V. KUMAR, XI (ALAN) ZHANG, and ANITA LUO*

The rise of new media is helping marketers evolve from digital to interactive marketing, which facilitates a two-way communication between marketers and customers without intruding on their privacy. However, while research has examined the drivers of customers' opt-in and opt-out decisions, it has investigated neither the timing of the two decisions nor the influence of transactional activity on the length of time a customer stays with an e-mail program. In this study, the authors adopt a multivariate copula model using a pair-copula construction method to jointly model opt-in time (from a customer's first purchase to the opt-in decision), opt-out time (from the opt-in decision to the opt-out decision), and average transaction amount. Through such multivariate dependences, this model significantly improves the predictive performance of the opt-out time in comparison with several benchmark models. The study offers several important findings: (1) marketing intensity affects opt-in and opt-out times, (2) customers with certain characteristics are more or less likely to opt in or opt out, and (3) firms can extend customer opt-out time and increase customer spending level by strategically allocating resources.

Keywords: interactive marketing, e-mail marketing, opt-in, opt-out, vine copulas, pair-copula construction

Online Supplement: http://dx.doi.org/10.1509/jmr.13.0169

Modeling Customer Opt-In and Opt-Out in a Permission-Based Marketing Context

Conventional wisdom suggests that customers do not welcome communications from marketers and consider their messages unwanted interruptions that are to be avoided by registering for do-not-mail or do-not-call lists. However, in today's digital age, it is increasingly apparent that customers can also enthusiastically interact with firms by joining their e-mail programs voluntarily, proactively downloading their mobile applications, and following their social media accounts. We therefore argue that customers are not reluctant to receive marketing materials if they are first asked for consent. In 1999, Seth Godin proposed an idea called "permission marketing" and advised marketers to seek customers' permission before sending them promotional messages. Permission marketing creates a channel for two-way interaction and engagement, which is considered crucial for firm value creation. Thus, permission marketing emerges as a solution to the challenge faced by conventional marketing.

Permission marketing typically relies on the use of "new media" channels (e.g., web, e-mail, mobile, social media), which are well suited for interactive marketing (e.g., Winer 2009). Forrester Research (2011) forecasts that marketers in the United States will spend $77 billion on interactive marketing by 2016, the same amount currently spent on television advertising. Among the channels of new media, e-mail and mobile have gained much attention due to their interactive, digital, and cost-effective features (e.g., Shankar and

*V. Kumar is the Regents' Professor, Chang Jiang Scholar—HUST, Richard and Susan Lenny Distinguished Chair & Professor of Marketing, Executive Director of Center for Excellence in Brand & Customer Management, and Director of the PhD Program in Marketing (e-mail: vk@gsu.edu); Xi (Alan) Zhang is a doctoral student in Marketing (e-mail: xzhang31@gsu.edu); and Anita Luo is Assistant Professor in Marketing (e-mail: aluo@gsu.edu), J. Mack Robinson College of Business, Georgia State University. All authors contributed equally to this effort. The authors thank the three anonymous *JMR* reviewers for their valuable guidance in revising this manuscript. They also thank David Schweidel, Peter Fader, and Eric Bradlow for sharing their programming code, which helped the authors to write their own code for model estimation. They thank the retail firm for providing the data for this study and thank Denish Shah, Gayatri Shukla, and Amber McCain for their feedback. Finally, the authors thank Renu for copyediting a previous version of the article. Michael Wedel served as associate editor for this article.

Balasubramanian 2009; Shankar et al. 2010). Forrester Research also forecasts that mobile marketing spending will increase by nearly three times, from $2.8 billion in 2012 to $8.2 billion in 2016. The Direct Marketing Association (2011) forecasts that commercial e-mail will drive up sales by $82.2 billion in 2016.

Previous literature has shown that various factors such as trust and previous experience can affect customers' willingness to accept permission-based marketing (e.g., Jayawardhena et al. 2009; Tezinde, Smith, and Murphy 2002), that trust is an important determinant of online and offline buyer–seller relationships (e.g., Bart et al. 2005; Ganesan 1994), that online habits and sociodemographics affect customers' interest in permission-based web or mobile marketing programs (e.g., Barnes and Scornavacca 2008; Brey et al. 2007), and that an improperly designed message can decrease the response rate and increase the unsubscribe rate (e.g., Marinova, Murphy, and Massey 2002). However, these studies were typically conducted in experimental settings and examined the opt-in and opt-out processes separately, neglecting the possibility that the same customers' opt-in and opt-out behavior could be interdependent. In addition, although some prior studies have discovered that permission marketing can increase customers' brand loyalty and purchase intentions (e.g., DuFrene et al. 2005; Jolley et al. 2013), they have not investigated the possibility that changes in customer loyalty could adversely affect the length of time a customer is willing to stay in a permission-based marketing program. Thus, it is imperative to ask whether customers' opt-in and opt-out behavior can be modeled jointly; how to incorporate the influence of transactional behavior into the modeling of opt-in and opt-out decisions; and how to quantify the influence of a firm's marketing activities on customers' opt-in time, opt-out time, and purchase decisions.

We attempt to bridge the gap in the permission marketing literature by addressing five research questions: (1) What types of customers are more likely to opt in to a permission-based marketing program? (2) How do firms' marketing activities influence the timing of customers' opt-in and opt-out decisions? (3) Is there a dependence between the opt-in and opt-out times? (4) How do transactional behavior and customers' willingness to stay in the marketing program influence each other? and (5) How can firms optimize their marketing contact strategy to both extend the length of time customers stay in the marketing program and increase customers' spending level?

To answer these research questions, we analyze a unique data set from a U.S. retailer spanning 47 months. This database records the time when a customer opts in and out of the firm's e-mail program, the transactions the customer makes, the e-mail open and click-through histories, and the retailer's marketing activities. To obtain each individual customer's online habits and sociodemographic information, we merge the sampled data from the retailer's database using key identifier information with an external database provided by a marketing research firm, Acxiom. The methodological challenge of the research is to jointly model three variables: opt-in timing, opt-out timing, and purchase behavior. We use a multivariate copula model, called "vine copulas" (e.g., Aas et al. 2009; Smith et al. 2010), to capture the dependence structure of the three variables. For the marginal distributions, we model the opt-in and opt-out times using Weibull hazard models and account for unobserved heterogeneity by incorporating a gamma random effect term. We model the average transaction amount using a random-effect log-normal model.

To the best of our knowledge, this is the first empirical study to examine the timing of customers' opt-in and opt-out decisions while accounting for their purchase behavior. In addition, we extend the bivariate copula model into a multivariate copula model by introducing the vine copula to the marketing literature for the first time. Therefore, our study contributes to the existing literature substantively and methodologically.

In the following sections, we first review the literature on (1) permission-based marketing, (2) the linkage between the opt-in and opt-out decisions, (3) capturing the dependence between the durations, and (4) incorporating purchase behavior. Second, we describe our data and present descriptive statistics. Third, we discuss the proposed modeling framework. Fourth, we present the model results and model validation. Finally, we discuss the managerial implications and offer some conclusions.

LITERATURE REVIEW

Permission-Based Marketing

Permission marketing (Godin 1999) proposes that marketers should seek their customers' permission to send them marketing messages. There are two types of permission marketing, namely, opt-in and opt-out marketing. Opt-in marketing refers to firms explicitly asking customers for permission, usually when an online account is created. Customers can opt out any time after they opt in. Opt-out marketing refers to firms sending promotional messages to customers without seeking their permission, including for the first message, but providing customers an option to opt out on each occasion. Because most marketers adopt the former approach, we focus on the opt-in type of permission marketing and directly examine the opt-in and opt-out behaviors in this study.

The three main characteristics of permission marketing are "anticipated, personal, and relevant" (Godin 1999, p. 40). In contrast to spam, a permission-based message is anticipated, and customers trust its sender (we believe that customers will not join the firm's e-mail program in the first place if they do not trust the firm). Firms can personalize the marketing messages according to customers' specific interests, which customers can indicate at the time of their opt-in decision. To improve targeting precision, marketers also can tailor the promotional information included in the message on the basis of the customer's past purchase behavior. Gartner (see Online Media Daily 2002) reports that unsolicited direct mail or e-mail has a response rate of 1%, whereas the average click-through rate of permission-based e-mails is between 6% and 8%. Jolley et al. (2013) show that a permission-based e-mail marketing program can extend a customer's lifetime value.

Firms must manage two critical aspects to ensure the success of a permission-based marketing program: the customer's opt-in and opt-out timing. Research on permission marketing has explored several factors that influence a customer's willingness to give permission to marketers, including brand equity, a previous relationship (Tezinde, Smith, and Murphy 2002), income, gender, advertising message volume, previous experience with mobile ads (Barnes and Scornavacca 2008), and brand image and trust (Jayawardhena et al. 2009). Whereas customers' opt-in decisions are influenced by the aforementioned factors, it is also important to identify the drivers of customers' opt-out decisions so that firms can make targeted efforts to retain their existing subscribers. Previous research on customers' opt-out decisions has discovered that message relevance and monetary benefit positively influence customers' interest in a permission marketing program (Krishnamurthy 2001), that highly personalized messages (e.g., using the customer's name in the e-mail subject line) tend to make customers opt out (Marinova, Murphy, and Massey 2002), and that lengthier e-mails and those with fewer links lead to higher unsubscribe rates (Chittenden and Rettie 2003).

Linkage Between Opt-In and Opt-Out

Although previous research has identified many factors that could influence customers' opt-in and opt-out behavior, it has mainly focused on the incidence of opting in and opting out but has not studied the timing of the two decisions or the possible linkage between the two. The timing of customers' opt-in and opt-out decisions depends on who they are (sociodemographics), how they live (lifestyle, online habits), how they are influenced (marketing contacts), and how satisfied they are (relevant messages). Some customers may opt in the first time they interact with the firm (i.e., made a purchase) and opt out at end of their customer life cycle. Some customers may need more time trying and testing the firm before they opt in, and they may only stay with the e-mail program for a limited time and withdraw as soon as they believe the program fails to meet their expectations. Although there is much heterogeneity in customer opt-in and opt-out behavior, we argue that there might be a dependence between the two variables and that ignoring this dependence can lead to biased inferences that can adversely affect the marketer's decision making.

Broadly speaking, customers' opt-in and opt-out times may be positively or negatively correlated. The nature and extent of their dependence could be determined by the following factors. First, opt-in and opt-out decisions have some drivers, such as marketing activities, in common. For example, if direct mail substitutes for e-mail before a customer opts in but complements e-mail after the customer opts in, direct mail would extend both the customer's opt-in time and his or her opt-out time, leading to a positive dependence between the two. In contrast, if direct mail is always a substitute for or complement of e-mail, customers' opt-in and opt-out times would be negatively correlated. Second, observed heterogeneity (e.g., customer characteristics) affects a person's decision to opt in and opt out. For example, customer "inertia" makes customers delay their decision to opt in, and after they have opted in, they tend to stay for a long time and do not bother to opt out. In this case, opt-in and opt-out times are positively correlated. In contrast, "variety-seeking" customers are reluctant to remain with one company, so they need more time to sign up, but after they have opted in, they will quickly switch to another program for a better offer. In such cases, customers will demonstrate a negative dependence between the opt-in and the opt-out times.

