event studies use long event windows, McWilliams and Siegel (1997) contend that very few provide justification for the length used. I use this short event window for the following reasons: (1) Using a long event window severely reduces the power of the test statistic, Z_t. And this reduction leads to false inferences about the significance of an event (Brown and Warner 1980, 1985); (2) It is much more difficult to control for confounding effects when long windows are used (McWilliams and Siegel 1997); (3) A short window will usually capture the significant effect of an event (Ryngaert and Netter 1990). For example, Dann, Mayers, and Raab (1989) find that the stock price fully adjusts within 15 minutes of the release of firm-specific information; and (4) Marketing researchers tend to use shorter event windows. For example, Chaney, Devinney and Winer (1991) use four event windows (-1 to +1, -3 to +3, -5 to +1, and -5 to +5) in their new product introduction study. Agrawal and Kamakura (1995) use a 21-day window (-10 to +10) in a celebrity endorsement study. Lane and Jacobson's (1995) use an 11-day window (-5 to +5) in a brand extension study, so does Geyskens et al's (2002) study on Internet channel additions.

3.4.2.3 The Estimation Window and the Normal Return Model

The estimation window in the study covers 220 days from the day of t = -250 to the day of t = -31. The day of t = 0 is the event day.

To model the normal return, the market model shown in equation (3) is used. Stock prices from the day of t=-250 to the day of t=120 are collected for each of the 78 retailers from the Center for Research in Security Prices (CRSP) database. Thus stock prices for 371 days are collected for each retailer, and the daily stock price returns are calculated accordingly. The stock market indices for the three stock markets on corresponding days are obtained from CRSP. The daily market returns in equation (3) are calculated accordingly. I then estimate the coefficients in equation (3), where R_{it} is the stock price return for stock i on day t, and R_{mt} is the market index return on day t for the stock market where retailer i is publicly traded. The values of R_{it} and R_{mt} from t = -250 to t = -31 are used to estimate the values of α_i and β_i for each stock. The error terms ε_{it} in equation (3) are also calculated. The market model works well in this study. Out of 78 regressions, 54 have significant F values (For the market models with insignificant F values, β_i are very small, so the market models are similar to the constant mean return model). The maximum R square is 54%.

3.4.2.4 The Abnormal Return and Cumulative Abnormal Return

Then I use the estimated α_i and β_i to calculate daily abnormal returns (AR_{it}) for each stock from the day of t=-15 to the day of t=15. I have calculated the error terms ε_{it} in equation (3) for the estimation window. The variances of the error terms are used as the variances for the abnormal returns. The average cumulative abnormal returns *CAAR (-15, t)* from the day of t=-15 to the day of t are also calculated. Table 3.2 summarizes the average abnormal returns AAR_t on the day t and average cumulative abnormal returns *CAAR (-15, t)*. Figure 3.2 illustrates the change of *CAAR (-15, t)* over time.

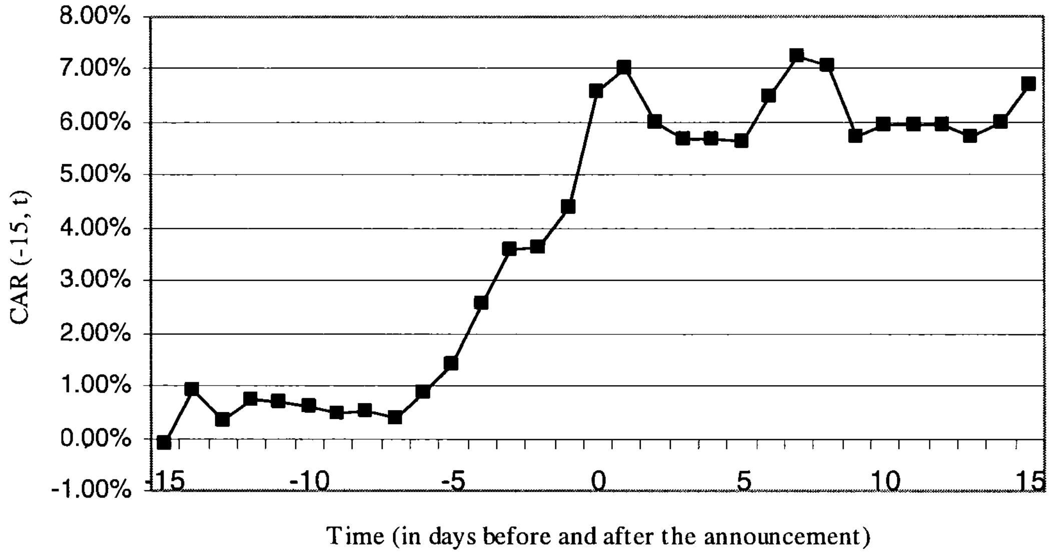
Table 3.2 shows that, on average, retailers' announcements of Internet sales channel additions obtained 2.18% abnormal return on the event day t=0, 0.77% on the day t=-1, and 0.44% on the day t=1. The magnitude of the stock market reaction to Internet channel announcements in the present study is much higher than that reported by Geyskens, Gielens, and Dekimpe (2002), whose study is about the newspaper industry. In their study, the average abnormal returns for t=-1 to t=1 are -0.27%, 0.35%, and 0.36% respectively. This means that stock markets react more positively to the Internet channel announcements of retailers than to those of newspaper companies. This disparity justifies my argument that the retail industry must be studied separately.

