The impact of the introduction and use of an informational website on offline customer buying behavior

J.E.M. van Nierop[1,2], P.S.H. Leeflang[2,3], M.L. Teerling[4], K.R.E. Huizingh[2]

February 2011

[1] Corresponding author: University of Groningen, Faculty of Economics and Business, P.O. Box 800, 9700 AV Groningen, The Netherlands, Tel: +31 50 363 6288, Fax: +31 50 363 3720, j.e.m.van.nierop@rug.nl. We thank Dennis Fok, Debra Trampe and Bernard Nijstad for useful comments on this paper, as well as participants at the Marketing Science conference 2009 in Ann Arbor, Michigan.

[2] University of Groningen, Faculty of Economics and Business

[3] Luiss Guido Carli University, Rome, Italy

[4] Telematica Institute, Enschede, the Netherlands

The impact of the introduction and use of an informational website on offline customer buying behavior

Do customers increase or decrease their spending in response to the introduction of an informational website? To answer this question, this study considers the effects of the introduction and use of an informational website by a large national retailer on offline customer buying behavior. More specifically, we study its effects on the number of shopping trips and the amount spent per category per shopping trip. The model is calibrated through estimation of a Poisson model (shopping trips) and a type-II tobit model (amount spent per category per shopping trip), with effect parameters that vary across customers. For the focal retailer, an informational website creates more bad than good news; most website visitors engage in fewer shopping trips and spend less in all product categories. The authors also compare the characteristics of shoppers who exhibit negative website effects with those few who show positive effects and thus derive key implications for research and practice.

Keywords: informational website; online and offline behavior, decomposition

1. Introduction

Virtually every company offers information through its website, and most of them enable customers to buy online as well. According to extant categorizations (e.g., Lee & Grewal, 2004; Teo & Pian, 2004), websites thus can be classified as either informational or transactional: An informational website offers commercial information but does not allow customers to make purchases online, whereas a transactional website does. Most academic research into the effectiveness of websites focuses on the impact of transactional websites (e.g., Moe & Fader, 2004; Sismeiro & Bucklin, 2004). Yet in reality, many firms maintain websites without transaction functionality, such as Zara; Bailey, Banks and Biddle; Dollar General; IKEA Netherlands, and the major electronics department store MediaMarkt in several European countries. Therefore, we address customer responses to the introduction and use of an informational website that supports an existing bricks-and-mortar retailer that does not have a transactional website.

When online information search leads to a purchase decision, customers can conduct that purchase in various ways. First, they might buy the product from the website on which they found the information, assuming the site allows them to do so. Such behavior is still relatively exceptional; online conversion rates rarely exceed 5% and are often much less (Moe & Fader, 2004). Second, customers may decide to leave the site and buy the product at a competitor's website (i.e., "free riding," Huang, Lurie, & Mitra, 2009). Third, online search might precede offline purchase, an option that is especially likely for search products (e.g., Alba & Lynch, 1997; Weathers, Sharma, & Wood, 2007) and for customers with technology anxiety or trust issues (e.g. Hoffman, Novak, & Peralta, 1999; Roy & Ghose, 2006). A majority of consumers still prefer to purchase in physical stores (i.e., 67%, Accenture, 2007) and use the Internet simply to gather information about product features and prices. That is, they prefer to research online and buy offline (Krillion, 2008; Mendelsohn, Johnson, & Meyer, 2006).

Several studies investigate the effects of new Internet channels on either aggregate firm performance or individual customer behavior (for an overview, see Neslin & Shankar, 2009). Various studies indicate that at the aggregate level, the Internet channel rarely cannibalizes existing channels (e.g., Deleersnyder, Geyskens, Gielens & Dekimpe, 2002; Geyskens, Gielens, & Dekimpe, 2002), and the effects of an informational channel on performance may be positive (Lee & Grewal, 2004). At the individual customer level, research thus far has been able to determine only the effects of online transactional channels, which have emerged as both positive and negative (Ansari, Mela, & Neslin, 2008; Kushwaha & Shankar, 2007).

But what happens to individual purchase behavior if an organization starts to provide online information to customers? Visiting an informational website seemingly should induce customers to make more store visits and spend more money. However, we posit that the effect could be negative if the online information makes those customers (1) more efficient buyers, who make fewer shopping trips and fewer impulse purchases in the store and/or (2) more critical buyers, who use the information but buy from competitors or consider the information provided insufficient. Gensler, Dekimpe, and Skiera (2007) and Pauwels, Leeflang, Teerling, and Huizingh (2011) also demonstrate that the effects of an additional transactional channel depend on the product category; certain product categories are more suitable for online buying than others. Yet no empirical results at the individual level reveal whether customers increase or decrease their spending *across product categories over time* as a result of their use of an informational website.