Third, the effectiveness of the e-mail program, such as the number of e-mail programs to which a customer has already subscribed and the relevance of the e-mail messages of the focal firm, may influence the customer's opt-in and opt-out likelihood. Customers who have already subscribed to a large number of e-mail programs are less likely to opt in to another one, and after they have opted in and are able to receive personalized relevant messages, they tend to stay for a long time. In this case, their opt-in and opt-out times are positively correlated. In contrast, if the same customers receive many nonrelevant messages after opting in, they will opt out quickly to reduce the pressure of information overload. In such cases, their opt-in and opt-out times will be negatively correlated. Although the dependence between opt-in and opt-out times may vary across firms and industries, researchers should empirically test the true dependence between them based on data. The scope of the current study is not to generalize whether the dependence should be positive or negative and offer explanations for such phenomena but simply to capture the dependence through an empirical model.

Capturing the Dependence Between Durations

Accounting for dependence between two durations, such as the dependence between acquisition and retention and between e-mail open and click, is not uncommon in the marketing literature. Chintagunta and Haldar (1998) adopt the Farlie–Gumbel–Morgenstern family of bivariate distributions to capture the dependence between customer purchase of products in two related categories, such as pasta and pasta sauce. Park and Fader (2004) adopt the Sarmanov bivariate distributions to investigate customer covisit timing behavior between the websites of two competing retailers. Schweidel, Fader, and Bradlow (2008) use the Sarmanov family to model the dependence between the time to customer acquisition and the subsequent duration of being "alive." Bonfrer and Drèze (2009) develop hazard models of e-mail open and click times with the Sarmanov family to capture the dependence between open and click rate.

Notably, Schweidel et al. (2008) develop their model and apply it to a context similar to that of this study by jointly modeling the timing of when a customer starts to engage and disengage with a firm. However, the model we propose in this study is distinguished from Schweidel et al.'s (2008) in several aspects. First, the Sarmanov families are limited in the dependence ranges for which they can account (Danaher and Smith 2011). Schubina and Lee (2004) note that the dependence range for the Sarmanov family depends on the specification of marginal distributions. They calculate the exact maximum dependence ranges that can be attained for several marginal distribution specifications—for example, the range of uniform is $[-3/4, 3/4]$ and of normal is $[-2/\pi, 2/\pi]$. While the Sarmanov family may be applied effectively in some contexts, such as Park and Fader's (2004) and Schweidel et al.'s (2008), we prefer to use copulas that can accommodate a wider range of dependence. In the empirical application of this study, we test both the Gaussian and the Frank copulas, two copulas that can account for nearly the full $(-1, 1)$ range of dependence (Trivedi and Zimmer 2005).

Second, although a bivariate model has advantages in solving marketing problems, real-world applications may require a model that can capture complex and multidimensional dependence structures. The Sarmanov families do not easily capture the dependence structure of three or more dimensions (Danaher and Smith 2011). In this study, we propose to use a vine copula (e.g., Aas et al. 2009), recently popularized in the statistics literature, to jointly model opt-in time, opt-out time, and purchase behavior. We discuss the method vine copula method in the "Proposed Modeling Framework" section.

Third, Schweidel et al. (2008) develop their model in a contractual telecommunications services context but do not consider the possibility that service subscription time is dependent on the types (low/medium/high margin) of service to which customers choose to subscribe. In this study, we investigate a noncontractual retailing context in which the duration of a customer staying in a marketing program and his or her purchases are two separate but interdependent behaviors (e.g., Ascarza and Hardie 2013; Netzer, Lattin, and Srinivasan 2008). While it may be argued that the effect of a customer's purchase behavior on the length of time he or she stays in a marketing program could be estimated by including it as a covariate in the marginal model, we determine that it would suffer from endogeneity because unobserved factors such as customer loyalty are highly likely to affect both variables. The vine copulas model we propose in this study avoids the potential endogeneity issue by modeling purchase behavior, opt-out time, and opt-in time simultaneously. In the next section, we discuss the substantive importance of jointly examining these phenomena.

Incorporation of Purchase Behavior

Krishnamurthy (2001) notes that customer interest in a permission marketing program is positively related to the customer's level of participation in the program. The author states that customers opt in to a marketing program to obtain information related to the products and promotions that add value to their lives by reducing the cost of information search and by providing monetary benefits. Most permission-based marketing programs allow customers to opt out or unsubscribe at any time if they are no longer willing to receive messages from the firm. The length of time a customer is willing to stay in a marketing program may depend on the relevance of the message, the intensity of marketing activities, and customer loyalty. We argue that customers who receive a higher proportion of relevant messages and/or who have a higher level of spending with the firm are more likely to stay in the marketing program longer. Firms can identify short-life customers at an earlier stage by analyzing their buying patterns (Reinartz and Kumar 2000).

From another angle, participation in a permission marketing campaign can change customers' attitudes and behaviors by increasing their purchase intention (DuFrene et al. 2005), encouraging them to spend more money (Jolley et al. 2013), and making them more responsive to firms' marketing messages (Marinova, Murphy, and Massey 2002). The longer customers stay in a marketing program, the more familiar they will be with the firm and the more likely they will be to shop with the firm. In summary, we argue that staying in a marketing program and actually making purchases are two interdependent processes (e.g., Danaher 2002) that should be jointly studied to avoid potential endogeneity issues. Firms should invest resources not only to encourage customers to stay longer in the marketing program but also to induce them to spend more money while they are still subscribed to the program. Because the timings of joining and withdrawing from the marketing program are also interdependent, we model the three processes jointly in an integrated model framework.

DATA DESCRIPTION

Our database comprises information from a U.S. retailer that sells multiple categories of home improvement products. The data set consists of information on the time a customer opts in and opts out of the firm's e-mail program, the transactions made by the customer, the e-mail open and click histories, and the marketing activities of the firm. We construct a calibration data set by randomly sampling a cohort of 9,180 customers who made their first purchases from the firm between February 2007 and July 2007. We construct a holdout data set by sampling another cohort of 9,180 customers to validate our proposed model.

To obtain information on customers' online habits and sociodemographics, a multinational marketing technology and services firm, Acxiom, merged the data we sampled with one of its databases using several key identifiers with a 100%, one-to-one match rate. The database provided by Acxiom, trademarked as PersonicX Digital, assigns people to 1 of 13 segments on the basis of how they use the Internet, how they shop online, when and where they access the Internet, and their demographic attributes (for a description of each cluster, see Table 1). We include this external segmentation to account for the customer characteristics that are useful in explaining customers' opt-in and opt-out propensities (Brey et al. 2007).

The retailer that provided the data currently operates a large-scale e-mail program with a substantial number of subscribers. The e-mail program is permission based in the sense that people must subscribe first to receive any e-mails from the retailer. Although purchase is not required to subscribe to the e-mail program, the majority of the existing e-mail subscribers have purchase histories with the firm before opting in, according to the retailer's management team. The number of e-mail subscribers who opt in on the same day as their first purchase is insignificant (approximately .03% of the sample), likely because, as we argue, customers need a period of time to develop trust with the firm before they agree to let the firm send messages to their e-mail inbox.

We view purchasing and subscribing as two separate decisions for customers of the focal retailer. The retailer does not have any policy in place to encourage customers to opt in to its e-mail program when they purchase from its physical stores. In addition, although customers can create an online account to manage their orders with convenience when they purchase online, they are considered to have opted in only after they click the check box "willing to receive further e-mail marketing messages." After they opt in, subscribers receive e-mails that contain instructions to opt out at the bottom of the message. Customers can opt out at any time by clicking the "unsubscribe" link, calling the customer service center, or writing to the retailer's office.

Table 1
DATABASE SEGMENTATION DESCRIPTIONS

Variables	*Segments Label*	*Average Age (Years)*	*Income/Wealth*	*Sociodemographics and Online Behavior*
Group 1	Superhighway Superusers	25–55	Medium to high	Extremely comfortable online user; likes sports, music, social networking, shopping, or investing
Group 2	Second Nature Surfers	24–39	Low to high	Frequent mobile user, no children or just started a family, online shopper, likes music, job search, online auctions, and social networking
Group 3	High-Speed Checkout	41–42	High	Online shopper; prefers apparel, toys, games, or travel
Group 4	Affluent Aficionados	56–68	High	Heavy online user, working, shopping, and investing
Group 5	Voluminous Variety	38–39	Medium to high	Heavy online user, either child-centric or pursuing personal hobbies such as news, sports, and travel
Group 6	My Internet, My Way	24–40	Medium to high	Fans of online social networking, job searches, and personal entertainment
Group 7	ECommerce Experts	55–70	Low to high	Heavy online shopper and online search (e.g., automobile category)
Group 8	Selective Surfers	54–58	Medium to high	Moderate online user; focuses on relaxing, social networking, investing, and shopping
Group 9	Rural Connections	41–58	Medium to high	Below-average online user; focuses on insurance quotes, sports apparel, or phone calls
Group 10	Senior Investors	67–78	High	Fans of online shopping and investing
Group 11	Functional Frequency	38–40	Low to medium	Home-centric, online usage mostly for job searches and some social networking
Group 12	Limited Logons	56–58	Low to medium	Low online usage, mostly evenings or weekends
Group 13	Sans Surfers	67–78	Low to medium	Very low online activities; prefers traditional channels such as direct mail and telephone

Source: Acxiom PersonicX Digital.

There are two characteristics of this study we need to clarify. First, this study focuses on the opt-in and opt-out behaviors of existing customers. We acknowledge that other firms may have a proportion of e-mail subscribers who have no purchase history before they opt in. Although it may be worthwhile to examine the opt-in behavior of prospective customers, the managerial implications we draw from this study apply to existing customers. Second, we focus on customers' first opt-in and opt-out decisions. We observe only a small number of customers who have multiple opt-in and opt-out records (approximately .1% of the sample).

DESCRIPTIVE STATISTICS

The key variables of interest in this study are the timings of the opt-in and opt-out decisions. We compute opt-in time as the number of days that have elapsed between a customer's first purchase and opting in. We compute opt-out time as the number of days elapsed between the opt-in and opt-out decisions. We observe that both opt-in and opt-out times may be right-censored. In the calibration sample, 22.8% of the customers did not opt in, 18.5% of the customers opted in but opted out before the end of the observation window, and 58.7% of the customers opted in and stayed until the end of the observation window. Of the customers who opted in, the mean opt-in time was 597 days, and the median opt-in time was 611 days. Of the customers who opted in but then opted out, the mean opt-out time was 410 days, and median opt-out time was 343 days.

To illustrate the differences in purchase behavior of e-mail subscribers and nonsubscribers, we randomly select two samples of equal size (subscribers and nonsubscribers) and report the descriptive statistics of several variables computed for the same period of time (see Table 2). Compared with nonsubscribers, e-mail subscribers spend more money, purchase more frequently, redeem more coupons, receive more direct mails, and return more items. The results are consistent with previous findings that permission-based

Table 2
SUMMARY STATISTICS OF BEHAVIORAL VARIABLES FOR SUBSCRIBERS AND NONSUBSCRIBERS

		E-Mail Nonsubscribers Sample		*E-Mail Subscribers Sample*[b]	
Variables	*Operationalization*[a]	*M*	*SD*	*M*	*SD*
Total_Money ($)	The total amount of money spent	4,440	5,755	5,430	6,503
Total_Freq	The total number of purchase occasions	41.50	52.65	51.05	52.37
Total_Coupon	The total number of coupons redeemed	1.17	2.57	1.77	3.25
Total_Dmail	The total number of direct mail campaigns targeted	4.44	5.31	4.81	5.30
Total_Return_Freq	The total number of return occasions	4.36	7.67	5.58	8.35
Total_Cross_Buy	The average number of product categories purchased per transaction	2.66	1.00	2.73	.93

[a]We computed the variables for the period February 2007 to December 2010.
[b]E-mail nonsubscribers did not opt in from February 2007 to December 2010. E-mail subscribers opted in on February 2007 and stayed in the e-mail program until December 2010.

marketing programs reinforce customer loyalty and induce more active customer engagement.