Table 3.2 SUMMARY OF ABNORMAL RETURNS FOR 31 DAYS

t	AAR_t (%)	CAAR (-15, t) (%)	Number of Retailers with Significant Positive AR_i[1]	Number of Positive AR_i	Percentage of Positive AR_i	z value[2]	p value
-15	-0.12	-0.12	3	30	0.38	-2.04	0.98
-14	1.01	0.90	8	43	0.55	0.91	0.18
-13	-0.57	0.33	3	31	0.40	-1.81	0.97
-12	0.41	0.74	7	37	0.47	-0.45	0.67
-11	-0.03	0.71	4	33	0.42	-1.36	0.91
-10	-0.1	0.61	5	39	0.50	0.00	0.50
-9	-0.16	0.45	5	31	0.40	-1.81	0.97
-8	0.07	0.52	5	38	0.49	-0.23	0.59
-7	-0.16	0.36	2	39	0.50	0.00	0.50
-6	0.52	0.89	6	42	0.54	0.68	0.25
-5	0.52	1.41	6	37	0.47	-0.45	0.67
-4	1.14	2.55	7	38	0.49	-0.23	0.59
-3	1.06	3.61	8	39	0.50	0.00	0.50
-2	0.01	3.62	6	35	0.45	-0.91	0.82
-1	0.77	4.39	8	43	0.55	0.91	0.18
0	**2.18**	**6.57**	**10**	**50**	**0.64**	**2.49**	**0.01**
1	0.44	7.01	5	37	0.47	-0.45	0.67
2	-1.02	5.99	2	35	0.45	-0.91	0.82
3	-0.31	5.68	6	33	0.42	-1.36	0.91
4	-0.01	5.67	8	34	0.44	-1.13	0.87
5	-0.02	5.65	7	31	0.40	-1.81	0.97
6	0.86	6.50	7	37	0.47	-0.45	0.67
7	0.73	7.23	6	39	0.50	0.00	0.50
8	-0.16	7.07	3	41	0.53	0.45	0.33
9	-1.36	5.71	2	27	0.35	-2.72	1.00
10	0.22	5.93	7	36	0.46	-0.68	0.75
11	0.04	5.97	8	34	0.44	-1.13	0.87
12	-0.02	5.95	5	38	0.49	-0.23	0.59
13	-0.22	5.73	5	38	0.49	-0.23	0.59
14	0.25	5.99	6	38	0.49	-0.23	0.59
15	0.72	6.70	7	41	0.53	0.45	0.33

[1]: An abnormal return bigger than 1.64 times standard deviation of abnormal returns is considered to be significantly positive.
[2]: z-value is calculated by equation (8).

