



What's in a "Name"? Impact of Use of Customer Information in E-Mail Advertisements

Author(s): Sunil Wattal, Rahul Telang, Tridas Mukhopadhyay and Peter Boatwright

Source: *Information Systems Research*, September 2012, Vol. 23, No. 3, Part 1 of 2 (September 2012), pp. 679-697

Published by: INFORMS

Stable URL: <https://www.jstor.org/stable/23276480>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Information Systems Research*

What's in a "Name"? Impact of Use of Customer Information in E-Mail Advertisements

Sunil Wattal

Fox School of Business, Temple University, Philadelphia, Pennsylvania 19122, swattal@temple.edu

Rahul Telang

Heinz College, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, rtelang@andrew.cmu.edu

Tridas Mukhopadhyay, Peter Boatwright

Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213
{tridas@cmu.edu, boatwright@cmu.edu}

In this study, we examine how consumers respond to firms' use of two types of information for personalization: product preferences and name. We collect a unique data set of over 10 million e-mail advertisements sent by a website to over 600,000 customers who could buy the advertised products from the online merchant. We estimate a two-stage hierarchical model using Bayesian analysis to account for observable and unobservable consumer heterogeneity. Our analysis suggests several interesting results regarding consumers' responses to firms' use of information. When firms use product-based personalization (where the use of information is not explicitly mentioned), consumers respond positively. On the other hand, consumers respond negatively when firms are explicit in their use of personally identifiable information (i.e., a personalized greeting). We also find that negative responses to personalized greetings are moderated by consumers' familiarity with firms. The main contribution of this study is that it not only indicates the economic benefits of personalization in e-mails but also highlights consumers' concerns over the use of information in personalization.

Key words: personalization; privacy; information use; hierarchical Bayesian model

History: Lorin M. Hitt, Senior Editor; Gautam Ray, Associate Editor. This paper was received on August 2, 2008, and was with the authors 15 months for 3 revisions. Published online in *Articles in Advance* November 3, 2011.

1. Introduction

Advances in information technology have enabled firms to use information about consumers at an unprecedented level. Firms use consumers' information to tailor interactions on an individual basis across sales, marketing, and customer service. For example, Amazon.com recommends books and other products to customers based on their prior browsing and purchase behavior. Firms such as Coca-Cola and Hewlett-Packard (HP) target e-mail advertisements to customers based on preferences and personal information.¹ Such use of information is an attempt by firms to understand each customer better and replicate a shopping environment reminiscent of the preindustrial revolution era when most businesses were neighborhood stores and owners greeted each customer by name and knew what each customer

liked. In the process, not only do customers benefit by receiving product or services that match their preferences closely but firms also benefit because targeted marketing enables them to charge higher prices and make more profits (Chen et al. 2001). Therefore, although the use of information for personalization sounds like a win-win proposition for firms and customers alike, such a view makes an important assumption, i.e., that consumers always see value in their information being used in exchange for personalized experiences. However, consumers may not see such value in certain scenarios. For example, Malhotra et al. (2004) argue that consumer information is a *double-edged sword* that can enhance consumer utility and at the same time cause privacy violation. The hue and cry over cases such as Lotus, Blockbuster (Culnan 1993), and DoubleClick (Culnan and Bies 2003) highlights the sensitive nature of using personal information in marketing.

In recent times, consumers are becoming increasingly aware of firms' information collection and use. Numerous types of cybercrimes over the Internet have heightened these concerns. Bot networks,

¹ Coca-Cola and HP use e-mail marketing services from <http://www.yesmail.com> to send targeted e-mail advertisements to their customers (as mentioned on <http://www.yesmail.com/Company/Overview.aspx>—accessed on January 15, 2011).

Trojans, phishing, data theft, identity theft, spyware, and credit card fraud are making daily headlines that scare many readers.² According to a Gartner (2007) report, around 15 million consumers experienced identity theft in 2006 and lost an average of \$3,257 each. Many states have explicitly passed laws to protect consumers from personal information breaches. Financial institutions regularly warn customers not to respond to e-mail requests for personal information. In light of these concerns, whether a customer would always respond positively to e-mails from an online merchant that bears her personal information (such as name) is unclear. Apart from personal information, firms also routinely use consumers' product preference information to tailor the content of advertisements in e-mails. Prior empirical work (such as Ansari and Mela 2003) has not explored in detail how firms' use of different types of information impacts consumer behavior.

One would expect some factors such as type of information used and familiarity of customers with firms to moderate consumers' responses to firms' use of information. In this research, we make a distinction between personally identifiable information (PII) and other information. According to the Federal Trade Commission (2000), PII constitutes any data that can be used to identify, locate, or contact any individual (such as name, address, telephone numbers, and e-mail address) as opposed to demographic and behavioral data (such as age, income, and ethnicity that characterize a person without making direct identification). Also, a customer's familiarity with a vendor would mitigate customer concerns about misuse of information. A customer who makes repeat purchases with a vendor or visits a vendor's website multiple times is less likely to be concerned when the vendor uses her name or other information for personalization. For example, *customers who regularly visit websites such as Amazon.com or Netflix.com know by experience that these websites will make recommendations based on their past interests*. In offline scenarios, too, it is reasonable to expect people to feel uncomfortable when complete strangers can access their personal information (Posner 2008).

With this motivation, we propose our research questions as follows: (1) How do consumers respond

when firms use their information to offer personalized products or services via e-mail? (2) How do these responses to firms' use of information differ for different types of information? (3) How does familiarity moderate consumers' responses to firms' use of information in e-mail advertising?

We collect a large data set of responses to 10 million e-mails sent to over 600,000 customers from a Web-based firm. The firm uses two types of information to personalize its e-mail advertisements: (1) e-mailing customers about products in which they expressed interest in the past (labeled *product-based personalization*) and (2) greeting customers by name (labeled *personalized greeting*). An important aspect of this data is the setting in which it was collected because customers received personalized as well as nonpersonalized e-mails on a pseudorandom basis. The response rates would reflect any concerns about personalization. We define the response to an e-mail advertisement in terms of an initial action (opening the e-mail) and four subsequent levels of response: unsubscribing, no action, click-through, and purchase. We test our hypotheses using a multi-stage ordered probit model in a hierarchical Bayesian framework to estimate individual-level parameters.

Our main results are as follows. We show that, when the use of information for personalization is not evident to customers, product-based personalization leads to a positive response. On the other hand, we also find a negative and significant response to personalized greetings. The level of a consumer's familiarity with a firm moderates this negative response: We find that customers who made prior purchases with a firm respond less negatively to personalized greetings. We also find that a higher level of current activity by a consumer (meaning a higher probability of click-through to e-mail advertisements or purchase) is associated with a less negative response to personalized greetings.

A key contribution of our paper is to show that personalization does not always lead to positive responses from consumers but, rather, that consumer response varies based on familiarity with a firm as well as the type of personalization. Our finding that consumers respond negatively to personalized greetings is unique and has not been highlighted in any empirical work. Whereas prior work (mainly surveys and experimental studies) examined the personalized content made available on merchants' websites or in controlled experiments, we study personalized e-mails sent directly by merchants to consumers. Additionally, a unique aspect of our study is that consumers had to respond to personalized offers with their own money on the line. Prior work has studied situations where customers had no monetary risks.

² Bot Networks, or Botnets for short, indicates a group of computers used for malicious purposes, often without the knowledge of the owners (<http://searchsecurity.techtarget.com/definition/botnet>). Trojans, or Trojan Horses is a type of computer program that appears to perform a certain function but actually performs other, often malicious, actions (<http://articles.winferno.com/antiviruses/trojan-horse>). Phishing is a type of a scam in which hackers send an email or pop-up message that claims to be from a business or organization that consumers may deal with in an attempt to steal personal information (<http://www.ftc.gov/bcp/edu/pubs/consumer/alerts/alt127.shtm>).

Our results also demonstrate that the type and manner of information used and consumer familiarity with vendors moderate consumer response to personalization. We also use a methodology that allows us to estimate parameters at an individual level and thus allows us to estimate individual user response. This is key not only to illustrating heterogeneity across individuals but also to gleaning interesting managerial insights. For example, our model can predict what type of information to use for personalization when sending an e-mails to individual customers.

The rest of the paper is organized as follows. In §2, we present a literature review and, in §3, we propose a conceptual model of how consumers are likely to respond when a firm uses different types of information for personalization. In §4, we describe our data in detail. We present our estimation model in §5, and in §6, we highlight the results of our model. Finally, in §7, we present our conclusions, discuss the limitations of our study, and propose areas of future research.

2. Literature Review

In this paper, we focus on analyzing how consumers react to firms' use of information through personalized e-mails. According to a survey by McKinsey, over 80% of firms worldwide use e-mail campaigns to reach customers (*McKinsey Quarterly* 2007). The Forester Group estimates that spending on e-mail direct marketing is expected to reach \$2 billion by 2014.³ Unfortunately, the rise of spam in recent years has cast a long shadow on direct e-mail marketing by legitimate firms.⁴ Additionally, the use of e-mails for phishing and identity theft has also become commonplace. In this context, we examine consumers' reactions to personalized e-mail messages.

Several researchers have examined personalization. Murthi and Sarkar (2004) summarize existing literature on personalization in the management sciences area and identify potential analytical and empirical research issues. Chen and Hitt (2002) find no evidence of beneficial consumer behavior through personalization in their empirical study with online brokerages.⁵ Prior analytical research (Chen et al. 2001, Shaffer and Zhang 1995, Dewan et al. 2003, Wattal et al. 2009) captures the strategic impact of personalization in terms of price competition and strategic entry-exit. A limitation of this line of research is that these studies do not model any consumer disutility that might arise from firms' use of information, and they assume

that consumers always derive a higher utility from personalization.

Researchers have also explored potential reasons for consumers' concerns over interacting online with firms. Ba and Pavlou (2002) show that consumers consider utility and trust important to shaping their decisions to purchase online. Two major consumer concerns about purchasing online are security and privacy (Gartner 2007). Although consumers want good deals and superior service, they are also leery about online merchants and their motivations. Although 75% of consumers see value in e-mails from firms with whom they do business, a much smaller fraction (17%) is willing to receive e-mail from unknown merchants (Laudon and Traver 2009).

Personalization and firms' use of information have also been closely associated with privacy concerns. Smith et al. (1996) suggest that firms' use of information can cause privacy concerns related to unauthorized secondary use of information, unauthorized collection of information, and the feeling that firms profit from consumer information. Chellappa and Sin (2005) show that privacy concerns negatively impact consumer willingness to use personalization services. Research has also shown that consumer willingness to provide information for personalization increases with information transparency (Awad and Krishnan 2006), trust, and personal interest (Dinev and Hart 2006).