To further illustrate the importance of studying opt-out time, we conducted a preliminary analysis to explore the relationship between the length of time a customer stays in an e-mail program and his or her purchase behavior. We randomly selected 103 customers who began their relationship with the retailer (first purchase) at the same time (February 2007), opted in to the e-mail program at the same time (June 2008), but opted out at different times. We split these customers into the following three cohorts according to the length of time the customer had been with the retailer: 1–6 months (cohort 1), 7–12 months (cohort 2), and 13–18 months (cohort 3). We summarize their purchase behavior for the same time window, June 2008–December 2010 (see Table 3). To ensure that the three cohorts of customers are comparable, we selected customers with a similar purchase pattern before opting in (e.g., making a purchase every 1.4–1.8 months).

Table 3 shows that, on average, the customers who stayed in the e-mail program for a longer period tend to purchase more frequently and spend more money. The statistics can be interpreted from two perspectives. From one point of view, the customers who choose to stay longer in the program demonstrate stronger interest in the product category and the brand, have a greater chance of being exposed to the firm's e-mail marketing, and are more active in their purchasing. From another point of view, the customers who have longer-lasting interests in home improvement products, who are more accustomed to reading e-mails to obtain information, and who have a greater intention to purchase are more likely to stay in the e-mail program for a longer period of time.

PROPOSED MODELING FRAMEWORK

Modeling Challenges

In this study, we jointly model three variables: the customer's opt-in time, opt-out time, and average transaction amount. We need a multivariate copula model that can capture the three-dimensional dependence structure. Multivariate copula models have received attention from the fields of statistics, finance, insurance (e.g., Smith et al. 2010; Zimmer and Trivedi 2006), and marketing (e.g., Danaher and Smith 2011; Kushwaha and Shankar 2013; Stephen and Galak 2012). However, the number of multivariate distributions that are readily applicable to three- or higher-dimensional problems is limited. The multivariate Gaussian copula, an example of the elliptical copula, has been used to model intermagazine exposures and page views of multiple websites (Danaher and Smith 2011) and model multivariate count data (Stephen and Galak 2012). In addition to the elliptical copula, several studies have attempted to extend the bivariate Archimedean copula to higher dimensions (e.g., Savu and Trede 2010; Zimmer and Trivedi 2006). The most commonly used Archimedean copulas include Clayton, Gumbel, and Frank (Trivedi and Zimmer 2005). However, these extensions are developed at the expense of dependence measures. A flexible n-variate copula should be able to accommodate $n(n - 1)/2$ dependence parameters for each pair of the marginal distributions. However, for example, a trivariate extension of a bivariate Frank copula, which Zimmer and Trivedi (2006) propose, only allows for two (rather than three) dependence parameters, which also must be positive. This restriction limits its application to many practical problems. Thus, a flexible multivariate copula model is needed.

Drawing on the work of Joe (1997) and Bedford and Cooke (2002), Aas et al. (2009) show that multivariate data can be decomposed into a cascade of bivariate copulas, called "pair-copula constructions." A pair-copula decomposition offers a highly flexible way to construct multivariate distributions and has been the focus of many recent studies (e.g., Hobæk Haff 2013; Kurowicka and Joe 2011; Min and Czado 2010; Panagiotelis et al. 2012; Smith et al. 2010). It has no restrictions on the number of dependence parameters. It allows the selection of any copulas to build bivariate copulas, such as Gaussian, t, Gumbel, and Frank. Compared with estimation of multivariate Gaussian copulas, pair-copula construction estimation is relatively easy because the parameters of each pair-copula can be estimated sequentially. However, a multivariate Gaussian copula requires the evaluation of multiple integrals without a closed-form solution, which can only be approximated numerically. In a simulation study, Smith et al. (2010) compare the vine copula and the multivariate Gaussian copula and show that the vine copula outperforms a multivariate Gaussian copula in forecasting. In this study, we construct a trivariate pair-copula model and test it in an empirical application with Gaussian and Frank copulas as pair-copulas. We choose the "best-fitting" copula among the two models using model selection criteria such as the Bayesian information criterion (BIC). In the next section, we discuss the marginal models for opt-in time, opt-out time, and average transaction amount and the modeling of the dependence with pair-copula construction.

Modeling the Opt-In and Opt-Out Times

Because the opt-in and opt-out times are both continuous survival data, we model these variables using the conditional hazard model (e.g., Jain and Vilcassim 1991), which is well suited for censored observations. Let (T_{i1}, T_{i2}) and (C_{i1}, C_{i2}) denote the paired opt-in and opt-out times and censoring times for customer $i = 1, \ldots, n$. Let $t_{ij} = \min(T_{ij},$

Table 3

PURCHASE COMPARISONS OF CUSTOMERS WITH DIFFERENT OPT-OUT TIMES

	Sample Size (N = 103)	*Average Interpurchase Time Before Opt-In (Months)*	*Time Elapsed Between Opt-In and Opt-Out (Months)*	*Total Purchase Amount Between Opt-In and the End of Observation*	*Total Number of Purchase Occasions Between Opt-In and the End of Observation*
Cohort 1	27%	1.4–1.8	1–6	$2,262	20
Cohort 2	43%		7–12	$2,584	23
Cohort 3	30%		13–18	$3,022	34

C_{ij}) denote the actual observed durations, $\delta_{ij} = I(t_{ij} = T_{ij})$, and $\mathbf{Z_{ij}}$ be a vector of covariates for customer i, where the subscript j denotes the opt-in time (j = 1) or the opt-out time (j = 2). Note that $\delta_{i1} = 0$ denotes the case in which customer i did not opt in during the observation period.[1] The opt-in or opt-out time t_j is assumed to follow the Weibull distribution, characterized by the distribution function $F_j(t_j)$. We use the Weibull distribution because it is highly flexible; can accommodate flat, monotonically increasing or decreasing hazard functions; and has been proved useful in marketing applications (e.g., Chintagunta and Halder 1998; Seetharaman and Chintagunta 2003).

The density function of the Weibull distribution is $f_j(t_j) = \alpha_j\lambda_{ij}t_j^{\alpha_j - 1}\exp(-\lambda_{ij}t_j^{\alpha_j})$, where α_j represents the shape parameter and λ_{ij} controls the scale parameters. We allow the scale parameter to be customer-specific, λ_{ij}, which is specified as a function of the corresponding vector of covariates $\mathbf{Z}_{ij}$ and parameter sets $\beta_\mathbf{j}$. To ensure that the scales are positive, we use exponential specifications as follows:

$$(1)\qquad \lambda_{ij} = \exp\Bigg(\beta_{0j} + \sum_{d=1}^{12}\beta_{dj}\text{PERSONICX}_{id} + \beta_{13j}\text{COUPON}_{ij} + \beta_{14j}\text{DMAIL}_{ij} + \beta_{15j}\text{RETURN}_{ij} + \beta_{16,\,j=2}\text{EMAIL}_{i,\,j=2} + \beta_{17,\,j=2}\text{OPEN}_{i,\,j=2} + \beta_{18,\,j=2}\text{CLICK}_{i,\,j=2}\Bigg),$$

for every i = 1, …, n; j = 1, 2; and d = 1, …, 12. Here, β_{0j} captures a customer i's intrinsic probability to opt in or opt out. The variables EMAIL, OPEN, and CLICK are related to activities that can only occur after a customer has opted in to a permission-based e-mail program, so $\beta_{16,j=2}$, $\beta_{17,j=2}$, and $\beta_{18,j=2}$ are specified only in the opt-out model. We explain all the variables in Equation 1 next.

PERSONICX represents a vector of binary variables that indicate the segment to which a customer is assigned according to PersonicX Digital, the database from Acxiom. The database assigns customers to 1 of the 13 segments according to their demographics and online behaviors (see Table 1). We use Group 1, labeled "Superhighway Superusers," as the reference group to create 12 dummy variables. We expect these variables to provide some explanatory power for the opt-in and opt-out model because online habits and sociodemographics affect customers' interest in permission-based web or mobile-marketing programs (e.g., Brey et al. 2007).

We operationalize COUPON_{ij} as the total number of coupons that customer i redeemed before opting in (j = 1) or between the opt-in and opt-out decision or the censoring time (j = 2). In addition to direct mail, a company's website, or a referral, an e-mail program is another option customers can use to obtain savings opportunities, such as coupon codes or price discount information. We expect that customers who are already active in coupon redemption have a smaller probability of opting in due to high information-processing costs and low incremental saving benefit. Meanwhile, for e-mail subscribers, we expect that coupon redemption activities could indicate the relevance of e-mails to customers' purchase needs, which could subsequently affect their interest in the e-mail program (Krishnamurthy 2001).

We operationalize DMAIL_{ij} as the average number of direct mail customer i received per month before the opting in (j = 1) or between the opt-in and the opt-out decisions or the censoring time (j = 2). Direct mail usually uses product information and coupons to attract customers to visit the stores. While direct mail and e-mail serve similar marketing purposes, it is uncertain how they affect each other. It is likely that customers who receive a substantial amount of direct mail are less motivated to participate in an e-mail marketing program due to the increase of information burden (Krishnamurthy 2001). We test the nonlinear forms (logarithmic and quadratic) of DMAIL_{ij} in both models because there could be an optimal level of marketing communications (Nash 1993).

We operationalize RETURN_{ij} as the total number of product return occasions customer i made before the opting in (j = 1) or between the opt-in and the opt-out decision or the censoring time (j = 2). Research has shown that customers who have a medium level of returns have the highest customer lifetime value (e.g., Petersen and Kumar 2009). Product return frequency signifies the relationship between customers and firms, which is important for the customer opt-in and opt-out decisions (e.g., Jayawardhena et al. 2009). We test the nonlinear effect (logarithmic and quadratic) of RETURN_{ij} and expect to find an optimal level of product return frequency.

We operationalize $\text{EMAIL}_{i,\,j=2}$ as the average number of e-mails customer i received per month, we compute $\text{OPEN}_{i,\,j=2}$ by dividing the number of e-mails opened by the total number of e-mails received, and we compute $\text{CLICK}_{i,\,j=2}$ by dividing the number of e-mails clicked by the total number of e-mails opened, between the opt-in and the opt-out decision or the censoring time. Krishnamurthy (2001) indicates that message processing costs and message relevance are two important factors that could affect customers' interest in permission marketing programs. Ha (1996) argues that, because of the intrusive nature of e-mail promotions, customers' attitudes toward e-mail marketing will decrease as firms' e-mailing frequency increases. We expect to find a U-shaped effect of e-mail quantity on customer opt-out probability. In addition, we use the e-mail open and click rates as a measure of message relevance that could indicate the category–message fit and the perceived attractiveness of advertisers (Krishnamurthy 2001). Firms that can consistently send messages relevant to customers' needs will be more appreciated, and customers will be less likely to opt out. However, we expect that the utilities derived from relevant messages increase up to a threshold, as customers usually have a spending limit or a share of wallet for a certain firm. Thus, we use the linear and quadratic form of $\text{EMAIL}_{i,\,j=2}$ and the logarithmic forms of $\text{OPEN}_{i,j=2}$ and $\text{CLICK}_{i,j=2}$.

[1]In this study, we assume that every customer will eventually opt in to the retailer's e-mail program given a long enough period of time. A split-hazard model can be used to account for the opt-in probability if one makes the assumption that a proportion of customers will never opt in (e.g., Schweidel, Fader, and Bradlow 2008; Sinha and Chandrashekaran 1992).