Figure 3.2 AVERAGE CUMULATIVE ABNORMAL RETURN
***CAAR (-15, T)* OVER TIME**

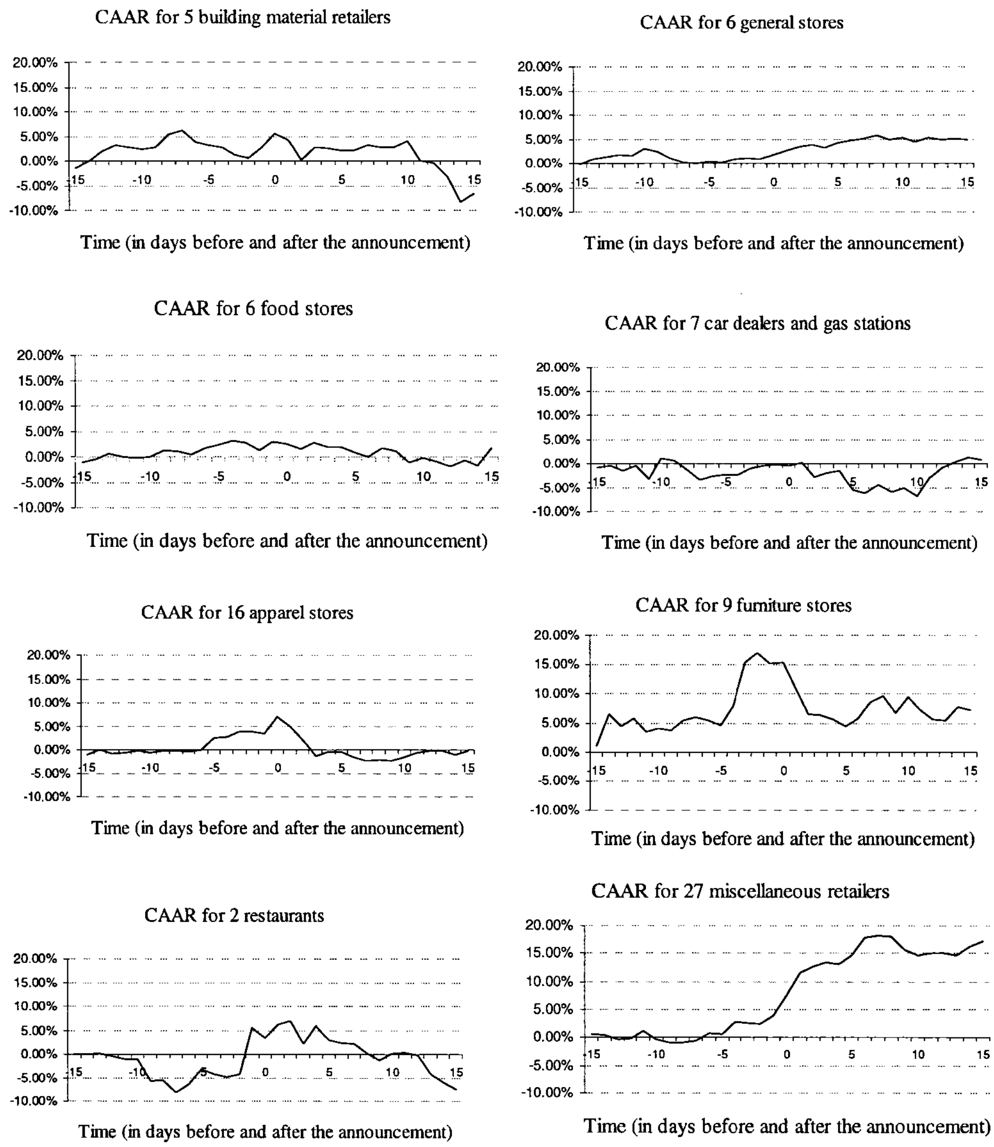


Of all windows surrounding the event day, the one from $t = -6$ to $t = 1$ shows the most significant jump of *CAAR*, which means that there has been some information leakage for some retailers before the event day. The number of retailers with significant positive abnormal returns ($p < 0.05$) also increases before and on the event day. The number is 7, 8, 6, and 8 from $t = -4$ to $t = -1$ respectively, and on the event day it hits 10. I then conduct the nonparametric test. Of 78 retailers, 50 (64%) got positive abnormal returns on the event day. Therefore, significantly more retailers get positive abnormal returns on the event day than get negative ones ($p < 0.01$).

When looking at abnormal returns for retailers in different retailing sectors, it is evident that stock markets react differently to Internet channel addition announcements by retailers in different sectors (Figure 3.3). Retailers in some sectors (*e.g.*, food store, car dealers, and gas stations) did not get significant

Figure 3.3 AVERAGE CUMULATIVE ABNORMAL RETURN

CAAR (-15, T) FOR 8 RETAILING CATEGORIES



abnormal returns from the announcements on the event day or during the whole month around the event days. Retailers in some other sectors (*e.g.*, building material retailers, apparel stores, furniture stores, and restaurants) got some positive abnormal returns on or before the event days, but the stock prices soon returned to normal. Only retailers in two sectors (general stores and miscellaneous retailers) got significant abnormal returns during the event window. General stores got a moderate cumulative abnormal return (5%), and miscellaneous retailers got a huge cumulative abnormal return (17%) within two weeks after the event days.

3.5 Explanation of Abnormal Return Variation across Retailers

3.5.1 Operationalization of Measures

3.5.1.1 Firm Characteristics

Firm Size. Following Geyskens et al. (2002), I compiled three measures to represent firm size: number of employees, annual sales, and the market values of the retailers. The values for the first two measures were obtained directly from COMPUSTAT North America database. The data are for the year before the event happened. The retailers' market values were calculated by the following formula:

$$\text{Market value} = \text{closing price} \times \text{common share outstanding} + \text{liquidating value (preferred stock)} + \text{long-term debt} \tag{9}$$

All the variables in the right hand of equation (9) were obtained from COMPUSTAT. The market values are data for the last trade day of the year before the event happened. I first standardized the three measures (number of employees, annual sales, and the market values) and then averaged the three items into a single scale of firm size.