However, most of the prior work on personalization and consumer response to firms' use of information was done through surveys and controlled experiments that measured consumers' preferences on such use for personalization (Lai et al. 2003, Tam and Ho 2003). As Strandburg (2005) points out, concerns such as privacy may not be accurately reflected in such settings because of the gap between consumers' claimed privacy levels in surveys and their actual behavior in taking steps to protect their own privacy. Participants willingly take part in surveys and experiments and may not be wary of providing names and other relevant information to an experimenter. They also may not have significant concerns if the experimenter uses their information in the experiment. Similarly, Acquisti et al. (2009) show that a disconnect exists between consumers' stated privacy concerns in surveys and their privacy-protecting behavior in real life. Thus, a unique aspect of our study is that consumers in our data set had to make decisions that had real financial consequences in that they had to pay if they accepted offers. Moreover, the consequences arising from the misuse of information in actual e-commerce transactions are likely to be worse than those in academic experiments. We know of no prior work on online personalization where consumers had to

³ See Parfeni (2009).

⁴ See *Australian* (2008).

⁵ Personalization was not the central theme of that paper.

respond to personalized offers with their own money on the line.

One prior study close to our research was done by Tam and Ho (2006). They describe two studies—a lab experiment and a field study—to test the effects of different types of personalization, namely, self-reference and content relevance. In the lab experiment, participants navigate through a hypothetical online shop for 12 minutes, followed by a 5-minute distraction phase, a memory recall test, and a questionnaire. Tam and Ho (2006) find that the effect of using consumers' names in banner advertisements is positive in the experimental setting. In the field study, they examine the effect of recommendations when subjects come to specific websites, log in using their user IDs and passwords, and download free music. Our study differs from those of Tam and Ho (2006) in the following ways:

We note the subtle difference between a subject in a controlled experiment visiting a website on her own and being greeted by name, and a consumer receiving an e-mail and being greeted by name. The distinction is between an e-mail that "pushes" the marketing message to the consumer and a website that consumers visit on their own and are therefore "pulled" in to its content. Tam and Ho (2006), for example, used the latter vehicle for their lab experiments. Their subjects also did not bear any financial risks.

In addition, subjects participated in the field study for six weeks for a specific product type (music) immediately after voluntarily consenting to do so, which reduced any chance of problems with memory recall. In other words, subjects knew the identity of the website and voluntarily identified themselves before receiving their personalized recommendations for music. Therefore, any concerns with information misuse are likely to be less in the settings studied by Tam and Ho (2006) than in advertisements by a commercial entity that collects consumer information and later uses the same for potential financial gain (through advertisements).

3. Conceptual Model and Hypotheses

We model the utility that a consumer derives from an e-mail advertisement as a combination of two main effects: (1) economic benefit, where the consumer compares the cost of responding to an e-mail in terms of time (spent opening the e-mail or clicking through) or money (making a purchase) with the expected benefit, and (2) utility associated with psychological cues, as suggested in the behavioral economics literature (Slovic et al. 1977, Bertrand et al. 2005). Therefore, one would expect consumer response to e-mail advertisements to depend on both economic and psychological cues triggered by different elements of the e-mail

advertisements. Mathematically, we can express this as a preliminary equation:

$$U_N = U_E + U_P, \quad (i)$$

where U_N is the net utility that the consumer derives from the e-mail, U_E is the economic benefit from the e-mail, and U_P is the utility derived from the psychological cues associated with the e-mail. We now look at different types of information separately.

3.1. Product Preference Information

In e-mail advertising, product personalization refers to a firm sending e-mails to customers about products that fit customers' preferences. Firms can send such e-mails in two ways: *explicit*, where the firm includes a statement such as "you liked product x in the past; you might also like product y ." The other method is *implicit*, where the firm features products in an e-mail advertisement based on a customer's preferences but does not disclose the use of customer information and the customer cannot infer that the product featured in the e-mail is based on her product preferences. We consider the latter type of product-based personalization in this paper.

In *implicit* product-based personalization, the consumer may not notice that product personalization has taken place and, hence, the psychological cues associated with product-based personalization are minimal in this scenario. On the other hand, it is well established that customers' past preferences are strong predictors of future purchases (Rossi et al. 1996). In online shopping, consumers are expected to get higher economic utility as a result of product-based personalization because they can save time and effort by finding the right information easily (reducing search costs). Product-based personalization enables firms to learn customers' preferences over time and recommend products based on this information. Tam and Ho (2006) show that users are more likely to accept offers associated with relevant Web content. In other experimental environments with controlled settings, research has shown that personalization based on targeted product recommendations leads to higher response rates. For example, such experiments were conducted with subjects downloading online music (Tam and Ho 2003), viewing online news (Lai et al. 2003), and locating items using search engines (Pitkow et al. 2002). Ansari and Mela (2003) find that customizing the links within an e-mail can potentially improve click through rates by 62%.

In terms of equation (i), the utility that a consumer derives from an e-mail with product-based personalization is

$$U_N^{PP} = U_E^{PP} + U_P^{PP}, \quad (ii)$$

and that of an e-mail without product-based personalization is

$$U_N^{nPP} = U_E^{nPP} + U_P^{nPP}. \quad (\text{iii})$$

The superscripts *PP* and *nPP* denote e-mails with product-based personalization and no product-based personalization, respectively. Comparing (ii) with (iii), we can infer the following: From the discussion on product preference information, the economic benefit of product-based personalization is clearly positive; that is, $U_E^{PP} > U_E^{nPP}$. For implicit product-based personalization, the utility from psychological cues associated with the use of consumers' information is likely to be minimal; that is, $U_P^{PP} = U_P^{nPP}$. The net result is that consumers are likely to derive a higher net utility from e-mails with product-based personalization where the elements of personalization are not explicit ($U_N^{PP} > U_N^{nPP}$). Therefore, the first hypothesis we propose is as follows.

HYPOTHESIS 1 (H1). *Product-based personalization in e-mail advertisements, where firms' use of information is not evident to the consumer, leads to positive customer response.*

3.2. Personalized Greetings

Using equation (i), the utility that consumers derive from an e-mail with a personalized greeting is $U_N^{PG} = U_E^{PG} + U_P^{PG}$, and that from an e-mail with no personalized greeting is $U_N^{nPG} = U_E^{nPG} + U_P^{nPG}$. The superscripts *PG* and *nPG* denote e-mails with a personalized greeting and with no personalized greeting, respectively.

The economic benefits of an e-mail with a simple personalized greeting ("Dear John") are likely to be the same as those of a similar e-mail without a personalized greeting (provided the rest of the e-mail is the same in terms of product, price, and message); that is, $U_E^{PG} = U_E^{nPG}$.

However, the psychological cues associated with personalized greetings are likely to play a significant role in predicting whether customer response would be different for such e-mails. The use of a name involves the concept of self, as mentioned in psychology literature (Rogers et al. 1977). Self refers to everything associated with an individual. Rogers et al. (1977) define self as a vague idea about who a person thinks he or she is, and it includes information relevant to an individual. Bugental and Zelen (1950) measure self in a survey by asking individuals a single question: "Who are you?" Their study suggests that self includes information such as name (e.g., "I am Joe"), occupation ("I am a painter"), social status ("I am a veteran"), gender ("I am a boy"), nationality or race ("I am a Japanese-American"), among others. Tam and Ho (2006) suggest that name is a self-referent concept associated with better recall and

a higher probability of accepting an offer. In marketing literature, too, the use of consumers' names has been associated with increased sales (Levy and Weitz 1992). However, in online environments, additional factors may come into play that also influence the psychological cues associated with the use of consumers' names. In online environments, consumers transact with many different merchants and leave personal information with them (voluntarily or sometimes involuntarily). In turn, they also receive a large number of e-mails from many such merchants. Because consumers tend to be more sensitive about personally identifiable information, a personal greeting in an e-mail is likely to capture the attention of the reader. Given the high level of cyber security concerns about phishing, identity theft, and credit card fraud, many consumers would be wary of e-mails, particularly those with personal greetings.

The use of name may also trigger concerns associated with privacy. In online marketing situations, a customer would immediately notice that a firm is using her personal information for product promotions. This would bring a range of privacy concerns to the fore. First, the *collection* issue discussed in the literature (e.g., Nowak and Phelps 1995) may come up. The customer may ask, "How did the firm get my name?" Second, she may question, "Does the firm have permission to use my name in e-mail advertisements?" This brings *unauthorized secondary use* to bear (Malhotra et al. 2004). Finally, she may wonder, "How interested am I in the message?" The issue of *fair exchange* is activated in this context (Milne and Gordon 1993).

Which of these effects will dominate consumer reactions to personal greetings? Would the use of names in e-mails induce the self-referent concept and lead to positive reactions? Or would consumers experience discomfort and concern about security and privacy on seeing their names in e-mail advertising? In most cases, customers themselves volunteer information to firms and "agree" to the policies on how firms might collect and use information about consumers. Wouldn't consumers recall that they authorized (by accepting the privacy policy of a website) the manner in which the website can use their personal information and therefore should not be vexed by personal greetings? If consumers were rational, we would expect them to remember the privacy policies of Web firms and the terms and conditions they agreed to in the past and thus not to react negatively. However, a rich literature on behavioral decision theory points to many ways that humans may stray from the rational model. For example, if a consumer cannot estimate the probability of a vendor abusing her personal information, she may substitute instances in her memory or even imagine instances as substitutes

(Kahneman and Tversky 1982). Alternatively, people who lack confidence in their judgments may play safe and judge e-mails with personal greetings as too close to preceding stimuli (Poulton 1989). In many cases, consumers may recall media reports on identity theft or even the infamous Nigerian e-mail scam and thus come to suspect a Web vendor's motive. (Consumers may not exhibit a negative reaction to reputable Web firms. However, the incidence of cyber fraud renders this to empirical validation.) Evidence also shows that consumers rarely read privacy policies, which are incomprehensible to a majority of Internet users (Jensen and Potts 2004).

In summary, we believe that the net utility derived by consumers from psychological cues associated with personalized greeting can be positive or negative. Although one may expect an intuitive positive effect, one can also argue for the possibility of a negative response when a consumer sees personal information such as his name in an e-mail (i.e., U_p^{PG} can be higher or lower than U_p^{nPG}). If consumers expect the use of personalized greetings as a norm in e-mail advertisements, then we should observe no significant reaction (Howard et al. 1995). Whether the net effect is positive, negative, or neutral remains an open empirical question. Therefore, the question of how consumers react to online firms' personalized greetings in e-mail advertisements remains an open empirical question and has implications for both academic theory and practice.