Heterogeneity

In addition, there is unobserved heterogeneity in terms of customers' hazard of the opt-in and the opt-out. For example, some customers may be more likely to opt in or opt out, but this heterogeneity is not directly measured. To account for this unobserved heterogeneity in both the opt-in and the opt-out models, we incorporate an unobservable multiplicative effect v_{ij}, called a frailty term, on the Weibull hazard functions (e.g., Han and Hausman 1990; Schmittlein and Morrison 1983). The conditional hazard function is specified as $h(t_{ij}|v_{ij}) = \alpha_j\lambda_{ij}t_{ij}^{\alpha_j-1}v_{ij}$. Following Sahu et al. (1997), we assume that the random variable v_{ij} follows a gamma distribution with a mean of 1 (for identification purposes) and a variance of $1/\gamma_j$, where γ_j is a parameter to be estimated. By integrating v_{ij}, we obtain the closed-form solutions for the unconditional Weibull survival function (e.g., Gutierrez 2002; Meade and Islam 2010),

$$(2) \qquad S(t_{ij}) = \left[1 + \gamma_j\lambda_{ij}t_{ij}^{\alpha_j}\right]^{-\frac{1}{\gamma_j}},$$

and the unconditional Weibull density function,

$$(3) \qquad f(t_{ij}) = S(t_{ij})^{1+\gamma_j}\alpha_j\lambda_{ij}t_{ij}^{\alpha_j-1}.$$

Modeling the Average Transaction Amount

We assume that the average transaction amount (in U.S. dollars), AMT_t, that customer i spent during the time he or she stayed with the e-mail program follows a log-normal distribution (e.g., Borle, Singh, and Jain 2008):

$$(4) \qquad \log AMT_i \sim Normal(\mu_i, \sigma^2),$$

where μ_i is the mean and σ^2 is the variance of the normal distribution. We assume that the mean parameter μ_i is a function of the individual-level covariates, as the following equation shows:

$$(5) \qquad \mu_i = \mu_{0i} + \mu_1 Avg\_Coupon_i + \mu_2 Avg\_Dmail_i + \mu_3 Avg\_Email_i + \mu_4 Avg\_Return_i + \mu_5 Avg\_CrossBuy_i + \mu_6 Avg\_Open_i + \mu_7 Avg\_Click_i + \mu_8 Avg\_IPT_i.$$

To account for the unobserved heterogeneity, we allow the intrinsic average transaction amount μ_{0i} to be customer specific. We assume that this heterogeneous parameter is normally distributed across customers as $\mu_{0i} = \mu_0 + \Delta\mu_{0i}$, where $\Delta\mu_{0i} \sim N(0, \sigma^2_{\mu_0})$ and $\sigma^2_{\mu_0}$ is the variance parameter. Thus, μ_0–μ_8, σ^2, and $\sigma^2_{\mu_0}$ are the parameters to be estimated from the data.

Note that log AMT is actually a mixture of two normals, one for the idiosyncratic variation and one for the random effect. Here, we use average transaction amount instead of total amount spent because total amount spent is a cumulative measurement that is likely to be a function of time elapsed. The joint modeling of opt-out time and total amount spent would create a positive dependence because of the shared effect of time elapsed. Because average transaction amount is calculated by dividing the total amount spent by the total number of purchase trips, the joint modeling of average transaction amount and opt-out time can capture the dependence that has teased out the shared effect of time elapsed. Next, we explain all the variables specified in Equation 5.

We operationalize AVG_Coupon_i as the average number of coupons customer i redeemed in every transaction. The use of coupons can lead to unplanned purchases and increase the amount of money a customer typically spends (e.g., Heilman, Nakamoto, and Rao 2002). However, highly price-conscious or deal-prone customers typically have budget constraints and tend to pay reduced prices (e.g., Völckner 2008). We expect that customers who redeem a medium level of coupons have the biggest shopping basket.

We operationalize Avg_Dmail_i and Avg_Email_i as the average number of direct mails or e-mails customer i received between two transactions. Marketing communications can retain existing customers and increase brand loyalty. However, excessive marketing contacts could be detrimental to the firm–customer relationship (e.g., Venkatesan and Kumar 2004). We expect to identify an optimal level of marketing contacts.

We computed Avg_Return_i as the average number of product return occasions customer i made for each transaction. We computed $Avg\_CrossBuy_i$ as the average number of product categories customer i purchased in each transaction. The product return frequency has an inverted U-shaped effect on the firm–customer relationship (Petersen and Kumar 2009). Customers who buy from multiple categories tend to shop from a wider range of products in a purchase occasion (Venkatesan and Kumar 2004). Similarly, we expect to find an inverted U-shaped effect of product return and a positive effect of cross-buying on the average transaction amount.

We operationalized Avg_Open_i and Avg_Click_i as the number of e-mails customer i opened or clicked between two transactions. Permission-based e-mail messages can increase customers' trust in the firm, their purchase intentions, and their lifetime values (e.g., DuFrene et al. 2005; Jolley et al. 2013). We expect that customers with higher e-mail open and click-through rates are more interested in the firm and spend more money with the firm.

We operationalized Avg_IPT_i as customer i's average interpurchase time, which is computed across the customer's purchase history between the opt-in and the opt-out or the censoring time. We use the average interpurchase time as a control variable and expect that customers who have shorter interpurchase time spend less money on each transaction.

Modeling the Dependence Using Pair-Copula Construction

Let $\mathbf{X} = (X_1, X_2, X_3)$ be a vector of random variables with a joint density function as $f(x_1, x_2, x_3)$. We demonstrate how to decompose the joint density into a cascade of pair-copulas. First, in line with Sklar's (1959) theorem, we can express the bivariate joint density as follows:

$$(6) \qquad f(x_1, x_2) = c_{12}[F_1(x_1), F_2(x_2)] \times f_1(x_1) \times f_2(x_2),$$

where $F_1(x_1)$ and $F_2(x_2)$ are continuous marginal distributions and $c_{12}(\cdot)$ is the bivariate pair-copula density. Drawing on basic probability theory, we can obtain the conditional density as follows:

$$(7) \qquad f(x_2|x_1) = c_{12}[F_1(x_1), F_2(x_2)] \times f_2(x_2).$$

The conditional density in a three-dimensional case is given by

$$(8)\ f(x_2|x_1, x_3) = \frac{f(x_2, x_3|x_1)}{f(x_3|x_1)}$$

$$= \frac{c_{23|1}[F(x_2|x_1), F(x_3|x_1)] \times f(x_2|x_1) \times f(x_3|x_1)}{f(x_3|x_1)}$$

$$= c_{23|1}[F(x_2|x_1), F(x_3|x_1)] \times f(x_2|x_1),$$

where $c_{23|1}(\cdot)$ is the suitable bivariate pair-copula density for $F(x_2|x_1)$ and $F(x_3|x_1)$. Note that $c_{23|1}(\cdot)$ captures the dependence between two reduced conditional distributions.

From Equations 6–8, we can record the joint density $f(x_1, x_2, x_3)$ using pair-copulas and the marginal densities. Applying the same logic to our empirical problem, we can construct the joint density function of opt-in time, opt-out time, and average transaction amount. Because the opt-in and the opt-out times are data of lifetimes, we use survival copulas in our specification (e.g., Nelsen 2006; Shih and Louis 1995):

$$(9)\quad f(t_{i1}, t_{i2}, AMT_i) = c_{23|1}[S(t_{i2}|t_{i1}), S(AMT_i|t_{i1}); \Omega_{23|1}]$$
$$\times\ c_{12}[S_1(t_{i1}), S_2(t_{i2}); \Omega_{12}]$$
$$\times\ c_{13}[S_1(t_{i1}), S_3(AMT_i); \Omega_{13}]$$
$$\times\ f_1(t_{1i}) \times f_2(t_{2i}) \times f_3(AMT_i),$$

where $c_{23|1}(\cdot)$, $c_{12}(\cdot)$, and $c_{13}(\cdot)$ are the density functions of the associated survival copulas[2]; $S_1(t_{i1})$, $S_2(t_{i2})$, and $S_3(AMT_i)$ are the marginal survival functions defined in Equations 2 and 4; $S(t_{i2}|t_{i1})$ and $S(AMT_i|t_{i1})$ are the conditional survival functions with a common conditioning variable t_{i1}; $f_1(t_{1i})$, $f_2(t_{2i})$, and $f_3(AMT_i)$ are the marginal densities defined in Equations 3 and 4; and $\Omega_{23|1}$, Ω_{12}, and Ω_{13} are the pair-copula parameters (see also Panagiotelis, Smith, and Danaher 2014).

Equation 9 only applies to the situation in which the opt-in and opt-out times are both observed; however, we must consider the cases in which the opt-in or opt-out times are censored. First, when the opt-in time is observed but the opt-out time is censored, we need to evaluate the conditional survival function $S(t_{i2}|t_{i1})$, which gives the probability that customer i stays in the e-mail program for at least time t_{i2} given that the customer's the opt-in time is t_{i1}. The conditional survival function is given by the first partial derivative of the bivariate copula function (He and Lawless 2003):

$$(10)\qquad S(t_{i2}|t_{i1}) = \frac{\partial C_{t_{i1}, t_{i2}}[S_1(t_{i1}), S_2(t_{i2}); \Omega_{12}]}{\partial S_1(t_{i1})},$$

where $C_{t_{i1},t_{i2}}(\cdot)$ is the bivariate copula function and Ω_{12} is the bivariate copula parameter defined in Equation 9. Second, when both the opt-in and the opt-out times are censored, the customer-level likelihood is given by the marginal survival function $S_1(t_{i1})$ (He and Lawless 2003).

The bivariate pair-copulas can be specified as Gaussian, t, Gumbel, Clayton, Frank, and so on. We empirically test Gaussian and Frank copulas in this study. We provide the copula distribution function, density, first partial derivative, and its inverse function for the Gaussian and Frank bivariate copulas in the Web Appendix.

Model Estimation

We have two sets of parameters to estimate, one of the marginal models and the other of the pair-copulas. Following Shih and Louis (1995) and Danaher and Smith (2011), we use a two-step procedure,[3] which yields consistent estimates for all parameters. In the first step, we estimate the parameters of each marginal model using maximum likelihood estimation. We specified the marginal likelihoods previously in Equations 1–5.

In the second step, we estimate the set of pair-copula parameters, assuming the parameters estimated from the first step to be fixed. We maximize the log-likelihood function given as follows:

$$(11)\quad LL(\Omega) = \sum_{i=1}^{N} \Big(\delta_{i1}\delta_{i2} \log\{c_{12}[S_1(t_{i1}), S_2(t_{i2}); \Omega_{12}]\}$$
$$+\ \delta_{i1}(1-\delta_{i2})\log[S(t_{i2}|t_{i1}); \Omega_{12}]$$
$$+\ \delta_{i1}\log\{c_{13}[S_1(t_{i1}), S_3(AMT_i)]; \Omega_{13}\}$$
$$+\ \delta_{i1}\delta_{i2}\log\{c_{23|1}[S(t_{i2}|t_{i1}), S(AMT_i|t_{i1})]; \Omega_{23|1}\}\Big),$$

where δ_{i1} and δ_{i2} are the indicator variables that equal 1 when t_{1i} and t_{2i} are observed and 0 otherwise; and $\boldsymbol{\Omega} = \{\Omega_{12}, \Omega_{13}, \Omega_{23|1}\}$ are the pair-copula parameters to be estimated. Note that for the customers who did not opt in ($\delta_{i1} = 0$), we use these observations in estimating the marginal opt-in model. However, we do not use them in the estimation of the copula dependence parameters, because the dependence relies on the observed opt-in time.