Direct Channel Experience. I used a dummy variable — whether a retailer had a direct channel before the event day — to measure the retailer's direct channel experience. If a retailer had a direct sales channel before the event day, it was coded as "1"; otherwise it was coded as "0." The values for this variable were obtained from retailers' annual reports of the year before the event day, which were obtained through US Securities and Exchange Commission Filings & Forms Service (EDGAR).

Level of Market Penetration. I used the number of stores to measure retailers' market penetration level. This information was also obtained from retailers' annual reports of the year before the event happened. The original value was standardized.

3.5.1.2 Introduction Strategies

Time of Entry. I measured time of entry using the number of days the retailer announced the online channel addition after January 1, 1994. The original value was standardized.

Owning or Renting Website. I measured this using a dummy variable, and the values were obtained from the announcements found in Factiva. If a retailer sells products through its own website, it is coded as "1", otherwise it is coded as "0". The descriptive statistics for these variables are given in Table 3.3.

Table 3.3 DESCRIPTIVE STATISTICS FOR VARIABLES

Dependent Variables	*Sum*	*Mean*	*Median*	*Minimum*	*Maximum*	*Standard Deviation*	*Transformation for the Models*
AR_0		0.024	0.008	-0.092	0.424	0.077	
Independent Variables							
Firm Characteristics							
Firm Size							
Number of Employees (Thousand)		41.287	4.520	0.010	622.000	93.951	Standardization
Annual Sales (Million $)		5922.941	459.370	9.090	82494.000	13209.236	Standardization
Market Value (Million $)		5312.184	299.448	1.608	91845.097	13792.498	Standardization
Direct Channel	34						
Number of Stores		977	303	1	22000	2649	Standardization
Channel Introduction Strategy							
Time of Entry		1869	1950	333	3938	532	Standardization
Own Website	64						

3.5.2 Results

The abnormal return on the event day (*i.e.*, AR_0) is regressed on the aforementioned independent variables. The results of the regression are given in Table 3.4. Retailers with direct channel experience benefit more from the Internet channel addition announcements (b = .204, $p < .001$). So H_1 is supported. Big retailers get significantly less abnormal returns than their small peers do (b = -.086, $p < .05$). So H_2 is supported. Level of market penetration does not significantly affect retailers' online channel potential ($p > .05$). H_3 is not supported. For the influence of channel introduction strategies, entry time does not significantly affect performance potential. Owning websites seems better than renting websites, but the effect is not significant ($p > .05$). H_4 is not supported. For the interaction effects, entry time decision does not significantly interact with firm characteristics ($p > .05$ for the three interaction coefficients). H_5, H_6, and H_7 are not supported. However, a retailer's website ownership decision interacts significantly with some firm characteristics. Retailers with direct channel experience have significantly less performance potential from owning sales websites than retailers without such experience (b = -.184, $p < .01$). H_8 is supported. Big retailers have significantly higher performance potential from owning sales websites than small retailers (b = .076, $p < .05$). H_9 is supported. Level of market penetration does not significantly interact with the website ownership decision. H_{10} is not supported.

Table 3.4 REGRESSION RESULTS

Independent Variables:	*Hypotheses*	*Hypothesized Signs*	*Parameter Estimates*
Intercept			-0.007
Firm Characteristics			
Direct Channel Experience	H_1	+	.204***
Firm Size	H_2	-	-0.086*
Number of Stores	H_3	-	0.007
Channel Introduction Strategy			
Entry Time		?	0.002
Owning Website	H_4	+	0.014
Interactions			
Entry Time×Direct Channel	H_5	+	0.007
Entry Time×Size	H_6	+	-0.012
Entry Time×Number of Stores	H_7	-	0.05
Owning Website×Direct Channel	H_8	-	-0.184**
Owning Website×Size	H_9	+	0.076*
Owning Website×Number of Stores	H_{10}	+	-0.018
$F(11, 63)$			1.839
R^2			0.243
R^2 (adjusted)			0.111

*: $p < 0.05$
**: $p < 0.01$
***: $p < 0.001$

3.6 Discussion and Future Research

Adding an Internet channel is an important strategic decision to retailers. The present study provides a new way to evaluate this strategy and shows that generally stock markets react positively to retailers' Internet channel additions. Retailers in different sectors, however, get different abnormal returns from the announcements, which means that investors have different attitudes toward the suitability of different products to the Internet. The abnormal returns on the event days are also influenced by the characteristics of a retail firm (size, direct channel experience, and level of market penetration) and channel introduction strategies (entry time and owning or renting website). The study finds that investors are more optimistic toward the Internet channel additions of small retailers and retailers with direct channel experience. Channel introduction strategies, however, do not affect stock market's reaction to these announcements. As to the decision of owning or renting a website, the study finds that investors expect large retailers to build their own websites to sell and retailers with direct channel experience to sell products through rented web space. Despite the interesting findings in the present study, this question deserves further research in the following ways:

First, the event study looks only at stock markets' short-term reaction to the announcements. It is important and interesting to look long-term financial and marketing performance of retailers with online sales channels.