3.3. Role of Familiarity

Although we argue that the impact of personalized greeting in an e-mail is not easy to hypothesize, the psychological cues associated with personalized greetings are likely to be different for different customers. Therefore, consumers' characteristics are likely to play an important role in their responses to personalized greetings. One such characteristic is familiarity with a firm. Familiarity is a proxy for learning, and an individual's tendency to substitute incomplete information with disapproval will be much lower for individuals who have had repeated interactions (especially purchases) with a firm and know it is not a phishing scheme, a too-good-to-be-true offer, or an outright fraud. Therefore, we examine the impact of familiarity on psychological cues associated with personalized greetings. Prior literature (Gefen 2000) establishes that familiarity is associated with increased levels of trust. For example, if a user has transacted with a firm in the past, then familiarity may alleviate the negative effect of a personalized greeting by the customer will realize that the firm is a genuine seller. Familiarity may accentuate the positive effect of a personalized greeting because customers are likely to be more amenable if they are

familiar with the entity using the name—akin to an acquaintance rather than a stranger greeting them by name. In either case, we expect a customer who is more familiar with a firm to experience a more positive (or less negative) change in utility as a result of a personalized greeting. Mathematically, we can write this as $U_p^{i,PG} > U_p^{j,PG}$, where $U_p^{i,PG}$ and $U_p^{j,PG}$ denote the net increase in utility as a result of psychological cues associated with personalized greetings for customers i and j , respectively, where customer i is more familiar with a firm than is customer j . Therefore, we propose our next hypothesis as follows.

HYPOTHESIS 2 (H2). *Familiarity will positively affect customer reaction to personalized greetings in e-mail advertisements.*

However, the arguments advanced by this hypothesis are applicable only when a consumer recognizes the personalization and information use. Thus, we do not hypothesize how familiarity moderates the response to *implicit* product-based personalization.

4. Data

We collected data for this research from a Web-based firm that acts as a distributor for a variety of products including long distance phone services, cellular plans, electricity, gas, health insurance, Internet connections, and mortgage lending. An advantage of data from such a diverse product set is that the disutility resulting from information overload is more prominent, and hence, the potential benefit from personalization is larger than that for a single-product firm. The firm wishes to remain anonymous as a precondition for sharing these data with us, so we refer to it as firm A.

Firm A mainly relies on e-mail advertisements to inform its customers about its products as well as to announce new offers. We collected data on more than 10,000,000 e-mails to about 600,000 customers over a nine-month period. For analytical tractability, we selected 19,661 customers at random for our analysis; these customers received 364,646 e-mails. We selected our random sample as follows: We used a random number generator to generate 50,000 numbers between 1 and 600,000. Comparing these randomly generated numbers with customer IDs in our data yielded 33,196 unique customers who received e-mails during our data collection period. Some customers may have unsubscribed before our data collection started and did not receive any e-mails. Further, the firm acquired 13,535 customers in our sample from an external database just prior to our data collection. We dropped all such customers from our analysis because prior information about these customers might not have been reliable. The remaining sample yielded 19,661 customers. For each e-mail

advertisement it sent out, firm A kept a record of the featured product, who received the e-mail, and what actions each user took after receiving the e-mail. For example, the firm kept track of whether users opened the e-mail (firms can track whether users open e-mail through active links in message bodies). If a user opened an e-mail, the firm could also record if he took any one of the following actions: unsubscribed (he can click the unsubscribe link at the bottom of the e-mail), took no action, clicked through, or made a purchase. Firm A also kept track of consumer preferences by grouping customers into pools. For example, if a customer viewed an offer about natural gas at some point,⁶ she was placed in the “natural gas pool.” A customer could belong to more than one pool depending on her interests.

The firm sent e-mail advertisements to its customers in the form of organized campaigns, where large groups of customers received the same e-mail. These organized campaigns were of three main types:

1. Customers in a “pool” received e-mails about a product. For example, a customer in the “long distance” pool received an e-mail about long distance phone plans. In our data analysis, we consider this as an instance of product-based personalization for all customers who received it.

2. At other times, customers in a pool received e-mails about a different product. For example, customers in the “long distance” pool received an e-mail about mortgage plans. In our analysis, we consider this e-mail an instance of product-based personalization for customers who are in both the “long distance” and “mortgage” pools and an instance of “no product-based personalization” for customers only in the “long distance” pool.

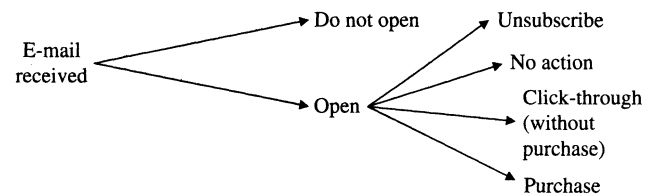
3. Finally, in some e-mail campaigns, all customers in a firm’s database received an advertisement about a product. For example, a firm sent an offer about a product (e.g., “satellite TV”) to all customers. Although the firm did not make an effort to target any particular customer, we still consider this e-mail an instance of product personalization for customers in the “satellite TV” pool and nonpersonalization for other customers.

The firm did not make explicit mention of this personalization in any e-mail; i.e., it did not use comments such as “Here are some recommendations for you.” To customers, these e-mails would appear as regular e-mail advertisements.

The firm also used consumers’ names to personalize its e-mail advertisements. The firm chose some e-mail campaigns at random and attached greetings

⁶ The customer had to log in to the website to view the offer completely. She was placed in a pool if she viewed an offer after logging in.

Figure 1 Two Stages of Consumer Decision Making



such as “Dear X” or “Dear Y.” We refer to this type of personalization as *personalized greetings*. As Figure A.1 in Appendix A shows, e-mails with personalized greetings have salutations at the top; e-mails without personalized greetings do not have any salutations.

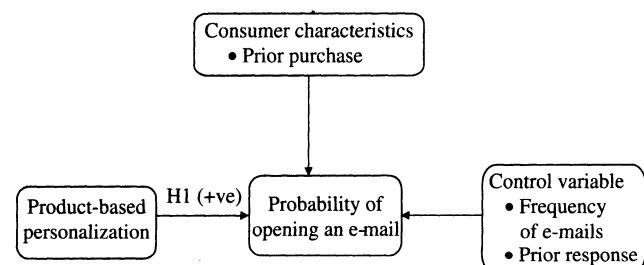
In our data setting, a consumer faces decisions at two levels: (1) when she receives an e-mail, she decides whether to open it; and (2) if she opens the e-mail, she can take any one of the following actions: unsubscribe, take no action, click through (without making a purchase), or make a purchase. A consumer’s decision process is represented in Figure 1.

In stage 1 of decision making, customers observe only the subject line of the e-mail but not the content. The subject line contains the product advertised in the e-mail. For example, a typical subject line reads “Protect your computer with Norton AntiVirus at an unbelievable price.” Customer names do not appear in the subject line; therefore, only *product-based* personalization can affect the consumer response of opening the e-mail. A positive response (as per H1) at this stage implies that consumers are more likely to open such e-mails. Personalized greetings do not play a role at this stage because consumers do not know if such e-mails contain personalized greetings.

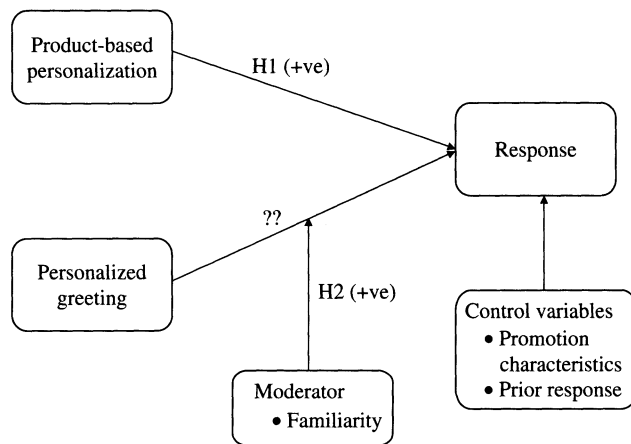
Figure 2 summarizes our model in the first decision phase, that is, when a customer decides whether to open an e-mail (stage 1). We use response to prior e-mails and frequency of e-mails as controls in our model. Berlyne (1970) suggests that the advertisement response function follows an inverted U shape with increases in advertising frequency.

If a customer opens the e-mail in stage 1, she has to make another decision in stage 2. At this stage, she can click the unsubscribe link, take no further

Figure 2 Stage 1 Model—Before Opening an E-mail



Note. +ve, positive.

Figure 3 Stage 2 Model for Response to Use of Consumer Information in E-mails

Note. +ve, positive.

action after opening, click through (but not make a purchase), or make a purchase. A clear ordinal ranking exists among these actions: If the customer utility from the e-mail is below some threshold, she may unsubscribe. If she receives a high level of utility, she may buy the product featured in the e-mail. A positive response to personalization would mean shifting a consumer's propensity toward the purchase direction of the ordered responses. Figure 3 summarizes the model in the second decision stage (stage 2).

When a customer opens an e-mail, she also becomes aware of other e-mail characteristics that are likely to impact her response to the e-mail. Prior studies on the effectiveness of advertisements in marketing literature have determined that consumer response to advertisements depends on the advertisement's characteristics (Walters and Rinne 1986). Therefore, we use promotion characteristics as control variables. We measure an e-mail's characteristics in three ways: (1) whether the advertisement mentioned the product price, (2) whether the e-mail offered a free gift or a discount, and (3) whether the e-mail offered a comparison with competitors' prices. E-mail characteristics (discount offers, comparisons with competitors, etc.) are applicable to both personalized as well as nonpersonalized e-mails. For example, some personalized e-mails may feature discount offers, other personalized e-mails may feature comparisons with competitors, and so on. Similarly, some nonpersonalized e-mails may feature discount offers and so on. As in stage 1, we use prior response as a control variable in stage 2.

What makes our data set interesting is that the firm sent a mix of personalized (either product-based e-mail or personalized greetings, or both) and nonpersonalized e-mails to its consumers more or less at random. Our discussion with the firm confirmed that the

decision whether to offer product personalization or personalized greetings in e-mail advertisements was made irrespective of product categories or customers' previous reactions. Such a pseudorandom experiment alleviates the significant selection concerns that one would encounter in a typical field study. Another interesting feature of our data is that we measure responses at multiple levels using both nonfinancial (opening, click-through, and unsubscribing) and financial (purchases) measures. Appendix A shows a sample e-mail that contains a personalized greeting. Appendix B gives detailed descriptive statistics from stages 1 and 2.

5. Estimation Model

We propose a choice model based on the random utility theory (McFadden 1974). When consumers receive e-mail advertisements about products, their utility depends on the messages featured in the e-mails. The dependent variables in stages 1 and 2 are binary and ordinal, respectively. Therefore, we use a probit model for stage 1 and an ordered probit model for stage 2. In addition, we modify these basic choice models to control for observable and unobservable consumer heterogeneity for the following reasons. First, in scenarios involving panel data where more than one data point is available for some individuals, the observations in the data are not independent and identically distributed (i.i.d.). The observable and unobservable characteristics of an individual are likely to influence observations pertaining to a given customer. Second, incorporating consumer heterogeneity enables us to study how different types of consumers differ in their responses to personalized greetings. For example, some consumers may not open any e-mail advertisements whereas others open most of the e-mail advertisements they receive. Bargain-hunting consumers may find more value in e-mails that feature comparisons with competitors. Prior literature also shows that response to firms' use of information depends on individual attitudes and demographic variables (Culnan 1993, Wang and Petrisson 1993). For these reasons, incorporating consumer heterogeneity in this setting is essential in order to obtain unbiased and consistent estimates of the variances (and unbiased *t*-statistics) (Hsiao 1986). Finally, in line with our hypotheses development (H2), we need to test whether observable consumer characteristics such as familiarity moderate the impact of personalized greetings. To achieve this end, we propose a model that incorporates both observable and unobservable consumer-specific heterogeneity (Chintagunta et al. 1991, Gonul and Srinivasan 1993).