Following Aas el al. (2009), we estimate these parameters sequentially. We first estimate Ω_{12} and Ω_{13} by maximizing the first three terms of Equation 11. Second, we calculate $S(t_{i2}|t_{i1})$ and $S(AMT_i|t_{i1})$ in a way analogous to that in Equation 10, using the estimates from the first step. Third, we estimate $\Omega_{23|1}$ by maximizing the fourth term of Equation 11. Finally, we use the estimates from the previous three steps as our starting point and maximize the full log-likelihood specified in Equation 11. We recover all the parameters in the simulation. We provide the data-generating algorithm for simulation study in the Web Appendix. We also provide estimation details of the proposed model with Frank pair-copulas in the Appendix.

RESULTS

Main Findings

We estimate the model specified in Equations 1–11 with Gaussian and Frank as pair-copulas using a maximum likeli-

[2]For distributions of high dimensions, the number of unique pair-copula decompositions increases significantly. Vine copulas, initially introduced by Joe (1997) and Bedford and Cooke (2002), provide a graphical way to organize the pair-copula construction conveniently. For a general way to construct a multidimensional distribution through vine representation, see the Web Appendix.

[3]Pair-copula constructions can also be estimated using a Bayesian method (Min and Czado 2010).

hood estimation in GAUSS. We compare the log-likelihood and the BIC of the two models to choose the "best-fitting" model. Table 4 gives the in-sample log-likelihood and the BIC. On the basis of the log-likelihood and the BIC, we determined that the proposed model with pair-copulas specified as Frank copulas provides a better fit to the calibration data. Thus, we chose the Frank copula specification in this study. Table 5 reports the estimates of the marginal models of the opt-in time, the opt-out time, and the average transaction amount. We discuss the results of each model in the following subsections.

Opt-in time model estimates. We examine how a customer's opt-in decision is affected. As the logarithm of γ_1 is –.600, we calculate that the variance of the gamma frailty term equals 1/exp(–.600) = 1.82, indicating a strong degree of heterogeneity in customers' opt-in decisions. Some customers are more prone to opt in than others with the same covariate value. After controlling for the unobserved heterogeneity, some notable findings emerged. Coupon redemption frequency has a negative effect on customers' opt-in probability. One of the main benefits of an e-mail program

Table 4

MODEL FIT COMPARISON

	The Proposed Model with Pair-Copulas Specified as	
	Gaussian	*Frank*
Log-likelihood	–1,281	–1,265
BIC	2,590	2,558

Table 5

PARAMETER ESTIMATES

Covariates	*Estimates*	*SE*	*T-Value*
Opt-In Time Hazard Function			
Intercept	.914	.069	13.192
$COUPON_1 \times 10^{-1}$	–.273	.091	–2.992
$DMAIL_1$ (log)	–1.294	.052	–24.691
$RETURN_1 \times 10^{-1}$ (log)	–.770	.062	–12.486
PERSONICX DUMMY 1 (Second Nature Surfers)	–.194	.088	–2.206
PERSONICX DUMMY 2 (High-Speed Checkout)	–.059	.076	–.776
PERSONICX DUMMY 3 (Affluent Aficionados)	.080	.068	1.172
PERSONICX DUMMY 4 (Voluminous Variety)	–.179	.070	–2.566
PERSONICX DUMMY 5 (My Internet, My Way)	–.123	.080	–1.533
PERSONICX DUMMY 6 (ECommerce Experts)	–.008	.065	–.121
PERSONICX DUMMY 7 (Selective Surfers)	–.047	.074	–.641
PERSONICX DUMMY 8 (Rural Connections)	–.091	.075	–1.204
PERSONICX DUMMY 9 (Senior Investors)	.081	.080	1.013
PERSONICX DUMMY 10 (Functional Frequency)	–.038	.082	–.465
PERSONICX DUMMY 11 (Limited Logons)	–.111	.088	–1.260
PERSONICX DUMMY 12 (Sans Surfers)	.023	.087	.261
Weibull Shape α_1 (log)	.517	.016	33.017
γ_1 (log)	–.600	.080	–7.467

Table 5

CONTINUED

Covariates	*Estimates*	*SE*	*T-Value*
PERSONICX DUMMY 9 (Senior Investors)	.081	.080	1.013
PERSONICX DUMMY 10 (Functional Frequency)	–.038	.082	–.465
PERSONICX DUMMY 11 (Limited Logons)	–.111	.088	–1.260
PERSONICX DUMMY 12 (Sans Surfers)	.023	.087	.261
Weibull Shape α_1 (log)	.517	.016	33.017
γ_1 (log)	–.600	.080	–7.467
Opt-Out Time Hazard Function			
Intercept	3.676	.425	8.642
$COUPON_2 \times 10^{-2}$	–6.825	2.406	–2.837
$DMAIL_2$	–2.224	.185	–12.052
$(DMAIL_2)^2$	.488	.046	10.656
$RETURN_2 \times 10^{-1}$	–1.713	.227	–7.562
$(RETURN_2 \times 10^{-1})^2$	.265	.065	4.046
$EMAIL_2 \times 10^{-1}$	–1.194	.331	–3.602
$(EMAIL_2 \times 10^{-1})^2$	.283	.136	2.081
$OPEN_2$	–1.256	.197	–6.374
$CLICK_2$	–2.413	.270	–8.933
PERSONICX DUMMY 1 (Second Nature Surfers)	–.552	.311	–1.777
PERSONICX DUMMY 2 (High-Speed Checkout)	–.238	.271	–.878
PERSONICX DUMMY 3 (Affluent Aficionados)	–.450	.255	–1.769
PERSONICX DUMMY 4 (Voluminous Variety)	–.381	.260	–1.468
PERSONICX DUMMY 5 (My Internet, My Way)	–.642	.281	–2.283
PERSONICX DUMMY 6 (ECommerce Experts)	–.027	.236	–.113
PERSONICX DUMMY 7 (Selective Surfers)	.106	.269	.395
PERSONICX DUMMY 8 (Rural Connections)	–.112	.272	–.413
PERSONICX DUMMY 9 (Senior Investors)	.533	.263	2.028
PERSONICX DUMMY 10 (Functional Frequency)	–.735	.290	–2.532
PERSONICX DUMMY 11 (Limited Logons)	–.229	.297	–.770
PERSONICX DUMMY 12 (Sans Surfers)	.209	.301	.695
Weibull Shape α_2 (log)	.670	.050	13.483
γ_2 (log)	1.994	.099	20.160
Average Transaction Amount[a]			
Intercept	–4.021	.052	–77.157
Avg_Coupon $\times 10^{-1}$	17.379	1.399	12.423
(Avg_Coupon $\times 10^{-1})^2$	–82.609	17.021	–4.853
Avg_Dmail (log)	.177	.020	8.729
Avg_Email $\times 10^{-3}$	11.040	1.128	9.786
(Avg_Email $\times 10^{-3})^2$	–74.068	8.070	–9.179
Avg_Return $\times 10^{-1}$	9.460	.906	10.437
(Avg_Return $\times 10^{-1})^2$	–46.566	9.284	–5.016
Avg_CrossBuy (log)	.829	.036	22.880
Avg_Open $\times 10^{-3}$ (log)	4.690	1.545	3.035
Avg_Click $\times 10^{-3}$ (log)	–6.444	6.326	–1.019
Avg_IPT $\times 10^{-1}$ (months)	.071	.006	11.557
σ^2 (log)	–.219	.009	–24.377
$\sigma^2_{\mu_0}$ (log)	–4.620	.850	–5.435
Pair-Copula Dependences			
Ω_{12}[b]	1.081 (.204)	.135	7.993
Ω_{13}	–.261 (–.043)	.153	–1.701
$\Omega_{23\|1}$	.567 (.094)	.128	4.444

[a]Average transaction amount is scaled by 10^{-3}.

[b]The corresponding Spearman's rho in the parentheses (for the transformation of the dependence measures, see Trivedi and Zimmer 2005).

is the savings opportunities delivered through e-mails. If a customer has already been active in redeeming coupons that may be distributed through other channels such as direct mail, a company website, or referral, it is unlikely that the customer will turn to another marketing program because the marginal benefit would not be high enough.

We find that the number of direct mails a customer receives has a negative effect on the opt-in probability, but this effect diminishes with an increase in the quantity of direct mail. Consistent with Barnes and Scornavacca (2008), this finding suggests that the marketing exposure a customer receives affects his or her decision to opt in. Because direct mail and e-mail share similar marketing functions, the customers who are already contacted through many direct mails are less likely to join another marketing program, which could increase their information-processing burden (Krishnamurthy 2001). In addition, the product return frequency has a negative but diminishing effect on the opt-in probability. This finding is consistent with the previous literature (e.g., Petersen and Kumar 2009) that customers with a moderate amount of product returns are the ones in which firms should invest resources to further build the relationship.

Furthermore, customers with different characteristics and online habits have different opt-in propensities, according to the estimates of the PersonicX variables. Customers who belong to the groups labeled “Second Nature Surfers” and “Voluminous Variety” are statistically significantly different in opt-in likelihood from those who belong to the reference group labeled “Superhighway Superusers,” while the rest of the customers do not show significant differences from the reference group. “Second Nature Surfers” are heavy online users; range in age from 24 to 39 years; either have no children or have just started a family; and prefer youth-oriented activities such as music, social networking, and online auctions (see Table 1). Consumers in this group are less likely to opt in to the retailer’s e-mail program, likely because they have less interest in or limited use for home improvement products. “Voluminous Variety” customers are familiar with and tend to use the Internet to obtain information on a daily basis (see Table 1). Consumers in this group would not join the retailer’s e-mail program easily, likely because they have better ways or alternatives to obtain product and promotion information. These findings are also consistent with previous studies (e.g., Brey et al. 2007) in that sociodemographics and information search behavior affect customers’ interest in permission marketing.

Opt-out time model estimates. We discuss how a customer’s opt-out decision is affected. Using the estimate of γ_1, we calculate the variance of the gamma frailty term as $1/\exp(1.994) = .14$, indicating a moderate level of heterogeneity in customers’ opt-out decisions. After controlling for the unobserved heterogeneity, we discovered that coupon redemption frequency has a strong negative effect on customers’ opt-out probability. If the savings opportunities delivered through e-mails are relevant to customers’ needs, customers would be more responsive by making more purchases with coupons. In such a case, customers have no reason to opt out from an effective e-mail program. In addition, the number of direct mails or e-mails a customer receives has a U-shaped effect on the opt-out probability. In line with previous research (e.g., Krishnamurthy 2001; Nash 1993), we find that too much communication is harmful to the firm–customer relationship and makes customers less interested in participation in the permission marketing program. Firms should plan an appropriate level of marketing intensity and avoid overmarketing to customers.

Furthermore, we find that e-mail open and click-through rate has a negative but diminishing effect on opt-out probability. It indicates that customers who open e-mail or click the links included in the e-mails more often are less likely to end their e-mail subscription. Consistent with Krishnamurthy (2001), this finding suggests that message relevance in terms of the category fit or incentive size is an important factor for customers to stay in a permission e-mail program. Firms may customize their e-mail messages on the basis of customers’ past purchases to increase the open and click-through rate. In line with Venkatesan and Kumar (2004), we also find that a moderate amount of product returns is healthy for the firm–customer relationship because it implies that customers are less likely to opt out.

From the estimates of the PersonicX variables, we discover some noteworthy results. Compared with customers of the reference and the other groups, customers who belong to the groups labeled “Second Nature Surfers,” “Affluent Aficionados,” “My Internet, My Way,” and “Functional Frequency” are statistically significantly less likely to opt out. It is also apparent that customers in these four groups likely stay in the e-mail program for different reasons. For example, “Functional Frequency” customers subscribe to the e-mail program to obtain promotional information because they are at an average age of 39 years, fall in the low- to middle-income bracket, and are raising a family. However, “Affluent Aficionados” likely opt in for new product information or gardening workshops because they typically are well educated and wealthy, are near retirement age, and have disposable time to pursue their personal hobbies.