Second, consumers are the final users of the new Internet channel. Further research therefore should also focus on consumer satisfaction with the new channel and the switching behavior between traditional channels and the new one.

Finally, the new Internet channel is both a substitute for and a complement to traditional ones. Therefore examination of the cannibalization and synergy between the two channels is another promising venue for further research.

Appendix 3.1: List of 78 Retailers in the Study

Retailer No.	Retailer Name	SIC Code	NAICS Code	Stock Market	Event Date
1	TRACTOR SUPPLY CO	5200	453998	NASDAQ	6/4/1999
2	TREND-LINES INC	5200	444130	NASDAQ	1/14/2000
3	HOME DEPOT INC	5211	444110	NYSE	6/29/1999
4	BUILDING MATERIALS HLDG CP	5211	421310	NASDAQ	9/20/2000
5	WICKES INC	5211	444110	NASDAQ	9/20/2000
6	DILLARDS INC	5311	452111	NYSE	5/16/1999
7	FEDERATED DEPT STORES	5311	452111	NYSE	6/26/1998
8	PENNEY (J C) CO	5311	452111	NYSE	11/27/1995
9	SEARS ROEBUCK & CO	5311	452111	NYSE	11/27/1995
10	WAL-MART STORES	5331	452990	NYSE	1/29/1996
11	COSTCO WHOLESALE CORP	5399	452910	NASDAQ	11/21/1995
12	ALBERTSONS INC	5411	445110	NYSE	10/13/2004
13	KROGER CO	5411	445110	NYSE	11/10/1995
14	MARSH SUPERMARKETS	5411	445110	NASDAQ	4/26/1999
15	SAFEWAY INC	5411	445110	NYSE	4/17/2000
16	COLES MYER LTD	5411	445110	NYSE	9/22/1997
17	WHOLE FOODS MARKET INC	5411	445110	NASDAQ	11/30/1994
18	HOLIDAY RV SUPERSTORES INC	5500	441210	NASDAQ	5/22/2000
19	COPART INC	5500	441120	NASDAQ	6/5/1998
20	RUSH ENTERPRISES INC	5500	441110	NASDAQ	9/8/2000
21	LITHIA MOTORS INC	5500	441110	NYSE	2/8/1999
22	MARINEMAX INC	5500	441222	NYSE	9/26/2000
23	HOMETOWN AUTO RETAILERS	5500	441222	NASDAQ	11/12/1999
24	O REILLY AUTOMOTIVE INC	5531	441310	NASDAQ	2/29/2000
25	WILSONS LEATHER EXPERTS INC	5600	448190	NASDAQ	12/14/1999
26	TOO INC	5600	448130	NYSE	6/8/2000
27	CACHE INC	5621	448120	NASDAQ	4/20/1999
28	DEB SHOPS INC	5621	448120	NASDAQ	1/5/1999
29	PAUL HARRIS STORES	5621	448120	NASDAQ	8/1/1999
30	ONE PRICE CLOTHING STORES	5621	448120	NASDAQ	6/6/2000
31	MOTHERS WORK INC	5621	448120	NASDAQ	1/7/1999
32	GAP INC	5651	448140	NYSE	11/6/1997
33	NORDSTROM INC	5651	448140	NYSE	1/14/1999
34	BUCKLE INC	5651	448140	NYSE	4/22/1999
35	PACIFIC SUNWEAR CALIF INC	5651	448140	NASDAQ	3/2/1999
36	GUESS INC	5651	448140	NYSE	3/26/1999
37	ABERCROMBIE & FITCH -CL A	5651	448140	NYSE	5/11/1999
38	BIG DOG HOLDINGS INC	5651	448140	NASDAQ	1/11/1999
39	PAYLESS SHOESOURCE INC	5661	448210	NYSE	7/28/1999
40	SHOE PAVILLION INC	5661	448210	NASDAQ	12/15/1998
41	WILLIAMS SONOMA INC	5700	442299	NYSE	11/1/1999
42	GUITAR CENTER INC	5700	451140	NASDAQ	11/8/1999
43	RESTORATION HARDWARE INC	5712	442110	NASDAQ	4/1/1999