We account for consumer heterogeneity by estimating individual-level parameters, which can be

viewed as draws from a super-population distribution (often referred to as the mixing distribution). We estimate a continuous heterogeneity model, where the mixing distribution is continuous (e.g., normal) and individual-specific parameters are drawn from the distribution. Monte Carlo Markov chain (MCMC) methods are commonly used to estimate the exact posterior distribution of individual-specific parameters (e.g., Rossi et al. 1996). We clarify that, even though we have two stages in consumer decision making, we do not need to estimate them jointly. Stage 2 is conditional on stage 1; therefore, we can simply multiply the estimates from stage 2 with the probability of opening e-mail (stage 1) to get joint probabilities.⁷

5.1. Stage 1

To test hypotheses H1 and H2, we propose a choice model based on the random utility theory (McFadden 1974). The utility assigned to consumer n for e-mail j can be specified as

$$U_{jn} = \theta_n \cdot X_{jn} + \varepsilon_{jn} \quad (1)$$

for $j = 1, 2, \dots, J_n$ and $n = 1, 2, \dots, N$, where J_n is the total number of e-mails that consumer n receives, N is the total number of consumers, and θ_n are the individual-specific parameters (which also include an intercept) to be estimated. An individual-specific parameter estimate means that we estimate the impact of X (our variables of interest) for each individual separately, unlike in a pooled regression where we estimate only the average impact across all consumers. The observable dependent variable y can take two discrete values (1 if the consumer opens the e-mail, and 0 otherwise). Specifically, $y_{jn} = 0$ if $U_{jn} < 0$, and $y_{jn} = 1$ otherwise. We assume that the stochastic component ε_{jn} is distributed normally with a mean 0 and variance 1. Under these assumptions, the probability that customer n selects choice i for e-mail j is given by a probit model:

$$P_i(U_{jn}, \theta_n, X_{jn}) = [1 - \Phi(\theta_n \cdot X_{jn})]^{y_{jn}} \cdot [\Phi(\theta_n \cdot X_{jn})]^{(1-y_{jn})}. \quad (2)$$

The independent variables X of interest in our model are as follows:

- *PRODUCT*, i.e., whether an e-mail features product-based personalization. This variable is 1 if so, and 0 otherwise. A positive coefficient for this variable would provide support to H1.
- *OPEN_{t-1}*, i.e., whether a customer opened the prior e-mail she received. The variable is 1 if the customer opened the prior e-mail, and 0 otherwise.

⁷ We do model the joint decision $P(B, A)$, where $A = \{\text{open, do not open}\}$ and $B = \{\text{unsubscribe, no action, click-through, purchase}\}$, but we have broken it into a series of conditionals; i.e., $P(A, B) = P(B|A) \times P(A)$.

- *FREQUENCY*, i.e., the number of e-mails a customer received from this firm in the last 30 days.

- *FREQUENCY_SQR*, i.e., square of *FREQUENCY*.

We further specify each individual-specific parameter θ_n to be drawn from a continuous normal distribution (Rossi et al. 1996). Such a specification provides a flexible random component specification to incorporate both observable and unobservable consumer-specific heterogeneity. We specify the multivariate regression as

$$\theta_n = \Psi \cdot Z_n + \mu_n \quad \text{and} \quad \mu_n \sim iid. N(0, V_\theta), \quad (3)$$

where the individual-specific coefficients θ_n are regressed on the observable consumer characteristics Z_n (including an intercept), μ_n is the unobservable component of the consumer heterogeneity, and V_θ is the variance-covariance matrix whose diagonal elements represent the fraction of the variance of θ that is unexplained by observable consumer characteristics. Recall that the parameter θ_n captures the impact of X on the probability of a customer opening an e-mail. Equation (3) highlights that observed and unobserved consumer heterogeneity affects these parameters. We use prior purchase behavior as an observed consumer characteristic as defined by the variable *PRIOR_PURCHASE*.

- *PRIOR_PURCHASE*, i.e., whether the customer made any purchase in the calendar year prior to the commencement of our study. The variable is 1 if the customer made a purchase, and 0 otherwise. We use the *PRIOR_PURCHASE* variable as a measure of familiarity with the firm because prior literature (Söderlund 2002) has used prior purchase as a common indicator of familiarity.

For inference, we use a hierarchical Bayesian model. For a discussion on the priors and conditional posteriors of this model, see the Technical Appendix A1.⁸ The advantages of using Bayesian analysis are, first, that some customers have sparse observations. For example, about 1/4 of customers in this study received only one e-mail; therefore, we have only one data point about them. Classical techniques rely on the asymptotic properties of large samples, and our data often involve small samples at the individual level. Second, because we estimate a large number of parameters (19,661 customers and 5 parameters for each customer), the partial pooling of data in the Bayesian method offers more information to estimate individual-specific parameters than would independent customer-level models. Additionally, the Bayesian model properly accounts for heterogeneity; it avoids biased variances and therefore avoids inferences about population parameters.

⁸ An electronic companion to this paper is available as part of the online version at <http://dx.doi.org/10.1287/isre.1110.0384>.

5.2. Stage 2

In stage 2, conditional on opening the e-mail, a consumer takes any one of the following actions: (1) unsubscribes from the mailing list by clicking on the unsubscribe link at the bottom of the e-mail, (2) opens the e-mail but takes no action, (3) clicks the link in the e-mail to visit the firm's website (but does not make a purchase), and (4) purchases the product featured in the offer. Because a clear ordinal ranking exists among these responses, we use an ordered probit model (Greene 2003). In stage 2, both product-based personalization and personalized greetings are likely to influence consumer choice. Therefore, we use two independent variables to represent personalization:

- *PRODUCT*, i.e., whether the e-mail features a consumer's product of interest. A positive coefficient for this variable would provide support to H1.

- *NAME*, i.e., whether the e-mail contains a personalized greeting. A positive (negative) coefficient for this variable would support (oppose) the use of names in e-mail advertisements.

In an ordered probit model, the latent utility function is specified as

$$U_{km} = \beta_m \cdot W_{km} + \xi_{km}. \quad (4)$$

Further, $m = 1, \dots, M$ is the number of consumers who opened at least one e-mail; $k = 1, 2, \dots, K_m$ is the number of e-mails that consumer m opens; β_m are the individual-level parameters to be estimated; and W are the independent variables. The consumer utility in stage 2 is conditional on opening the e-mail in stage 1.⁹ In an ordered probit model, consumer m chooses h if $\alpha_{m,h-1} < U_{km} < \alpha_{m,h}$, where α s are the cut-off points estimated along with β s. Because ξ is distributed normal, the probability of observing choice h is given as $P_h(U_{km}, \beta_m \cdot W_{km}) = \Phi(\alpha_{hm} - \beta_m \cdot W_{km}) - \Phi(\alpha_{(h-1)m} - \beta_m \cdot W_{km})$.

Because consumers in our sample have four choices, we use $\alpha_{0m} = -\infty$ and $\alpha_{4m} = \infty$ restriction for identification (i.e., the lower and upper bounds on the cutoff points) as is common in the literature. The parameters to be estimated include the coefficients β_m and the three cutoff points of the ordered probit model. We define $\delta_m = [\alpha_{hm}, \beta_m]$ as the set of individual-level parameters to be estimated, where $h = 1, 2, 3$. Note that we estimate the vector of parameters δ_m for each consumer separately. Here, we jointly estimate the cutoff points and coefficients using a form of the Gibbs sampler known as the "collapsed Gibbs" sampler (Liu 1994). Cowles (1996)

⁹ The U_{km} term should actually be written as $U_{km|o}$ to indicate that the utility is conditional on a customer opening the e-mail in stage 1. For the sake of notational convenience, we drop the $|o$ subscript.

shows that such a model is more efficient than a pure Gibbs sampler-based method for estimating an ordered probit model (as specified in Albert and Chib 1993). The independent variables of interest W at this stage include the two personalization variables *PRODUCT* and *NAME* as well as control variables such as

- *RESPONSE* _{$t-1$} , i.e., whether the customer made a click through or purchase on the prior e-mail she opened.

- *PRICE*, i.e., whether the e-mail mentioned the price offer.

- *GIFT*, i.e., whether the advertisement offered a free gift, price off, or cash back with the purchase.

- *COMPARISON*, i.e., whether the advertisement contained a comparison of competitors' prices.

Similar to the hierarchical Bayesian model in §5.1, we specify each individual-specific parameter δ_m to be drawn from a continuous normal distribution:

$$\delta_m = \Delta \cdot Z_m + v_m \quad \text{and} \quad v_m \sim iid. N(0, V_\delta), \quad (5)$$

where Z_m are the observable customer-specific variables including an intercept term. As in stage 1, we use the *PRIOR_PURCHASE* variable as an observed consumer characteristic. This model specification allows us to test for the moderating effect of familiarity on personalized greetings. A positive coefficient (Δ) for the *PRIOR_PURCHASE* variable in Equation (5) when β_{NAME} is the dependent variable suggests that consumers who made prior purchases with a firm are more likely than customers who did not make prior purchases with that firm to respond positively to personalized greetings. Therefore, a positive coefficient for this variable would provide support to H2.

6. Results

The results of an MCMC output are in the form of draws from a distribution for each parameter instead of point estimates. For example, for each individual-level parameter θ_n and β_m in Equations (2) and (4), respectively, the MCMC output gives us the distribution of each parameter for every customer.

6.1. Stage 1 Estimates

The sample size in stage 1 is 364,646 e-mails. We run the MCMC simulation for 9,000 draws and discard the first 6,000 as burn-in. Further, we use a thinning parameter of 3 (that is, out of the remaining 3,000 draws, we retain every third draw for our posterior distribution).¹⁰ The mean of the rejection rate for the Metropolis-Hastings (M-H) algorithm is 0.8, which is

¹⁰ The only purpose for thinning the data is to reduce storage space and the computational burden of analyzing the stored draws.