Average transaction amount model estimates. We discuss the factors that determine the dollar amount a customer spends per transaction. The findings related to the average transaction amount are all consistent with previous research (e.g., Petersen and Kumar 2009; Venkatesan and Kumar 2004; Völckner 2008). We expect customers with a moderate level of coupon redemption or product return history to spend the most. A moderate level of marketing such as direct mail and e-mail is healthy for the firm–customer relationship, but overmarketing will decrease a customer’s purchase spending. Customers who buy across multiple product categories tend to spend more in each transaction. In addition, we find that the number of e-mails opened between two transactions has a positive but diminishing effect while the number of e-mail links clicked between two transactions has no statistically significant effect on the amount spent. Although the current available information does not allow us to link the e-mail open and click-through rate directly to purchases, we suspect that the retailer’s current e-mail strategy does not result in a good conversion rate. The e-mail messages sent probably have more of an advertising effect than they do an instantaneous promotional effect (e.g., Li, Sun, and Montgomery 2011). If so, we suggest the retailer should customize its e-mail messages, which could help improve the conversion rate.

Pair-copula dependences. Table 5 reports both the Frank copula parameter estimates and the transformed Spearman’s rho coefficients (in parentheses). Spearman's rho measures the rank-order correlation coefficient, which is not affected by the specification of the marginal distributions of the raw data (see Danaher and Smith 2011). The bivariate copula parameter, Ω_{12}, which has an estimate of .204 of Spearman's rho, shows a moderate dependence between opt-in and opt-out times. Customers who take a longer time to opt in tend to stay in the e-mail program for a longer time. The bivariate copula parameter, $\Omega_{23|1}$, which has an estimate of .094 of Spearman's rho, measures the dependence between opt-out time and average transaction amount, conditional on opt-in time. Caution should be exercised in interpreting $\Omega_{23|1}$ because it measures the dependence between two conditional distributions. However, if the interest is in the unconditional dependence, such as Ω_{12}, it can be obtained by permuting the variables specified in the vine structure in Equation 9 and reestimating the model.

Opt-Out Time Prediction

The key questions the manager of a permission-based e-mail program faces are as follows: When do e-mail subscribers opt out? And how can companies prevent them from leaving? In this section, we use the pair-copula model proposed in this study to predict the customer opt-out time at an individual level. Because the opt-out time is dependent on the opt-in time and the average transaction amount, to predict the mean opt-out time, we must evaluate the conditional expectation, which can be expressed as follows (e.g., Yeo and Valdez 2006):

$$(12) \qquad E\left(t_{i2}\middle|t_{i1}, AMT_i\right) = \int_0^\infty t_{i2} \times f_{2|13}\left(t_{i2}\middle|t_{i1}, AMT_i\right) dt_{i2},$$

where $f_{2|13}(\cdot)$ is given by $c_{23|1}[S(t_{i2}|t_{i1}), S(AMT_i|t_{i1})]$ $c_{12}[S_1(t_{i1}), S_2(t_{i2})]f_2(t_{i2})$ (see Equation 8). Because the integral in Equation 12 does not have a closed form, we solve the integral in a numerical way and obtain the prediction of mean opt-out time at the individual customer level.

We also compare the proposed model with four benchmark models for validation on the holdout sample. For comparison purposes, we only predict the opt-out time for customers who have completely observed opt-in and opt-out data. The four benchmark models are (1) univariate model of opt-out time (here, univariate means the model does not consider any dependence with other variables), (2) univariate model of the opt-out time with the observed opt-in time as a covariate, (3) bivariate model of the opt-out time and the average transaction amount with the observed opt-in time as a covariate (here, bivariate means a bivariate Frank copula model with the marginal models specified as in Equations 1–5), and (4) bivariate model of opt-in and opt-out time (see Table 6).

Armstrong, Morwitz, and Kumar (2000) define relative absolute error (RAE) as the mean absolute deviation of the error values of a model relative to that of the benchmark model. So a higher RAE indicates better predictive performance. As Table 6 shows, the proposed vine copula model gives the best prediction of mean opt-out time, as all the RAEs of the benchmark models are less than 1. If we compare the RAEs among the benchmark models, we find that by including the observed opt-in time as a covariate in the univariate model of opt-out time (RAE = .58), we do not improve the prediction accuracy compared with the univariate model without opt-in time as a covariate (RAE = .65). In addition, we discover that modeling the opt-in and opt-out times together (RAE = .70) slightly improves the predictive performance over the univariate model of opt-out time (RAE = .65). More important, the model that considers the dependence among the opt-in time, the opt-out time, and the average transaction amount improves the predictive performance significantly. In the next section, we demonstrate the changes of customer opt-in and opt-out times by simulating different levels of marketing activities.

Table 6

COMPARISON OF THE PROPOSED MODEL WITH BENCHMARK MODELS

	RAE[a] Relative to the Benchmark Model (with Respect to Opt-Out Time)
Proposed Model	
Trivariate model of opt-in, opt-out, and average transaction amount	—
Benchmark Models	
1. Univariate model of opt-out time	.65
2. Univariate model of opt-out time with observed opt-in time as a covariate	.58
3. Bivariate model of opt-out time and average transaction amount with observed opt-in time as a covariate	.61
4. Bivariate model of opt-in and opt-out time	.70

[a]Relative absolute error is defined as the mean absolute deviation in the opt-out time (in months) prediction of the proposed model relative to that of the benchmark model (see Armstrong, Morwitz, and Kumar 2000).

MANAGERIAL IMPLICATIONS

Marketing Policy Simulation

Using the parameter estimates (in Table 5), we can assess how a firm's marketing policy affects the opt-in and opt-out behavior of its customers. We conduct several simulations to show that firms can adjust their marketing contact frequency to strategically manage customers' opt-in and opt-out decisions. To do so, we select the customers who have opt-in and opt-out data from the holdout sample, constructing a sample of 1,696 customers. The average opt-in time of the sample is 15.9 months, and the average opt-out time is 13.4 months. By doing the simulations, we attempt to answer two questions: (1) What is the impact on customers' opt-in time if the retailer changes the direct mail marketing intensity before customers opt in? (2) What is the impact on customers' opt-out time if the retailer changes the direct mail and e-mail marketing intensity after customers opt in?

To answer the first question (see Scenario 1 in Table 7), we vary the average number of direct mail pieces a customer receives per month before opting in and predict the changes in customer opt-in time using the hazard model specified in Equations 1–3. As Table 7 shows, the increase in the frequency of direct mail contact could increase the time a customer takes to opt in to the e-mail program. For

Table 7
MARKETING POLICY SIMULATION RESULTS

	Scenario 1: E-Mail Program Opt-In Time (in Months)		Scenario 2: E-Mail Program Opt-Out Time (in Months)	
	Prediction	Changes	Prediction	Changes
No change on marketing intensity	23.0 (9.2)	—	17.5 (12.8)	—
Average number of direct mails per month +1	36.6 (10.2)	13.6	25.4 (12.1)	7.9
Average number of direct mails per month +2	48.7 (11.4)	25.7	27.0 (12.7)	9.5
Average number of direct mails per month +3	59.9 (12.8)	36.9	21.9 (13.4)	4.4
Average number of direct mails per month +4	70.6 (14.2)	47.6	12.1 (10.2)	–5.4
Average number of e-mails per month +5			19.7 (13.0)	2.2
Average number of e-mails per month +10			21.1 (12.9)	3.6
Average number of e-mails per month +15			21.7 (12.7)	4.2
Average number of e-mails per month +20			21.4 (12.6)	3.9
Average number of e-mails per month +25			20.1 (12.4)	2.6
Average number of e-mails per month +30			17.9 (21.1)	.4

Notes: Standard deviations are in parentheses.

example, one more direct mail per month would prolong the customer opt-in time for an average of 13.6 months, while four more direct mails per month would prolong the opt-in decision an average of 47.6 more months. This suggests that firms should be cautious of the cannibalizing effect that direct mail contact can have on their e-mail program subscription rate. If a firm already markets to its customers with a massive number of direct mails every month, its customers probably will not want to join the firm's e-mail program. While most firms may treat direct mail and e-mail as two separate marketing activities in practice, we suggest that firms should coordinate the two. Firms should closely monitor the return on investment of their marketing activities and allocate resources accordingly among different programs, such as direct mail, e-mail, web, and mobile to maximize company profits.

To address the second question (see Scenario 2 in Table 7), we vary the marketing contacts (direct mail and e-mail) the retailer sends to its customers after they opt in. Because the marketing contacts also influence purchase behavior, we predict the opt-out time using our proposed vine copula model, which considers the influence of the opt-in time and purchase (see Equation 12). As Table 7 shows, the increase of direct mail contact can initially make customers stay with the e-mail program for longer time. For example, one more direct mail per month would extend customer opt-out time for an average of 7.9 months. However, this positive effect diminishes and even reverses after the increase reaches a threshold. Table 7 shows that three more direct mails per month can extend the opt-out time for a shorter length of time than one or two more direct mails. Similarly, increasing the number of e-mails sent to customers also has an inverted U-shaped effect on extending the opt-out time. For example, 10 more e-mails sent per month can extend the opt-out time for an average of 3.6 months, while 25 more e-mails sent per month can only extend the opt-out time for an average of 2.6 months. These findings suggest that firms should focus not on increasing marketing intensity per se but rather on the content of the marketing messages they deliver to customers. Customers are willing to receive marketing materials that match their needs, including category fit and monetary incentives (e.g., Krishnamurthy 2001). Firms should customize their marketing efforts on the basis of customers' interests, preferences, and purchase histories. Undifferentiated mass marketing not only results in poor targeting but also generates negative feelings and leads to e-mail opt-out and customer churn.

Optimal Resource Allocation

The parameter estimates in Table 5 show that a firm's marketing contact policy influences both the length of time a customer stays in an e-mail program and the average amount a customer spends on a transaction while he or she is subscribing to the e-mail program. In this section, we address the following question: Under the current budget constraint, how can the firm optimally reallocate its budget to different customers and across different marketing channels (direct mail and e-mail) to maximize both customers' e-mail subscription time and sales revenue? We randomly selected four customers who are predicted to opt out between one and two years after they opt in and simulate the optimal marketing contact decisions. Our simulation is based on the assumption that customers will demonstrate a similar purchase frequency to that during their subscription to the e-mail program if the length of subscription time were extended. We argue that the assumption is reasonable because customers who subscribe to the e-mail program are exposed to persuasive e-mail messages aimed at retaining them and cross-selling and up-selling to them. Therefore, customers in such a relationship are likely to be more active purchasers than those who are not enrolled in the e-mail program.

The retailer estimates that one direct mail costs $.67 and one e-mail costs $.25.[4] Therefore, we calculate that the firm spent an average of $12.36 in total on the four customers in each month during the period they stayed in the e-mail program. Under the firm's current marketing contact policy, Customers 1, 2, 3, and 4 are expected to stay in the e-mail program for approximately 17, 17, 16, and 24 months, respectively, and then opt out of the e-mail program. During their e-mail subscription time, Customers 1, 2, 3, and 4 contribute an average of $25.35, $6.05, $34.28, and $12.77 in

[4]We acknowledge that the average unit cost of an e-mail is higher than the average industry standard. The cost is estimated by the retailer, which typically runs e-mail marketing campaigns. Each campaign may include the costs of initial content selection and design, marketing research, pretest operation, modification and redesign, mailing, and postevaluation.

profit every month, respectively (see Table 8). Here, although we acknowledge that the gross margin varies across product categories, we apply a constant 30% gross margin to calculate profit and the potential profit mentioned afterward with the consent from the retailer.