44	BEST BUY CO INC	5731	443112	NYSE	1/6/2000
45	CIRCUIT CITY STORES INC	5731	443112	NYSE	6/15/1999
46	HARVEY ELECTRONICS	5731	443112	NASDAQ	4/16/1999
47	NATIONAL RECORD MART INC	5735	451220	NASDAQ	12/1/1998
48	CD WAREHOUSE INC	5735	451220	NASDAQ	9/9/1998
49	HASTINGS ENTERTAINMENT INC	5735	451220	NASDAQ	4/1/1999
50	MCDONALDS CORP	5812	722211	NYSE	6/22/2000
51	HOST AMERICA CORP	5812	722110	NASDAQ	12/1/1999
52	EZCORP INC -CL A	5900	522298	NASDAQ	8/10/1999
53	ABLE ENERGY INC	5900	454311	NASDAQ	5/11/2000
54	CVS CORP	5912	446110	NYSE	4/29/1999
55	RITE AID CORP	5912	446110	NYSE	1/7/1999
56	WALGREEN CO	5912	446110	NYSE	6/28/1999
57	NYER MEDICAL GROUP INC	5912	446110	NASDAQ	11/2/1999
58	HANCOCK FABRICS INC	5940	451130	NYSE	12/2/1997
59	OFFICE DEPOT INC	5940	453210	NYSE	6/15/1998
60	STAPLES INC	5940	453210	NASDAQ	11/17/1998
61	SPORT CHALET INC	5940	451110	NASDAQ	5/10/1999
62	BARNES & NOBLE INC	5940	451211	NYSE	1/28/1997
63	PAPER WAREHOUSE INC	5940	453220	NASDAQ	8/26/1999
64	SPORTS AUTHORITY INC	5940	451110	NYSE	5/19/1999
65	PREMIER CONCEPTS INC	5944	448310	NASDAQ	7/23/1999
66	ZANY BRAINY INC	5945	451120	NASDAQ	10/19/1999
67	YOUTHSTREAM MEDIA NETWORKS	5960	454390	NASDAQ	8/5/1999
68	K TEL INTERNATIONAL	5961	454113	NASDAQ	4/9/1998
69	BLAIR CORP	5961	454113	AMEX	7/17/2000
70	SPIEGEL INC -CL A	5961	454110	NASDAQ	1/30/1995
71	SPORT SUPPLY GROUP INC	5961	454111	NYSE	1/8/1999
72	UNAPIX ENTERTAINMENT INC	5961	454110	AMEX	5/5/1998
73	J JILL GROUP INC	5961	454111	NASDAQ	6/16/1999
74	BLUEFLY INC	5961	454111	NASDAQ	11/19/1998
75	SCHOOL SPECIALTY INC	5961	454113	NASDAQ	1/26/1999
76	ALBERTO-CULVER CO	5990	446120	NYSE	3/12/2000
77	PETSMART INC	5990	453910	NASDAQ	5/13/1999
78	PETCO ANIMAL SUPPLIES INC	5990	453910	NASDAQ	7/14/1999

Bibliography

Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski (1997). Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces. *Journal of Marketing*, 61 (3), 38-53.

Agrawal, Jagdish and Wagner A. Kamakura (1995), "The Economic Worth of Celebrity Endorsers: An Event Study Analysis," *Journal of Marketing*, 59 (July), 56-62.

Brown, Stephen J., Jerold B. Warner (1980), "Measuring Security Price Performance", *Journal of Financial Economics*, 8, 205-258.

Brown, Stephen J., Jerold B. Warner (1985), "Using Daily Stock Returns: the Cases of Event Studies", *Journal of Financial Economics*, 14, 3-31.

Chaney, Paul K., Timothy M. Devinney, Russell S. Winer (1991), "The Impact of New Product Introductions on the Market Value of Firms", *Journal of Business*, 64 (4), 573-610.

Dann, Larry Y., David Mayers, and Robert J. Raab, Jr. (1997), "Trading Rules, Large Blocks, and the Speed of Price Adjustment." *Journal of Financial Economics*, 4, 3-22.

Deleersnyder, Barbara, Inge Geyskens, Katrijn Gielens, Marnik G. Dekimpe (2002), "How Cannibalistic is the Internet Channel? A Study of the Newspaper Industry in the United Kingdom and the Netherlands", *International Journal of Research in Marketing*, 19, 337-348.

Frazier, Gary L. (1999). Organizing and Managing Channels of Distribution. *Journal of the Academy of Marketing Science*. 27 (2) 226-240.

Geyskens, Inge, Katrijn Gielens, Marnik G. Dekimpe (2002), "The Market Valuation of Internet Channel Additions", *Journal of Marketing*, 66 (April), 102-119.

Horsky, Dan and Patrick Swyngedouw (1987), "Does It Pay to Change Your Company's Name? A Stock Market Perspective", *Marketing Science*, 6 (4), 320-335.

Jarrell, Gregg and Sam Peltzman (1985), "The Impact of Product Recalls on the Wealth of Sellers," *Journal of Political Economy*, 93 (3), 512-536.

Kalyanaram, Gurumurthy, William T. Robinson, and Glen L. Urban (1995), "Order of Market Entry: Established Empirical Generalizations, Emerging Empirical Generalizations, and Future Research", *Marketing Science*, 14 (3, Part 2 of 2), G212-G221.