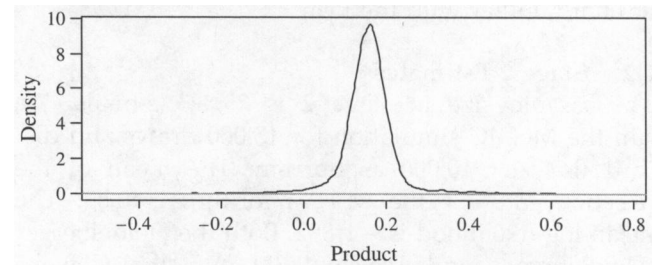
within the desired rejection rate of 0.6–0.9. The mean log-likelihood is $-38,625$. We conducted two tests to check for the convergence of our MCMC output—the Heidelberg (Heidelberger and Welch 1983) and the Geweke (1992) diagnostic tests. Both tests indicate adequate convergence. We calculate the pseudo $R^2 = (L_0 - L_M)/L_0$ measure defined by McFadden (1974). L_M is the log-likelihood of our model, and L_0 is the log-likelihood of a constant-only model (no independent variables). For the stage 1 model, we find that $L_M = -38,625$, $L_0 = -42,666$, and pseudo $R^2 = 9.5\%$. Table 1 summarizes the posterior distribution of the individual-specific means of the parameters (θ_n) in Equation (2).

For all tables, the significant coefficients are denoted in boldface type. The significance of the coefficients can be interpreted as follows: $^{+ (*)}$ denotes that more than 90% of the values in the posterior distribution are positive (negative). This is equivalent to a one-tailed significance test at $p < 0.1$ level in classical regression. Similarly, $^{++}$, $^{+++}$ (or ** , ***) are equivalent to significance at the $p < 0.05$ and $p < 0.01$ levels, respectively.

Each value in column 2 of Table 1 is calculated by averaging the mean values of the parameter estimates for each consumer. We continue with this notation for the rest of the paper. Table 1 can be interpreted as follows: The value 0.16 for $\theta_{PRODUCT}$ suggests that the mean of the posterior distribution of individual-specific coefficients for the product-based personalization variable is 0.16. Column 3 of Table 1 (% MCMC draws greater than 0) suggests that 99% of customers have a positive mean for $\theta_{PRODUCT}$ (which is also evident from Figure 4). This finding suggests that product-based personalization positively impacts the probability that an e-mail will be opened by more than 99% of customers. This finding provides support for H1.

The coefficient $\theta_{OPEN1-1}$ is, on average, positive for more than 99% of consumers, which implies that consumers who opened a previous e-mail are more likely to open the current e-mail. Additionally, the coefficient $\theta_{FREQUENCY}$ is positive and significant, whereas the coefficient $\theta_{FREQUENCY_SQ}$ is negative and significant. This observation implies that, although sending more e-mails to customers does improve response rates, increasing the frequency beyond a certain

Figure 4 Distribution of Individual-Specific Means for $\theta_{PRODUCT}$



threshold may actually result in a decline in response rates. This conclusion confirms findings in prior literature that the response to increased frequency of advertisements follows an inverted U-shaped curve (Berlyne 1970). This holds true for e-mail advertising as well.

Next, we report the results from estimating the parameters from Equation (3). In Equation (3), we let the vector θ (estimated in Table 2) itself be regressed on observed individual characteristics, namely, whether a consumer had made a prior purchase with the firm. Table 2 presents the means of the posterior distribution of the hierarchical regression coefficients ψ .

The coefficients in Table 2 can be interpreted as follows: Each row in Table 2 estimates the impact of the observable consumer variable Z_n on the model parameters θ . For example, the coefficient of *PRIOR_PURCHASE* for $\theta_{PRODUCT}$ is 0.006. This implies that the estimated coefficient Δ in the equation $\theta_{PRODUCT} = \text{Intercept} + \Delta \times \text{PRIOR_PURCHASE}$ is 0.006. This coefficient is not significant, however, which suggests that there is no significant difference in response to product personalization between customers who made any prior purchases and those who did not. In other words, product-based personalization increases the likelihood that an e-mail will be opened to the same degree for both familiar and nonfamiliar customers. As discussed earlier, familiarity is not likely to affect the efficacy of *implicit* product personalization. Similarly, the coefficient of *PRIOR_PURCHASE* for $\theta_{FREQUENCY}$ is 0.12 (see Table 2). Thus, the coefficient Δ in the equation $\theta_{FREQUENCY} = \text{Intercept} + \Delta \times \text{PRIOR_PURCHASE}$ is 0.12, which implies that customers familiar with the firm are more likely to respond positively to an

Table 1 Posterior Distribution of θ_n

θ_n	Mean	% MCMC draws greater than 0
Constant	-2.88^{***}	0.00
$\theta_{PRODUCT}$	0.16⁺⁺⁺	0.99
$\theta_{OPEN1-1}$	0.85⁺⁺⁺	1.00
$\theta_{FREQUENCY}$	0.20⁺⁺⁺	0.99
$\theta_{FREQUENCY_SQ}$	-0.15^{**}	0.02

Table 2 Posterior Distribution of ψ

θ	Intercept	<i>PRIOR_PURCHASE</i>
Constant	-2.91^{***}	0.27⁺⁺⁺
$\theta_{PRODUCT}$	0.16⁺⁺⁺	0.006
$\theta_{OPEN1-1}$	0.87⁺⁺⁺	-0.17^{***}
$\theta_{FREQUENCY}$	0.19⁺⁺⁺	0.12⁺⁺⁺
$\theta_{FREQUENCY_SQ}$	-0.15^{***}	0.02⁺⁺⁺

increased frequency of e-mails than customers who are not familiar with the firm.

6.2. Stage 2 Estimates

The sample size in stage 2 is 23,323 e-mails. We run the MCMC simulation for 15,000 draws and discard the first 10,000 as burn-in. The mean of the rejection rate for the M-H algorithm is 0.83. The mean log-likelihood is $-4,882$. Both the Heidelberger (Heidelberger and Welch 1983) and the Geweke (1992) diagnostic tests indicate adequate convergence. Again, we first present the estimates from the ordered probit model. The parameter estimates in stage 2 follow from Equations (4) and (5) and can be interpreted in a manner similar to the interpretation of the stage 1 results in Tables 1 and 2. We also calculate the pseudo R^2 as defined in §6.1. For the stage 2 regression, $L_M = -4,882$, $L_0 = -5,739$, and pseudo $R^2 = 15\%$. Table 3 summarizes the posterior distribution of the individual-specific means of the model parameters (β_m) in Equation (4) and the individual-specific cutoff points (α_{1m} , α_{2m} , α_{3m}) of the ordered probit model.

The key variables of interest for our hypotheses testing are the personalization variables. The coefficients of $\beta_{PRODUCT}$ and β_{NAME} are 0.64 and -0.32 , respectively (see Table 3). We also find that the mean coefficient of $\beta_{PRODUCT}$ is positive for 98% of consumers, and the mean coefficient of β_{NAME} is negative for 95% of consumers. We plot the posterior means of the individual-level coefficients of the personalization variables in Figure 5. Figure 5(a) illustrates the results in Table 3 by showing that almost all customers have a positive posterior mean for $\beta_{PRODUCT}$, which suggests that customers respond positively to e-mails with *implicit* product-based personalization. Thus, H1 is supported in stage 2 as well. The bimodality of this distribution also suggests the presence of consumer segments, which we explore below in §6.3. Figure 5(b) represents the density of the individual-specific posterior means of β_{NAME} . Our MCMC output shows that 95% of customers have a negative mean value for β_{NAME} . This finding suggests

that consumers on average respond negatively to personalized greetings. The coefficients of *COMPARISON* and *GIFT* are negative and significant, which suggests that consumers are less likely to make a click-through or purchase if the e-mail contains information about price or free gifts. Although a detailed analysis of e-mail characteristics is beyond the scope of this paper, a likely reason for the negative coefficients of *COMPARISON* and *GIFT* could be that providing more details in an e-mail may reduce the consumer's need to visit the firm's website by making a click-through.

The negative response to name provides evidence of consumers' discomfiture with seeing their names being used in e-mail advertisements. This may come as a surprise to many in the academic and practitioner communities, where it has been mainly assumed that personalization (especially personalized greetings) leads to positive responses from consumers. One may wonder why consumers would respond negatively to the use of name in e-mail advertisements when they themselves registered with the firm and provided their names and e-mail addresses. In fact, sending personalized greetings is considered the first step of e-mail personalization.¹¹ There are many possible explanations for consumers' negative responses to firms' use of information (in this case, name). First, given the rampant use of e-mails for phishing and identity theft, consumers may become suspicious when they see their names explicitly used to hawk products. Consumers may not expect a firm to use information they provided in the past for some other purpose (to make a purchase or create an account) for possible monetary gains as in e-mail advertisements. This fits with the *unauthorized secondary use* principle of information sharing between consumers and firms; see Malhotra et al. (2004). As we explain in §3, if people do not recall exactly what information and authorization they provided to an online vendor, they may commit availability and simulation fallacies and simulate or imagine negative instances. Or, if they lack confidence, they may fall for anchoring and adjustment biases as they play safe and turn away from vendors' personalized offers. Second, although customers routinely provide name and other personal information online, they still value anonymity on the Internet.¹² Using a consumer's name compromises that consumer's anonymity and

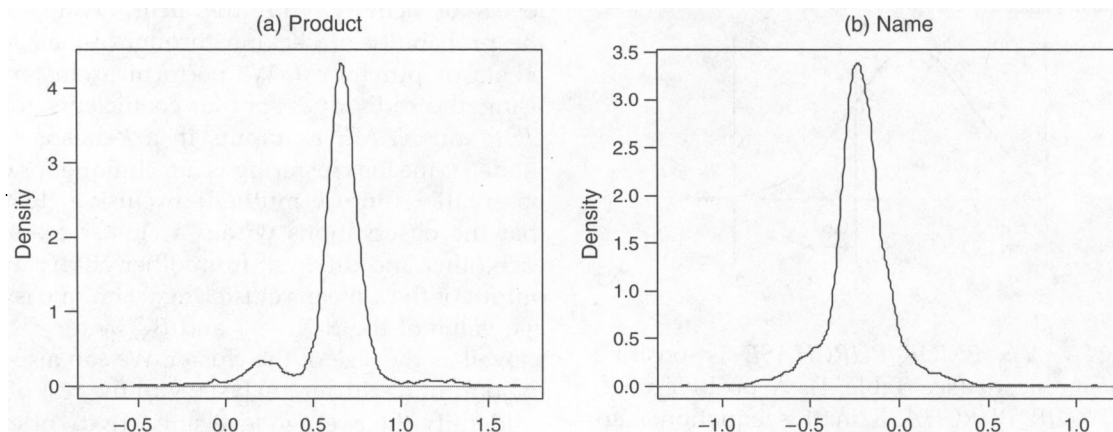
Table 3 Posterior Distribution of δ_m

δ_m	Mean	% draws greater than 0
α_1	-3.97*	0.08
α_2	2.26+++	1.00
α_3	4.43+++	0.99
β_{PRICE}	-0.02	0.29
$\beta_{COMPARISON}$	-0.68**	0.02
β_{GIFT}	-0.29*	0.06
$\beta_{RESPONSE1-1}$	0.19	0.81
$\beta_{PRODUCT}$	0.64++	0.98
β_{NAME}	-0.32**	0.05

¹¹ EmailMarketingPro. Targeted email marketing: Segmentation and personalization. Accessed November 12, 2010, <http://emailmarketingpro.org/targeted-email-marketing-segmentation-and-personalization>.