To find the optimal marketing contact decisions, we keep the current budget constraint of $12.36 per month unchanged and set the marketing contacts as changing cells to be optimized using Excel Solver. We calculate the expected length of e-mail subscription time and the average transaction amount using our proposed model (Equations 1–11) and the estimated parameters (Table 5). In the optimization process, we set the objective to maximize the total profit generated from the four customers. Under the current budget constraint, we find that the optimal contacting strategy is to increase the marketing contact for Customers 1 and 4 through both direct mail and e-mail but to decrease spending for Customers 2 and 3 (see Table 8). For example, the retailer should target Customers 1 and 2 with both one direct mail every other week and approximately one e-mail every other day or every week, respectively. The expected benefits include extending Customers 1 and 2's e-mail subscription time to the maximum of approximately 30 and 39 months and increasing their potential profit to $50.07 and $41.43 per month, respectively.

In summary, the two additional analyses discussed in this section indicate that marketing intensity has a significant influence on customers' opt-in and opt-out times. By strategically reallocating resources across different communication channels, firms can extend the length of time their e-mail subscribers stay with them and maximize customers' spending.

CONCLUSIONS AND FURTHER RESEARCH

The objective of this article is to explore the factors that are critical for managing an effective permission-based marketing program (e.g., e-mail marketing). To maximize return on investment, firms always desire to increase customers' opt-in rate and decrease their opt-out rate. To achieve these goals, marketers would benefit from understanding what makes customers willing to grant permission to firms, what triggers them to withdraw, and how to influence their decisions.

When customers have joined the permission marketing program, firms can send marketing messages to customers' e-mail inboxes or mobile devices to influence their purchase behavior. A customer's decision to stay in the marketing program is associated with his or her purchase behavior, and capturing such dependence helps predict when the customer is likely to opt out. This research proposes a trivariate copula model that can jointly model a customer's opt-in time, opt-out time, and average transaction amount. The empirical study reveals a positive dependence between the opt-in and the opt-out times and a positive dependence between the opt-out time and the average transaction amount conditional on the opt-in time. By capturing such dependence, the proposed model improves the predictive performance of the opt-out time over several benchmark models.

In addition, this research yields several important findings that have managerial relevance. We find that customers with certain characteristics are more likely to opt in or opt out. We find that customers under high marketing intensity are less likely to opt in. After customers have joined the marketing program, overmarketing can make them withdraw more quickly. Through a simulation study, we demonstrate how firms can optimally allocate resources to different channels such as direct mail and e-mail under the current budget constraint. Furthermore, we find that higher e-mail open rates lead to higher spending levels, suggesting that firms should focus on delivering marketing messages that are relevant to their e-mail subscribers.

To the best of our knowledge, this is the first study to model customer opt-in and opt-out times while incorporating purchase behavior. Nevertheless, there are some limitations we should address. First, because permission marketing pertains to customers' interests and needs, our findings may be constrained by the industry of analysis. For example, if we were to analyze a permission-based e-mail program in another industry, such as music, the people who are most likely to have an active interest in the category are probably those who are younger; in the low- to medium-income bracket; and heavy mobile and online users who are fans of social media, online shopping, and so on. Because the music industry is increasingly digital, traditional marketing channels such as direct mail may not be as influential as digital channels such as online search, social media, or e-mail marketing. Thus, to generalize our findings from this study, further research could apply our model to other product categories or industries.

Second, in this study, our main objective is neither to generalize under what conditions opt-in and opt-out times are positively or negatively correlated nor to offer explanations for such phenomena. Our main focus is to capture such dependence through an empirical model and provide a better prediction of opt-out time. While we recognize the

Table 8

COMPARISON OF CURRENT AND OPTIMAL CONTACTING STRATEGY

	Current Resource Allocation (per Month)				*Optimal Resource Allocation (per Month)*			
	Direct Mail	*E-Mail*	*Time Expected to Stay in the E-Mail Program (in Months)*	*Current Profit*	*Direct Mail*	*E-Mail*	*Time Expected to Stay in the E-Mail Program (in Months)*	*Potential Profit*
Customer 1	1	8	17	$25.35	2	13	30	$50.07
Customer 2	4	2	17	$6.05	2	4	39	$41.43
Customer 3	2	12	16	$34.28	1	6	40	$47.80
Customer 4	1	6	24	$12.77	2	7	35	$26.83

importance of identifying the causes of possible dependence, we leave this question for future study.

Third, in the optimization process, while we capture the effects of marketing contact on purchase amount through the log-normal model, we hold customers' purchase frequency and e-mail open and click-through rates constant. Nevertheless, customers may increase or decrease their purchase frequency and e-mail response rate if the marketing intensity changes. Further research could take these factors into consideration.

Fourth, to apply our proposed model to other permission-based contexts, such as mobile or social media, researchers should consider other factors. Mobile-based permission marketing may depend on factors such as the design of the mobile website, cell phone screen size and resolution, and the ease of mobile payment. Social networking–based permission marketing could depend on the number of "friends," consumers' profile and activeness, and privacy concerns. E-mail providers such as Google have redesigned the e-mail inbox to allow users to categorize their e-mails using tabs. For example, different tabs can be created to organize e-mails from different sources, such as "Work," "Social networks," and "Promotions." Such a feature could make customers less likely to open an e-mail from a less important tab such as "Promotions." It would be worthwhile for further research to study how such environmental factors and changes in interfaces affect the participation of customers in permission-based marketing programs.

APPENDIX: ESTIMATION OF THE PROPOSED PAIR-COPULA MODEL

We show the estimation of the trivariate copula model with Frank pair-copulas. We assume that there are three-dimensional data simulated using the algorithm (see the Web Appendix) with N observations. Let (T_{i1}, T_{i2}) denote the simulated paired opt-in and opt-out times. Because opt-in and opt-out times are sequentially observed events, if the observation window is of a fixed length C_i and the opt-in time is observed as T_{i1}, the maximum (censoring) lifetime that can be observed for opt-out time T_{i2} is $C_{i2}^{max} = C_i - T_{i1}$. Let $(t_{i1}, t_{i2}) = [\min(T_{i1}, C_i), \min(T_{i2}, C_i - T_{i1})]$ denote the observed opt-in and opt-out times. Let x_{i3} denote the simulated log-normally distributed variable that is only observed when the opt-in time is observed.

Using the two-step procedure, we first estimate the parameters of each marginal model (see Equations 1–5 in the main text) using maximum likelihood estimation. Second, we plug in these estimates to evaluate the pair-copula parameters by maximizing the log-likelihood function specified in Equation 11 in the main text. Following Aas et al. (2009), we estimate the parameters sequentially. First, we estimate the parameter Ω_{12} from the original data by maximizing the first two terms of Equation 11 given by

$$\text{(A1)} \quad \ell_1(\Omega_{12}) = \sum_{i=1}^{N}\Big(\delta_{i1}\delta_{i2}\log\big\{c_{12}[S_1(t_{i1}), S_2(t_{i2}); \Omega_{12}]\big\} + \delta_{i1}(1-\delta_{i2})\log\big[S(t_{i2}|t_{i1}); \Omega_{12}\big]\Big),$$

where the bivariate copula density $c_{12}(\cdot)$ is given by

$$\text{(A2)} \quad c_{12}[S_1(t_{i1}), S_2(t_{i2}); \Omega_{12}] = \frac{e^{-\Omega_{12}(u_{i1}+u_{i2})}\left(1-e^{-\Omega_{12}}\right)\Omega_{12}}{\left[e^{-\Omega_{12}} - 1 + \left(e^{-\Omega_{12}u_{i1}} - 1\right)\left(e^{-\Omega_{12}u_{i2}} - 1\right)\right]^2},$$

and the conditional survival function $S(t_{i2}|t_{i1})$ is given by

$$\text{(A3)} \quad S(t_{i2}|t_{i1}; \Omega_{12}) = \frac{e^{-\Omega_{12}u_{i1}}\left(e^{-\Omega_{12}u_{i2}} - 1\right)}{e^{-\Omega_{12}} - 1 + \left(e^{-\Omega_{12}u_{i1}} - 1\right)\left(e^{-\Omega_{12}u_{i2}} - 1\right)}.$$

We also estimate the parameter Ω_{13} from the original data by maximizing the third term of Equation 11 given by

$$\text{(A4)} \quad \ell_2(\Omega_{13}) = \sum_{i=1}^{N}\delta_{i1}\log\big\{c_{13}[S_1(t_{i1}), S_3(x_{i3}); \Omega_{13}]\big\},$$

where the bivariate copula density $c_{13}(\cdot)$ is given by

$$\text{(A5)} \quad c_{13}[S_1(t_{i1}), S_3(x_{i3}); \Omega_{13}] = \frac{e^{-\Omega_{13}(u_{i1}+u_{i3})}\left(1-e^{-\Omega_{13}}\right)\Omega_{13}}{\left[e^{-\Omega_{13}} - 1 + \left(e^{-\Omega_{13}u_{i1}} - 1\right)\left(e^{-\Omega_{13}u_{i3}} - 1\right)\right]^2}.$$

Here, we use u_{i1}, u_{i2}, and u_{i3} to denote the marginal survival functions $S_1(t_{i1})$, $S_2(t_{i2})$, and $S_3(x_{i3})$ for variables t_{i1}, t_{i2}, and x_{i3}.

Second, we estimate the parameter $\Omega_{23|1}$ by maximizing the fourth term of Equation 11:

$$\text{(A6)} \quad \ell_3(\Omega_{23|1}) = \sum_{i=1}^{N}\delta_{i1}\delta_{i2}\log\Big\{c_{23|1}\big[S(t_{i2}|t_{i1}), S(x_{i3}|t_{i1}); \Omega_{23|1}\big]\Big\},$$

where the bivariate copula density $c_{23|1}$ is given by

$$\text{(A7)} \quad c_{23|1}\big[S(t_{i2}|t_{i1}), S(x_{i3}|t_{i1}); \Omega_{23|1}\big] = \frac{e^{-\Omega_{23|1}(u_{i2|1}+u_{i3|1})}\left(1-e^{-\Omega_{23|1}}\right)\Omega_{23|1}}{\left[e^{-\Omega_{23|1}} - 1 + \left(e^{-\Omega_{23|1}u_{i2|1}} - 1\right)\left(e^{-\Omega_{23|1}u_{i3|1}} - 1\right)\right]^2},$$

where $u_{i2|1}$ and $u_{i3|1}$ denote the conditional survival functions $S(t_{i2}|t_{i1})$ and $S(x_{i3}|t_{i1})$, which are computed as

$$\text{(A8)} \quad S(t_{i2}|t_{i1}) = \frac{e^{-\Omega_{12}u_{i1}}\left(e^{-\Omega_{12}u_{i2}} - 1\right)}{e^{-\Omega_{12}} - 1 + \left(e^{-\Omega_{12}u_{i1}} - 1\right)\left(e^{-\Omega_{12}u_{i2}} - 1\right)},$$

and

$$\text{(A9)} \quad S(x_{i3}|t_{i1}) = \frac{e^{-\Omega_{13}u_{i1}}\left(e^{-\Omega_{13}u_{i3}} - 1\right)}{e^{-\Omega_{13}} - 1 + \left(e^{-\Omega_{13}u_{i1}} - 1\right)\left(e^{-\Omega_{13}u_{i3}} - 1\right)},$$

where Ω_{12} and Ω_{13} are estimated from the previous steps. Third, using the estimates of Ω_{12}, Ω_{13}, and $\Omega_{23|1}$ from the previous steps as the starting value, we maximize the full log-likelihood specified in Equation 11 to obtain the final estimates $\hat{\Omega}_{12}$, $\hat{\Omega}_{13}$, and $\hat{\Omega}_{23|1}$.