Lane, Vicki, and Robert Jacobson (1995), "Stock Market Reactions to Brand Extension Announcements: The Effects of Brand Attitude and Familiarity", *Journal of Marketing*, 59 (January), 63-77.

Lee, Ruby P., Rajdeep Grewal (2004), "Strategic Response to New Technologies and Their Impact on Firm Performance", *Journal of Marketing*, 68 (October), 157-171.

Lilien, Gary L., Philip Kotler, K. Sridhar Moorthy. 1992. Marketing Models. Prentice-Hall, Inc. Upper Saddle River, NJ.

MacKinlay, A. Craig (1997), "Event Studies in Economics and Finance", *Journal of Economic Literature*, 35 (1), 13-39.

McWilliams, Abagail, Donald Siegel (1997), "Event Studies in Management Research: Theoretical and Empirical Issues", *Academy of Management Journal*, 40 (3), 626-657.

Peltzman, Sam (1981), "The Effects of FTC Advertising Regulation", *Journal of Law and Economics*, 24 (December), 403-448.

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.

Ryngaert, M, J Netter (1990), "Shareholder Wealth Effects of the 1986 Ohio Antitakeover Law Revisited: Its Real Effects", *Journal of Law, Economics, & Organization*, 6, 253-262.

Chapter 4

General Discussion

The two essays of my dissertation look at some important issues in e-commerce and retailing such as how retailers use the Internet to better communicate with and serve their customers.

The first essay examines the question of how retailers use the Internet to better communicate with consumers. I use permission email as an example to study this question. In my model of consumer response to emails, email offers are interpreted as having an information structure similar to financial options. Thus, communications design can be viewed from the perspective of crafting a low option price and a high expected upside. The consumer incurs the option price to gain the possibility, but not the obligation, of taking an offer once its value is revealed. The option price in the current study consists of the consumer's costs of attending to, processing, retaining, and following up on an email message. The expected upside is determined by the perceived mean and variance of the underlying email offer. My model motivates a set of hypotheses about communications design features that are tested in the empirical analysis.

The empirical approach described in the study provides an accessible way for practitioners to measure how email effectiveness is influenced by design features of the subject line, the email body, and the targeting and timing of the email campaign. Marketing communications managers can use the findings in this study to manage their email lists and generate permission email campaigns with high opening rates, high click-through rates, and lower opt-out rates.

In addition to interesting findings about permission email, the perspective of financial options developed in the study may also have applicability for other marketing communications and promotions design problems.

Despite of the aforementioned contributions, there are some limitations that need to be acknowledged and addressed regarding the first study of my dissertation.

One of the limitations is that the data used in the current study are aggregated at the email level and an individual recipient's responses to the email cannot be tracked. It would be interesting to build on Ansari and Mela's (2003) study of individuals' clicks on email links to consider other forms of individual response such as their email opening behavior, email list opt-out behavior, and email forwarding behavior. And individual level data will allow me to study the impact of consumer demographics on consumer response. Phillip and Suri's (2004) study of the impact of gender differences on the evaluation of promotional email is one example of this direction. Longitudinal individual level data could also provide an opportunity to examine consumer response to permission email over time. Results from such research could be used to help construct more effective email lists, customize email, and optimize the time interval between emails.

Another limitation of this study is in the measurement of email effectiveness, which is limited to variables such as the opening rate, the click-through rate, and the opt-out rate. Future research could look at some other measures such as daily online visit/purchase, offline store traffic and sales, brand awareness and advertising recall, and customer loyalty. These data will provide us richer information about the effectiveness of permission email.

The second study of my dissertation is an example of using the event study to evaluate the influence of marketing strategic actions on the performance of firms in financial markets. I conduct an event study to examine how adding online stores by brick-and-mortar retailers affects their stock price returns. I find that the 78 public retailers' announcements of online channel additions, on average, generated positive stock price returns. I also find that retailers in different sectors get different abnormal returns from the announcements.

Regression analysis of stock market reactions on the characteristics of firms and introduction strategies finds that investors are more optimistic toward the Internet channel additions of small retailers and retailers with direct channel experience. When influencing market reactions, firm characteristics also interact with retailers' channel introduction strategies.

Future research along the line of this study could focus on research questions such as the long-term marketing performance of retailers with online sales channels, consumer satisfaction with the new online channel, consumers' switching behavior between traditional channels and the new one, and cannibalization and synergy between the two channels. Among these potential research opportunities, I think cannibalization and synergy between the two channels is the most exciting and promising topic. At the end of my dissertation, I would like to discuss this topic in more detail.