¹² GVU Center. GVU's 10th WWW user survey. Accessed March 7, 2007, http://www-static.cc.gatech.edu/gvu/user_surveys/survey-1998-10.

Figure 5 Individual-Specific Means on the Personalization Coefficients



may raise a privacy flag and thus create a negative reaction. According to prior research, anonymity is a mechanism to protect privacy.¹³ Third, seeing their names in e-mail advertisements may raise customers' expectations, and they may expect more personalization from firms. This argument follows from the concept of fair exchange (Smith et al. 1996). A customer's response to personalized e-mails could be similar to the following reaction of an eBay customer: "I *already know* that eBay knows my account ID and e-mail address, and I *don't care*. The fact that they can pull this information from a database and slap it into a bulk e-mail doesn't impress me in the slightest: the content of the e-mail that they're sending to me is still totally generic, reflecting nothing about my interests or history with eBay" (see McNamara 2005).

One can argue that negative response to personalized greetings is not a result of consumers' discomfiture with seeing their names but rather because of nuances in data collection or other explanations. We conduct several robustness checks to disprove several alternate explanations. For example, one can argue that a firm uses personalized greetings only for less popular products. We performed a subsample analysis by grouping products into high sales and low sales based on median sales. We analyzed the response rates and personalized greetings for each subsample and found no evidence of such a claim.¹⁴ Another argument about negative response to personalized greetings is that customers who never respond might get a higher fraction of e-mails with personalized

greetings. We did not find evidence of this in our data but we controlled for it by estimating an individual-level coefficients model where the intercepts will capture any unobserved individual-level differences such as whether a customer responds more or less often. Moreover, our extensive discussions with the firm ruled out the possibility of these biases in the data. We also ran our model after controlling for different product categories and confirmed that our results hold. Finally, our empirical model, which controls for observable and unobservable consumer-specific heterogeneity, rules out the possibility that systematic differences between customers drive our results.

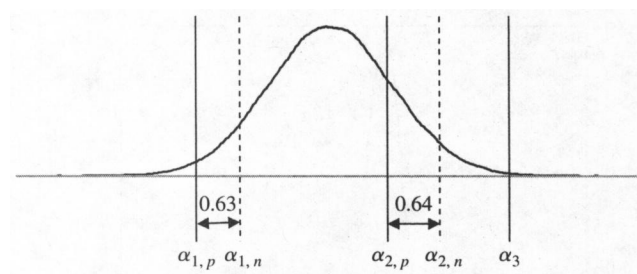
To test H2, we analyze how the negative response for NAME differs for familiar customers (those who made prior purchases with the firm) versus unfamiliar customers (those who did not). To do this, we examine the results from our estimation of Equation (5), which outlines the relationship between estimated parameters and observable consumer characteristics. Table 4 provides the means of the posterior distribution of Δ of the multivariate regression specified in Equation (5).

From Table 4, notice that the coefficients of the *PRIOR_PURCHASE* variable are negative and significant for α_1 and α_2 . Thus, customers who made any prior purchases are less likely to unsubscribe and more likely to click through. This is also highlighted in Figure 6, where the coefficients $\alpha_{i,n}$ (dashed vertical line) represent the cutoff points for customers that did not make prior purchases, and the coefficients $\alpha_{i,p}$ (solid vertical line) represent the cutoff points for customers who made prior purchases. From Figure 6, we see that the latter category of customers have a higher probability of clicking through and a lower probability of unsubscribing. This also supports prior literature that familiarity is associated with positive outcomes for a firm (Bhattacharjee 2002).

An interesting result is that the coefficient Δ of *PRIOR_PURCHASE* on β_{NAME} in the equation β_{NAME}

¹³ The *Stanford Encyclopedia of Philosophy* summarizes the relation between privacy and anonymity as follows: "Privacy can be gained in three independent but interrelated ways: through secrecy, when no one has information about one, through anonymity, when no one pays attention to one, and through solitude, when no one has physical access to one" (DeCew 2008).

¹⁴ We thank the anonymous reviewers for suggesting these alternative explanations.

Figure 6 Impact of *PRIOR_PURCHASE* Variable on the Ordered Probit Cutoff Points

= Intercept + $\Delta \times \text{PRIOR_PURCHASE}$ is positive (0.27) and significant (see Table 4). A positive coefficient of *PRIOR_PURCHASE* in this equation suggests that the value of β_{NAME} is less negative for customers with *PRIOR_PURCHASE* = 1 than for customers with *PRIOR_PURCHASE* = 0. Mathematically, $\beta_{\text{NAME}} = -0.36$ for customers who did not make prior purchases and $-0.09(-0.36 + 0.27 \times 1)$ for customers who made prior purchase(s). This finding implies that customers who made prior purchases also respond less negatively to personalized greetings in e-mail advertisements than customers who did not make purchases. Because past purchase behavior is a commonly used indicator of familiarity (Söderlund 2002), this finding also supports H2 in that familiarity with a firm indeed leads to less negative reaction to personalized greetings in e-mail advertisements.

Surprisingly, the coefficient of *PRIOR_PURCHASE* on β_{PRODUCT} is negative (see Table 4), which suggests that customers who did not purchase previously respond more positively to product personalization. A likely reason is the nature of products considered in this study, such as gas and electricity, where consumers usually sign long-term contracts. Also, it can imply that nonfamiliar customers are more likely to click only on e-mails that feature products of interest to them, whereas familiar customers are more open to clicking on e-mails that feature other products as well.

6.3. Consumer Segments

Next, we characterize the heterogeneity in consumer response to product-based personalization and per-

sonalized greetings in terms of customers' current levels of activity with the firm (as measured by the probability of clicking through e-mail advertisements or purchases). We perform a cluster analysis using the individual-specific coefficients for *PRODUCT* and *NAME* as inputs in a *k*-means clustering model; *k*-means clustering is a technique for grouping observations into *k* mutually exclusive clusters such that the observations within a cluster are similar to each other and different from other observations. The output of the *k*-means clustering technique is the average value of the β_{PRODUCT} and β_{NAME} for each cluster as well as the size of the cluster. We can also combine the output of cluster analysis with the rest of the data to identify the average levels of activity of each cluster in terms of click-through and purchase probability. The results are shown in Table 5.¹⁵

We can interpret Table 5 as follows. Customers in segment A have an average coefficient of 0.66 for the product personalization variable (*PRODUCT*) and -0.6 for the personalized greeting variable (*NAME*). Both coefficients are significant at the 0.01 level. Further, 71.5% of our sample of customers belong to segment A. Customers in segment A have a 3.8% probability of clicking through and a 0.3% probability of making purchases.

Table 5 suggests some interesting results about how the firm can predict customers' responses to personalized greetings based on their current levels of activity. Customers in segment A respond more negatively to personalized greetings than do customers in the other two segments (the response is also significant at the 0.01 level). They also have the lowest click-through and purchase probability of all the other segments. On the other hand, the response of customers in segments B and C to personalized greetings is not significantly different from zero. This observation suggests that consumers in these segments do not respond negatively to personalized greetings. We also find that customers in segments B and C also have higher click-through and purchase probabilities than customers in segment A. Overall, our results suggest that customers who have high levels of current activity at the website in terms of higher click through on e-mail advertisements and purchase rates also have less negative responses to personalized greetings than customers with low levels of current activity.

Table 4 Posterior Distribution of Δ

δ	Intercept	<i>PRIOR_PURCHASE</i>
α_1	-3.89***	-0.63***
α_2	2.34***	-0.64***
α_3	4.41***	0.25
β_{PRICE}	-0.01	-0.09
$\beta_{\text{COMPARISON}}$	-0.70***	0.10
β_{GIFT}	-0.27***	-0.11 ⁺
$\beta_{\text{RESPONSE1-1}}$	0.22***	-0.27***
β_{PRODUCT}	0.70***	-0.37***
β_{NAME}	-0.36***	0.27***

7. Conclusions

This research is among the first studies to empirically analyze how consumers respond to firms' use

¹⁵ In *k*-means clustering, we have to specify the number of segments ex ante. In our analysis, we start with two segments and iteratively proceed until further analysis reveals very small segment sizes (with membership of less than 5%).

Table 5 Segment Parameters

SEGMENT	β_{PRODUCT}	β_{NAME}	SIZE (%)	Click-through	Purchase
A	0.66 ⁺⁺⁺	−0.6 ^{***}	71.5	3.8	0.3
B	0.72 ⁺⁺⁺	−0.008	15.9	12.0	1.0
C	0.24	−0.31	12.6	6.7	3.7

of different types of information in e-mails. We consider firms' use of two types of information: customer product preferences and customer names. The firm in our data used customer product preference information to offer product recommendations in e-mails without explicitly informing the consumers about the use of their information. We measure consumer response to e-mail advertisements in two stages: (1) whether the customer opens an e-mail, and (2) conditional on opening the e-mail, whether the customer unsubscribes, takes no further action, clicks through, or makes a purchase. For each stage, we estimate a random utility model using a hierarchical Bayesian framework to understand how individual consumers react to firms' use of information and how observable consumer characteristics such as familiarity with the firm moderate this response.

Our analysis suggests several interesting results regarding consumer response to firms' use of information. One, when firms use product-based personalization (where the use of information is not explicitly mentioned), consumers respond positively. On the other hand, consumers respond negatively when firms are explicit in their use of personally identifiable information (i.e., personalized greetings in e-mail advertisements). Two, the negative response to personalized greetings is moderated by consumer familiarity with the firm. The negative response to personalized greetings is less for consumers who made any prior purchases with the firm. We also show that the negative response to personalized greetings is not significant for consumers whose current level of activity is high (in terms of higher click through to e-mail advertisements and purchases).

We advance arguments to suggest that a likely reason for the negative response to personalized greetings is consumers' discomfiture with seeing their information used for e-mail advertising. This discomfiture could be related to privacy concerns such as violation of anonymity, absence of fair exchange, or perception of unauthorized secondary use of data. However, our secondary data do not enable us to measure privacy concerns directly, and thus we cannot claim that privacy is the sole reason for the negative response to personalized greetings. We therefore use theory and surrounding evidence in support of our arguments. We also perform robustness checks to rule out the possibility that the negative response

to personalized greetings is a result of nuances in the data.