REFERENCES

Aas, Kjersti, Claudia Czado, Arnoldo Frigessi, and Henrik Bakken (2009), "Pair-Copula Constructions of Multiple Dependence," *Insurance: Mathematics and Economics*, 44 (2), 182–98.

Armstrong, Scott J., Vicki G. Morwitz, and V. Kumar (2000), "Sales Forecasts for Existing Consumer Products and Services: Do Purchase Intentions Contribute to Accuracy?" *International Journal of Forecasting*, 16 (3), 383–97.

Ascarza, Eva and Bruce G.S. Hardie (2013), "A Joint Model of Usage and Churn in Contractual Settings," *Marketing Science*, 32 (4), 570–90.

Barnes, Stuart J. and Eusebio Scornavacca (2008), "Uncovering Patterns in Mobile Advertising Opt-In Behaviour: A Decision Hierarchy Approach," *International Journal of Mobile Communications*, 6 (4), 405–416.

Bart, Yakov, Venkatesh Shankar, Fareena Sultan, and Glen L. Urban (2005), "Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large-Scale Exploratory Empirical Study," *Journal of Marketing*, 69 (October), 133–52.

Bedford, Tim and Roger M. Cooke (2002), "Vines—A New Graphical Model for Dependent Random Variables," *Annals of Statistics*, 30 (4), 1031–1068.

Bonfrer, André and Xavier Drèze (2009), "Real-Time Evaluation of Email Campaign Performance," *Marketing Science*, 28 (2), 251–63.

Borle, Sharad, Siddharth S. Singh, and Dipak C. Jain (2008), "Customer Lifetime Value Measurement," *Management Science*, 54 (1), 100–112.

Brey, Eric T., Siu-Ian So, Dae-Young Kim, and Alastair M. Morrison (2007), "Web-Based Permission Marketing: Segmentation for the Lodging Industry," *Tourism Management*, 28 (6), 1408–1416.

Chintagunta, Pradeep K. and Sudeep Haldar (1998), "Investigating Purchase Timing Behavior in Two Related Product Categories," *Journal of Marketing Research*, 35 (February), 43–53.

Chittenden, Lisa and Ruth Rettie (2003), "An Evaluation of E-Mail Marketing and Factors Affecting Response," *Journal of Targeting, Measurement and Analysis for Marketing*, 11 (3), 203–217.

Danaher, Peter J. (2002), "Optimal Pricing of New Subscription Services: Analysis of a Market Experiment," *Marketing Science*, 21 (2), 119–38.

——— and Michael S. Smith (2011), "Modeling Multivariate Distributions Using Copulas: Applications in Marketing," *Marketing Science*, 30 (1), 4–21.

Direct Marketing Association (2011), *The Power of Direct Marketing: ROI, Sales, Expenditures, and Employment in the US, 2011–2012 Edition*. New York: Direct Marketing Association.

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

Forrester Research (2011), *US Interactive Marketing Forecast, 2011 to 2016*. Cambridge, MA: Forrester Research.

Ganesan, Shankar (1994), "Determinants of Long-Term Orientation in Buyer–Seller Relationships," *Journal of Marketing*, 58 (April), 1–19.

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

Gutierrez, Roberto G. (2002), "Parametric Frailty and Shared Frailty Survival Models," *The Stata Journal*, 2 (1), 22–44.

Ha, Louisa (1996), "Observations: Advertising Clutter in Consumer Magazines: Dimensions and Effects," *Journal of Advertising Research*, 36 (4), 76–84.

Han, Aaron and Jerry A. Hausman (1990), "Flexible Parametric Estimation of Duration and Competing Risk Models," *Journal of Applied Econometrics*, 5 (1), 1–28.

He, Wenqing and Jerald F. Lawless (2003), "Flexible Maximum Likelihood Methods for Bivariate Proportional Hazard Models," *Biometrics*, 59 (4), 837–48.

Heilman, Carrie M., Kent Nakamoto, and Ambar G. Rao (2002), "Pleasant Surprises: Consumer Response to Unexpected In-Store Coupons," *Journal of Marketing Research*, 39 (May), 242–52.

Hobæk Haff, Ingrid (2013), "Parameter Estimation for Pair-Copula Constructions," *Bernoulli*, 19 (2), 462–91.

Jain, Dipak C. and Naufel J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10 (1), 1–23.

Jayawardhena, Chanaka, Andreas Kuckertz, Heikiki Karjaluoto, and Teemu Kautonen (2009), "Antecedents to Permission Based Mobile Marketing: An Initial Examination," *European Journal of Marketing*, 43 (3/4), 473–99.

Joe, Harry (1997), *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

Jolley, William, Alvin Lee, Richard Mizerski, and Saalem Sadeque (2013), "Permission Email Messages Significantly Increase Gambler Retention," *Journal of Business Research*, 66 (9), 1617–22.

Krishnamurthy, Sandeep (2001), "A Comprehensive Analysis of Permission Marketing," *Journal of Computer-Mediated Communication*, 6 (2), (accessed May 14, 2014), [available at http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2001.tb00119.x/abstract].

Kurowicka, Dorota and Harry Joe (2011), *Dependence Modeling: Vine Copula Handbook*. London: World Scientific Publishing Company.

Kushwaha, Tarun and Venkatesh Shankar (2013), "Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics," *Journal of Marketing*, 77 (July), 67–85.

Li, Shibo, Baohong Sun, and Alan L. Montgomery (2011), "Cross-Selling the Right Product to the Right Customer at the Right Time," *Journal of Marketing Research*, 48 (August), 683–700.

Marinova, Ana, Jamie Murphy, and Brian L. Massey (2002), "Permission E-Mail Marketing as a Means of Targeted Promotion," *Cornell Hotel and Restaurant Administration Quarterly*, 43 (1), 61–69.

Meade, Nigel and Towhidul Islam (2010), "Using Copulas to Model Repeat Purchase Behavior: An Exploratory Analysis via a Case Study," *European Journal of Operation Research*, 200 (3), 908–917.

Min, Aleksey and Claudia Czado (2010), "Bayesian Inference for Multivariate Copulas Using Pair-Copula Constructions," *Journal of Financial Econometrics*, 8 (4), 511–46.

Nash, Edward L. (1993), *Database Marketing: The Ultimate Marketing Tool*. New York: McGraw-Hill.

Nelsen, Roger B. (2006), *An Introduction to Copulas*, 2nd ed. New York: Springer.

Netzer, Oded, James M. Lattin, and V. Srinivasan (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*, 27 (2), 185–204.

Online Media Daily (2002), "GartnerG2 Says Email Threatens Traditional Direct Mail Promotions," (March 2), (accessed May 6, 2014), [available at http://www.mediapost.com/publications/article/7287/email-threatens-traditional-direct-mail-promotions.html#axzz2JP9JxAQ6].

Panagiotelis, Anastasios, Claudia Czado, and Harry Joe (2012), "Pair Copula Constructions for Multivariate Discrete Data," *Journal of American Statistical Association*, 107 (499), 1063–72.

———, Michael S. Smith, and Peter Danaher (2014), "From Amazon to Apple: Modeling Online Retail Sales, Purchase Incidence, and Visit Behavior," *Journal of Business & Economic Statistics*, 32 (1), 14–29.

Park, Young-Hoon and Peter S. Fader (2004), "Modeling Browsing Behavior at Multiple Websites," *Marketing Science*, 23 (3), 280–303.

Petersen, Andrew J. and V. Kumar (2009), "Are Product Returns a Necessary Evil? Antecedents and Consequences," *Journal of Marketing*, 73 (May), 35–51.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17–35.

Sahu, Sujit K., Dipak K. Dey, Helen Aslanidou, and Debajyoti Sinha (1997), "A Weibull Regression Model with Gamma Frailties for Multivariate Survival Data," *Lifetime Data Analysis*, 3 (2), 123–37.

Savu, Cornelia and Mark Trede (2010), "Hierarchies of Archimedean Copulas," *Quantitative Finance*, 10 (3), 295–304.

Schmittlein, David C. and Donald G. Morrison (1983), "Modeling and Estimation Using Job Duration Data," *Organizational Behavior and Human Performance*, 32 (1), 1–22.

Schubina, Maria and Mei-LingTing Lee (2004), "On Maximum Attainable Correlation and Other Measures of Dependence for the Sarmanov Family of Bivariate Distributions," *Communications in Statistics: Theory and Methods*, 33 (5), 1031–52.

Schweidel, David A., Peter S. Fader, and Eric T. Bradlow (2008), "A Bivariate Timing Model of Customer Acquisition and Retention," *Marketing Science*, 27 (5), 829–43.

Seetharaman, P.B. and Pradeep K. Chintagunta (2003), "The Proportional Hazard Model for the Purchase Timing: A Comparison of Alternative Specifications," *Journal of Business & Economic Statistics*, 21 (3), 368–82.

Shankar, Venkatesh and Sridhar Balasubramanian (2009), "Mobile Marketing: A Synthesis and Prognosis," *Journal of Interactive Marketing*, 23 (2), 118–29.

———, Alladi Venkatesh, Charles Hofacker, and Prasad Naik (2010), "Mobile Marketing in the Retailing Environment: Current Insights and Future Research Avenues," *Journal of Interactive Marketing*, 24 (2), 111–20.

Shih, Joanna H. and Thomas A. Louis (1995), "Inferences on the Association Parameter in Copula Models for Bivariate Survival Data," *Biometrics*, 51 (4), 1384–99.

Sinha, Rajiv K. and Murali Chandrashekaran (1992), "A Split Hazard Model for Analyzing the Diffusion of Innovations," *Journal of Marketing Research*, 29 (February), 116–27.

Sklar, A. (1959), "Fonctions de Répartition à n Dimensions et Leurs Marges," *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229–31.

Smith, Michael, Aleksey Min, Carlos Almeida, and Claudia Czado (2010), "Modeling Longitudinal Data Using a Pair-Copula Decomposition of Serial Dependence," *Journal of the American Statistical Association*, 105 (492), 1467–79.

Stephen, Andrew T. and Jeff Galak (2012), "The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace," *Journal of Marketing Research*, 49 (October), 624–39.

Tezinde, Tito, Brett Smith, and Jamie Murphy (2002), "Getting Permission: Exploring Factors Affecting Permission Marketing," *Journal of Interactive Marketing*, 16 (4), 28–39.

Trivedi, Pravin K. and David M. Zimmer (2005), "Copula Modeling: An Introduction for Practitioners," *Foundations and Trends in Econometrics*, 1 (1), 1–111.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (October), 106–125.

Völckner, Franziska (2008), "The Dual Role of Price: Decomposing Consumers' Reactions to Price," *Journal of the Academy of Marketing Science*, 36 (3), 359–77.

Winer, Russell S. (2009), "New Communications Approaches in Marketing: Issues and Research Directions," *Journal of Interactive Marketing*, 23 (2), 108–117.

Yeo, Keng Leong and Emiliano A. Valdez (2006), "Claim Dependence with Common Effects in Credibility Models," *Insurance: Mathematics and Economics*, 38 (3), 609–629.

Zimmer, David M. and Pravin K. Trivedi (2006), "Using Trivariate Copulas to Model Sample Selection and Treatment Effects," *Journal of Business & Economic Statistics*, 24 (1), 63–76.