As the latest multiple-channel innovation, the dual-channel of both a bricks-and-mortar store and an Internet store is becoming increasingly popular. The discussion of possible conflicts and synergy between these two channels is naturally an important topic to both practitioners and academics. One concern is channel conflict. Alba et al. (1997) argue that online activities may cannibalize a retailer's offline business and hurt profits. Follow-up research, however, has not found evidence for this concern. Geyskens, Gielens, and Dekimpe's (2002) study of the market valuation of Internet channel additions assesses the impact of adding an Internet channel on a firm's stock market return. They conclude that, on average, the stock market reacts positively to Internet channel additions. Deleersnyder et al. (2002) examines the long-term effect of the Internet channel addition on a firm's performance. They find that the added Internet channel in the newspaper industry rarely cannibalizes the traditional one (in terms of circulation and advertising revenues). Biyalogorsky and Naik (2003) study the effect of online activities on offline sales. They examine the extent to which online sales cannibalize offline sales and whether online activities build online equity for the firm. They find that online sales do not significantly cannibalize offline retail sales and that the online activities build long-term online equity. Lee and Grewal (2004) find that the adoption of the Internet as a communication channel positively influences firm performance (measured as Tobin's q). The adoption of the Internet as a sales channel, however, seems to matter only to firms that have preexisting catalog operations. The second study of my dissertation examines the immediate response of stock market to retailers' online channel additions. I find that retailers significantly increase their performance potential (measured as abnormal stock price returns) by adding Internet sales channels. All these studies support or at least partly support Frazier's (1999) contention that there are occasions when channels might complement each other.

A review of the previous literature indicates that it is promising to look at interactions between bricks-and-mortar and Internet channels from a new perspective - communication and promotion effects across channels. The following effects will be interesting: (1) the impact of traditional forms of communications and promotions on Internet channel performance, (2) the impact of Internet communications and promotions on traditional store performance, and (3) the mutual impact of online and offline store performance on each other.

The evaluation of promotion effects is a well-developed area. In marketing literature, the common approach to assess consumer promotions has been to compare sales or market share before, during, and after a promotion activity. Examples include Rao-Lilien's (1972) model, Little's (1975) BRANDAID promotional model, and Narasimhan's (1984) model of coupons. One limitation of this research is that it relies on a single indicator of retail performance as measured by sales or market share. Lam et al. (2001) consider multiple indicators of retail performance, including front traffic, store entry ratio, closing ratio, and average spending. They use these indicators to assess three categories of promotion effects: attraction effects, conversion effects, and spending effects. Their research is made possible by the developments of in-store technologies.

For online stores, consumer browsing and shopping path tracking technologies make it possible to measure multiple-level online promotion effects, which are similar to those used by Lam *et al.* (2001). In addition, with access to both online and offline sales and traffic data, we can further examine the interactions between the two channels. Findings about these interactions could provide some managerial guidelines for marketing communications managers and help them best design marketing communications under an environment of both online and offline operation.

Bibliography

Alba, Joseph, John Lynch, Barton Weitz, Chris Janiszewski (1997). Interactive Home Shopping: Consumer, Retailer, and Manufacturer Incentives to Participate in Electronic Marketplaces. *Journal of Marketing*, 61 (3), 38-53.

Ansari, A., & Mela, C. F. (2003). E-customization. *Journal of Marketing* Research, 40 (2), 131-145.

Biyalogorsky, Eyal, Prasad Naik (2003). Clicks and Mortar: The Effect of On-line Activities on Off-line Sales. *Marketing Letters*, 14 (1), 21-32.

Deleersnyder, Barbara, Inge Geyskens, Katrijn Gielens, Marnik G. Dekimpe (2002), "How Cannibalistic is the Internet Channel? A Study of the Newspaper Industry in the United Kingdom and the Netherlands", *International Journal of Research in Marketing*, 19, 337-348.

Frazier, Gary L. (1999). Organizing and Managing Channels of Distribution. *Journal of the Academy of Marketing Science*. 27 (2) 226-240.

Geyskens, Inge, Katrijn Gielens, Marnik G. Dekimpe (2002). The Market Valuation of Internet Channel Additions. *Journal of Marketing*. 66 (April) 102-119.

Lam, Shunyin, Mark Vandenbosch, John Hulland, Michael Pearce (2001). Evaluating Promotions in Shopping Environments: Decomposing Sales Response into Attraction, Conversion, and Spending Effects. *Marketing Science*. 20 (2) 194-215.

Lee, Ruby P., Rajdeep Grewal (2004), "Strategic Response to New Technologies and Their Impact on Firm Performance", *Journal of Marketing*, 68 (October), 157-171.

Little, John D.C. (1975). BRANDAID: A Marketing Mix Model, Part I: Structure; Part II: Implementation. *Operations Research.* 23 (4), 628-673.

Phillip, M. V., & Suri, R. (2004). Impact of Gender Differences on the Evaluation of Promotional Emails. *Journal of Advertising Research*, 44(4), 360-368.