Prior literature establishes that the use of consumer information can sometimes lead to privacy concerns (e.g., when the firm uses information collected from a consumer for a totally different purpose than it was collected for, or if the firm uses a customer's personal identification information). "Name" is an obvious and overt use of personal identification information. Because the e-mail solicits a consumer to spend money to buy the advertised product, it may also evoke security concerns. Personalized greetings are unlike product-based personalization where a firm uses consumers' past preference data to offer targeted products; in our data, product-based personalization is done tacitly without a consumer's awareness of the e-mail's personalization. We also argue that familiarity should reduce consumer concerns about information use and hence should improve reaction to name-based personalization. Our data provide support for this contention. We find that customers who made any prior purchases with a firm or have higher levels of activity currently (in terms of higher click-through to e-mail advertising or for purchases) also respond less negatively to personalized greetings.

Our result that the use of name leads to a negative response is interesting because it shows the need for more research on personalization and privacy using secondary data. We contrast our results with Tam and Ho (2006), who found that personalized greetings in banner advertisements lead to a positive response from consumers. This contrast raises an interesting question: Why is our result different? Tam and Ho (2006) tested their hypothesis regarding response to name in a laboratory setting where undergraduate students reacted positively if their names appeared in banner ads. We can fairly assume that subjects had a high level of trust with the experimenters, given the academic setting. Similarly, in the field study of Tam and Ho (2006), subjects were not e-mailed with product offers; they volunteered to participate in the study and had to visit a specific website and log in if they wished to download free music. In contrast, our paper examines the effects of name-based personalization when advertisers "push" information to users. Subjects in Tam and Ho (2006) had to bear little or no financial risks participating in an academic study,

but e-mail recipients in our study had to deal with the real or perceived risks of responding to online vendors who sought our participants' business using personally identifiable information. One could argue that their subjects were endowed with a high level of familiarity. Hence, the findings of Tam and Ho (2006) are not contradictory to our results. Rather, our work extends their study by showing that consumers' responses to personalized greetings are not always positive in the online context and depend, among other things, on the level of familiarity with a firm.¹⁶ The fact that the difference in problem settings leads to a different result is interesting and suggests the need for more studies on the topic of consumers' responses to information use with secondary data.

Our paper has several interesting implications for managers and policy makers:

1. The negative response to personalized greetings is interesting to policy makers because one assumes that, after consumers provide their information and "accept" the terms and conditions of a privacy policy, a firm owns the information and can use it in any way it sees fit. Our study does confirm, though, that consumers are still concerned about the use of their information. A likely reason is that many consumers do not read privacy policies and click "accept" to prevent exclusion from the features of a website (most online firms do not allow registration or transactions unless customers "accept" their terms). This suggests that policy makers should be cognizant that consumers care about privacy even after they provide information "voluntarily" to firms. Firms should design safeguards to prevent the use of information in ways not intended by consumers.

2. We also recommend that managers take the following strategy for personalized greetings: For customers who make few or no purchases and have a lower probability of click-through (segment A), the firm should send e-mail advertisements that feature only their product(s) of interest. These customers should not receive personalized greetings. For customers who have higher click-through and purchase rates (segments B and C), the firm should use product-based personalization. The use of personalized greetings has no significant impact on consumer response, so the firm can use its discretion to judge whether or when to use personalized greetings for these customers. We characterize these segments in terms of their current levels of activity (as measured by click-through rates and purchases). However, managers can use additional information about consumers to characterize these segments in terms of demographic variables such as age, income, gender, and occupation.

3. Another implication from our results is that managers should send a higher frequency of e-mails to familiar customers than to nonfamiliar customers.

4. Finally, we show that implicit product-based personalization yields higher response rates from customers and that the increase in response rates is higher for customers who are not familiar with the firm. This finding implies that customers who are familiar with a firm are more open (i.e., respond positively) to receiving e-mails that do not feature their products of interest. Therefore, managers should send e-mails with no product-based personalization mainly to customers who are familiar with the firm.

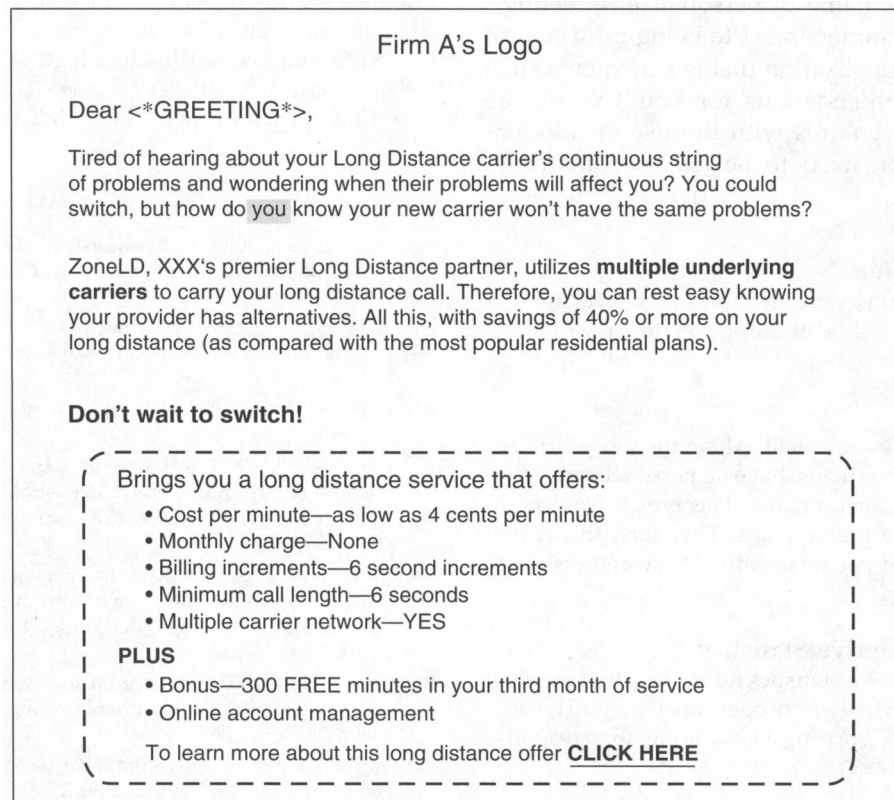
Increasingly, firms are getting access to more and more sensitive personal information such as location (Foursquare.com), networks of friends, and daily activities (Facebook.com). The recent fiasco over Facebook's Beacon system highlights the thin line that companies must walk while using customers' information and shows that there is a difference between *access* to information and *license* to use that information.¹⁷ Just because consumers post their activities and friends' lists on Facebook.com or their current location on Foursquare.com does not mean that firms have a license to use that information in advertising. Our study shows that firms should not use this information blindly to offer personalized services. However, firms should seek to categorize consumers into segments and use different types of information to offer personalized products and services to each segment.

Although our study takes an interesting first step in this area, the results need to be carefully interpreted in light of some limitations. One might argue that some customers never click the unsubscribe link even if they receive little utility (or disutility) from the e-mail advertisement. They may prefer to do nothing. Therefore, "unsubscribe" does not fit the ordinal classification used by our analysis. To check the robustness of our results, we combine unsubscribe and no action as a single category and run the stage 2 model with three values of the dependent variable: 0 for unsubscribe/no action, 1 for click-through (not a purchase), and 2 for purchase. We find that our results are qualitatively unchanged in the revised model. One can also argue that the negative response to name can be a result of nuances in the data collection. The data collection was pseudorandom, so standard concerns such as selection bias (basing a decision to send e-mails on potential outcomes) appear here.

¹⁶ The authors thank an anonymous reviewer for pointing this out.

¹⁷ Beacon was a system designed to report back to Facebook on members' activities on third-party sites that participate in Beacon, even if the users are logged off from Facebook and have declined having their activities broadcast to their Facebook friends. See Perez (2007).

Figure A.1 Sample E-mail Advertisement



For example, if the firm sent personalized greetings for e-mail advertisements for unpopular products or to customers with low rates of response, the results could be biased.¹⁸ We adjust for these biases by quantitative analysis and by gathering qualitative information from the firm about the practices it followed for e-mail advertising. In our discussion with the firm, the management explained that the firm did not knowingly follow any specific policies regarding the types of products and which consumers received personalized greetings. The firm also admitted to a lack of resources and skills to analyze past responses to personalization in planning any e-mail strategy or in targeting specific groups of customers. So, although our firm did not randomize (so the proportion of e-mails sent across segments may not be equally distributed), it did not explicitly select specific segments, either. Further, we did not find the use of personalized greetings only for unpopular products in our data. We also ran our model after controlling for different product categories and confirmed that our results hold. Finally, our empirical model estimates individual-level coefficients and controls for potential biases arising out of observable or unobservable consumer characteristics.

¹⁸ We thank the associate editor and anonymous reviewers for pointing this out.

Although our study suggests the possibility that consumers' discomfiture with personalized greetings is likely a result of privacy and security concerns, future field studies would benefit by measuring these constructs directly. Also, we measure familiarity by whether a customer made a purchase in the past. However, past experiences can be positive as well as negative. A negative past experience may not moderate the relation between personalized greetings and consumer response. Future studies can study familiarity and the impact of positive and negative prior experience on response to personalized greetings. More research is also needed to understand how the negative response to personalized greetings dynamically evolves over time. Another limitation of this research is that our domain is limited to e-mail advertisement, and no studies have tested how these findings would hold in other contexts. The nature of the firm (small and relatively less known) may also influence the results. Studying our model for well-known firms and brands such as Amazon.com and Netflix.com would be interesting. We might expect to see a less negative response to personalized greetings sent by well-known firms versus little-known websites. Of course, customers may still be wary of e-mail advertisements even from well-known firms. Also, we look at the use of information in two specific contexts: the use of preference information tacitly (without informing the

consumer that the message is personalized) and the use of a consumer's name in personalized greetings. Observing how consumers react to being informed of product-based personalization (using statements such as "we have recommendations for you") would be interesting. Other scenarios with the use of different types of information need to be explored in future research.

Electronic Companion

An electronic companion to this paper is available as part of the online version at <http://dx.doi.org/10.1287/isre.1110.0384>.

Appendix A

Figure A.1 shows a sample e-mail. Along the top is firm A's corporate logo. Some e-mails have a personalized greeting, which is the customer's name. This e-mail features an offer for long distance phone plans. Therefore, this e-mail also has product-based personalization for customers in the "long distance" pool.

Appendix B. Descriptive Statistics

Table B.1 gives descriptive statistics for stage 1, that is, when a customer decides whether to open an e-mail. The low response rate is not surprising considering that response

Table B.1 Descriptive Statistics in Stage 1

Total number of customers selected to receive e-mail	19,661
Total number of e-mails sent	364,646 ^a
Mean number of e-mails received per consumer	18.54
Total number of personalized e-mails	36,444
Total number of e-mails opened by consumers	23,323
Fraction of nonpersonalized e-mails opened	5.6%
Fraction of personalized e-mails opened	10.8%

^aThe management at the firm informed us that it made an effort to restrict the number of e-mails a customer received in a given month to five or fewer.

Table B.2 Descriptive Statistics in Stage 2

Variable	Fraction (out of 23,323 e-mails opened by customers)
Fraction of opened e-mails resulting in a click-through (not a purchase)	0.05
Fraction of opened e-mails resulting in an unsubscribing	0.04
Fraction of opened e-mails resulting in a purchase	0.01
Fraction of opened e-mails resulting in a click-through (not a purchase)	0.05
Fraction of e-mail advertisements where price was mentioned	0.61
Whether comparison with competitors was mentioned	0.15
Whether discount offers were mentioned	0.56
Whether the e-mail had product-based personalization	0.17
Whether the e-mail had a personalized greeting	0.2

rates in e-commerce are typically in the single digits. For example, a study by Opt-In News in 2001 shows that CTRs (click-through rates) for opt-in e-mail marketing campaigns ranged from a low of 0.9% to a high of 8.5% (Weil 2001).

The descriptive statistics for stage 2 (i.e., based on e-mails that were opened in stage 1) are shown in Table B.2.

References

- Acquisti, A., L. K. John, G. Loewenstein. 2011. The impact of relative standards on the propensity to disclose. *J. Marketing Res.* Forthcoming.
- Albert, J. S., S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88**(422): 669–679.
- Ansari, A., C. Mela. 2003. E-customization. *J. Marketing Res.* **40**(2): 131–145.
- Australian, The. 2008. Cisco says 90 per cent of email is spam. (December 17), <http://www.theaustralian.com.au/australian-it-old/cisco-says-90-per-cent-of-email-is-spam/story-e6frgmx-1111118343852>.
- Awad, N. F., M. Krishnan. 2006. The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quart.* **30**(1): 13–28.
- Ba, S., P. Pavlou. 2002. Evidence on the effect of trust building in electronic markets: Price premium and buyer behavior. *MIS Quart.* **26**(3): 243–268.
- Berlyne, D. E. 1970. Novelty, complexity, and hedonic value. *Attention, Perception Psychophysics* **8**(5): 279–286.
- Bertrand, M., D. Karlin, S. Mullainathan, E. Shafir, J. Zinman. 2005. What's psychology worth? A field experiment in the consumer credit market. Working Paper Series 11892, National Bureau of Economic Research, Cambridge, MA.
- Bhattacharjee, A. 2002. Individual trust in online firms: Scale development and initial trust. *J. Management Inform. Systems* **19**(1): 213–243.
- Bugental, J., S. Zelen. 1950. Investigations into the self-concept: The W-A-Y technique. *J. Personality* **18**(4): 483–498.
- Chellappa, R. K., R. Sin. 2005. Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Inform. Tech. Management* **6**(2–3): 181–202.
- Chen, P.-Y., L. M. Hitt. 2002. Measuring switching costs and the determinants of customer retention in Internet-enabled businesses. *Inform. Systems Res.* **13**(3): 255–274.
- Chen, Y., C. Narasimhan, Z. J. Zhang. 2001. Individual marketing with imperfect targetability. *Marketing Sci.* **20**(1): 23–41.
- Chintagunta, P. K., D. C. Jain, N. J. Vilcassim. 1991. Investigating heterogeneity in brand preferences in logit models for panel data. *J. Marketing Res.* **28**(3): 417–428.
- Cowles, M. K. 1996. Accelerating Monte Carlo Markov chain convergence for cumulative-link generalized linear models. *Statist. Comput.* **6**(2): 101–111.
- Culnan, M. J. 1993. How did they get my name? An exploratory investigation of consumer attitudes toward secondary information use. *MIS Quart.* **17**(3): 341–363.
- Culnan, M. J., B. J. Bies. 2003. Consumer privacy: Balancing economic and justice considerations. *J. Soc. Issues* **59**(2): 323–342.
- DeCew, J. 2008. Privacy. *Stanford Encyclopedia of Philosophy*, Fall 2008 ed. Edited by E. N. Zalta. Accessed November 10, 2009, <http://plato.stanford.edu/archives/fall2008/entries/privacy/>.
- Dewan, R., B. Jing, A. Seidmann. 2003. Product customization and price competition on the Internet. *Management Sci.* **49**(8): 1055–1070.

- Dinev, T., P. Hart. 2006. An extended privacy calculus model for e-commerce transactions. *Inform. Systems Res.* 17(1) 61–80.
- Federal Trade Commission. 2000. Online profiling: A report to Congress. Report, Federal Trade Commission, Washington, DC. <http://www.ftc.gov/os/2000/06/onlineprofiling-reportjune2000.pdf>.
- Gartner. 2007. Gartner says number of identity theft victims has increased more than 50 percent since 2003. Press release (March 6), <http://www.gartner.com/it/page.jsp?id=501912>.
- Gefen, D. 2000. E-commerce: The role of familiarity and trust. *Omega* 28(6) 725–737.
- Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. J. M. Bernardo, J. O. Berger, A. P. Dawid, A. F. M. Smith, eds. *Bayesian Statistics 4*, Oxford University Press, Oxford, UK, 169–193.
- Gonul, F., K. Srinivasan. 1993. Modeling multiple sources of heterogeneity in multinomial logit model: Methodological and managerial issues. *Marketing Sci.* 12(3) 213–229.
- Greene, W. H. 2003. *Econometric Analysis*, 5th ed. Prentice Hall, Upper Saddle River, NJ.
- Heidelberger, P., P. D. Welch. 1983. Simulation run length control in the presence of an initial transient. *Oper. Res.* 31(6) 1109–1144.
- Howard, D. J., C. Gengler, A. Jain. 1995. What's in a name? A complimentary means of persuasion. *J. Consumer Res.* 22(2) 200–211.
- Hsiao, C. 1986. *Analysis of Panel Data*. Cambridge University Press, Cambridge, UK.
- Jensen, C., C. Potts. 2004. Privacy policies as decision-making tools: A usability evaluation of online privacy notices. *Proc. Conf. Human Factors Comput. Systems (CHI '04)*, ACM Press, New York, 471–478.
- Kahneman, D., A. Tversky. 1982. The simulation heuristic. D. Kahneman, P. Slovic, A. Tversky, eds. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge, UK, 201–208.
- Lai, H. J., T. Liang, Y. C. Ku. 2003. Customized Internet news services based on customer profiles. *Proc. 5th Internat. Conf. Electronic Commerce*, ACM Press, New York, 225–229.
- Laudon, K. C., C. G. Traver. 2009. *E-Commerce: Business, Technology, Society*. Prentice Hall, Upper Saddle River, NJ.
- Levy, M., B. A. Weitz. 1992. *Retailing Management*. McGraw-Hill/Chicago.
- Liu, J. S. 1994. The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *J. Amer. Statist. Assoc.* 89(427) 958–966.
- Malhotra, N. K., S. S. Kim, J. Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Inform. Systems Res.* 15(3) 336–355.
- McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. P. Zarembka, ed. *Frontiers in Econometrics*. Academic Press, New York, 105–142.
- McKinsey Quarterly. 2007. How companies are marketing online: A McKinsey global survey. (September), <http://www.sata.ca/file/HowCompaniesAreMarketingOnline.pdf>.
- McNamara, W. 2005. Personalizing, my ass. *Sea Monkey Rodeo* (blog), January 17, <http://seamonkeyrodeo.blogspot.com/2005/01/personalizing-my-ass.html>.
- Milne, G. R., M. E. Gordon. 1993. Direct mail privacy-efficiency trade-offs within an implied social contract framework. *J. Public Policy Marketing* 12(2) 206–215.
- Murthi, B. P. S., S. Sarkar. 2004. Role of management sciences in research. *Management Sci.* 49(10) 1344–1362.
- Nowak, G. J., J. Phelps. 1995. Direct marketing and the use of individual-level consumer information: Determining how and when "privacy" matters. *J. Direct Marketing* 9(3) 46–60.
- Parfeni, L. 2009. Spending on email marketing will increase to \$2 billion by 2014. Softpedia (June 15), <http://news.softpedia.com/news/spending-in-Email-Marketing-Will-Increase-to-2-Billion-by-2014-114217.shtml>.
- Perez, J. C. 2007. Facebook's Beacon more intrusive than previously thought. *PCWorld* (November 30), http://www.pcworld.com/article/140182/facebooks_beacon_more_intrusive_than_previously_thought.html.
- Pitkow, J., H. Schutze, T. Cass, R. Cooley, D. Turnbull, A. Edmonds, E. Adar, T. Breuel. 2002. Personalized search. *Comm. ACM* 45(9) 50–55.
- Posner, R. A. 2008. Privacy, surveillance and law. *Univ. Chicago Law Rev.* 75(1) 245–260.
- Poulton, E. C. 1989. *Bias in Quantifying Judgments*. Lawrence Erlbaum Associates, Hove, UK.
- Rogers, T. B., N. A. Kuiper, W. S. Kirker. 1977. Self-reference and the encoding of personal information. *J. Personality Soc. Psych.* 35(9) 677–688.
- Rossi, P., R. E. McCulloch, G. M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* 15(3) 321–340.
- Shaffer, G., Z. J. Zhang. 1995. Competitive coupon targeting. *Marketing Sci.* 14(4) 395–416.
- Slovic, P., B. Fischhoff, S. Lichtenstein. 1977. Behavioral decision theory. *Annual Rev. Psych.* 28 1–39.
- Smith, H. J., S. J. Milberg, S. J. Burke. 1996. Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quart.* 20(2) 167–196.
- Söderlund, M. 2002. Customer familiarity and its effects on satisfaction and behavioral intentions. *Psych. Marketing* 19(10) 861–879.
- Strandburg, K. J. 2005. Social norms, self control, and privacy in the online world. K. J. Strandburg, D. Raicu, eds. *Privacy and Technologies of Identity: A Cross-Disciplinary Conversation*. Springer Science + Business Media, New York. http://works.bepress.com/katherine_strandburg/8.
- Tam, K. Y., S. Y. Ho. 2003. Web personalization: Is it effective? *IT Professional* 5(5) 53–57.
- Tam, K. Y., S. Y. Ho. 2006. Understanding the impact of Web personalization on user information processing, behavior and judgment. *MIS Quart.* 30(3) 865–890.
- Walters, R. G., H. J. Rinne. 1986. An empirical investigation into the impact of price promotions on retail store performance. *J. Retailing* 62(3) 237–266.
- Wang, P., L. Petrison. 1993. Direct marketing activities and personal privacy. *J. Direct Marketing* 7(1) 7–19.
- Wattal, S., R. Telang, T. Mukhopadhyay. 2009. Information personalization in a two-dimensional product differentiation model. *J. Management Inform. Systems* 26(2) 69–95.
- Weil, D. 2001. Typical response rates for email marketing. *1to1 Media* (November 19), <http://www.1to1media.com/View.aspx?docid=20461>.