In this study, we determine the effects of the introduction and usage of an informational website on purchases at the individual customer level. We relate the online search behavior of individual customers of a large retailer to their offline buying behavior using customer panel data that measure how often and how extensively customers visit the website over time, as well as how much and how often they shop at offline stores, and how much they spend in six different categories. We also study how offline buying behavior changes with the use of the website, using purchase data available for the periods before and after the implementation of the website.

We decompose the amount of money spent by a customer in a certain time period into the number of shopping trips in that period and the amount spent per category per trip to answer the following questions:

- Does the use of an informational website change the number of offline shopping trips conducted by individual customers?
- Does the use of an informational website alter the purchase amounts of individual customers in different product categories?

Our results show that the majority of registered website users engage in fewer shopping trips and spend less in all six product categories. Some consumers exhibit some positive behavioral effects, but overall, the effect is negative.

In the remainder of this article, we begin by reviewing prior literature. After discussing the methodology, we describe the data. Then we present the findings and discuss them in light of previous studies. We end with a summary of the main conclusions and their implications.

2. Literature review

2.1. The impact of the introduction of a transactional website

Most studies that consider the impact of an added transactional website on firm performance have investigated the effects at either the aggregate (firm) sales level or the individual customer level. The majority focus on the aggregate level. For example, Biyalogorsky and Naik (2003) investigate Tower Records' sales figures during 1989–1999 to determine the extent to which its added online transactional channel cannibalized offline sales. They find a cannibalization rate of 2.8% from online sales, indicating negligible contemporaneous cannibalization. Coelho, Easingwood, and Coelho (2003) demonstrate that when a company starts using a new channel, it can expect stronger sales growth from this channel than from its traditional channel, likely because the firm is reaching new customer segments. However, Coelho et al. (2003) also indicate that as penetration into these segments increases, growth diminishes, and cannibalization might begin between channels.

In the newspaper industry, Deleersnyder et al. (2002) find hardly any cannibalization between online and offline channels, possibly because different market segments prefer either a hard copy or information on the Internet. They also show that depending on the positioning of the channel portfolio, cannibalization or synergy between the channels is possible. Geyskens et al. (2002), also focusing on the newspaper industry, conclude that firms with fewer direct channels can gain more from using the Internet as an additional channel than can firms with a broader direct marketing offering. The newspaper industry can easily take advantage of the special economics of information goods delivered over the Internet, but these results may differ for retailers (Bakos & Brynjolfsson, 2000). In a retail setting, Lee and Grewal (2004) show that adding the Internet as a transactional channel does not have an effect on Tobin's Q. Wu, Mahajan, and Balasubramanian (2003), in a study of companies from a broad range of industries, find no performance effect of the addition of the Internet as a transactional channel. In another study, using data from an online grocery store (Peapod), Wu and Rangaswamy (2003) demonstrate that website features can either decrease or increase the amount of search and thereby influence consumers' consideration sets.

From the few studies based on individual customer data, we can infer that (1) marketing efforts can move customers into a particular channel (Ansari et al. 2008), (2) most customers use multiple channels after the addition of an Internet channel (Dholakia, Zhao, & Dholakia, 2005; Gensler et al. 2007), and (3) adding a transactional Internet channel may either decrease (Ansari et al. 2008; Gensler et al. 2007) or increase (Kushwaha & Shankar, 2007) customer buying behavior. Marketing efforts not only influence customer channel choice but also may explain the increase in buying by multichannel customers (Neslin et al., 2006). This argument could hold for informational websites as well, because customers who use both channels gain exposure to more marketing efforts and brands than do those using a single channel. Ansari et al. (2008) provide empirical support for this effect, and Wallace, Giese, and Johnson (2004) show that retailers may receive a loyalty payoff because customers perceive an enhanced portfolio of service outputs provided by multiple channels.

2.2. The effect of the introduction and usage of informational websites

Only a few studies consider the effects of introducing an informational website on offline customer behavior. Viswanathan, Kuruzovich, Gosain, and Agarwal (2007) study consumers' use of online infomediaries and subsequent purchase of cars offline, using an extensive secondary data set gathered from a survey of new automobile purchasers. They find clear differences across consumer segments: Those who obtain online price information pay lower prices for the same car than do consumers who obtain online product information.

Firms may invest in informational websites to obtain positive effects in terms of, for example, consumer knowledge, brand perceptions, or buying behavior, but they also run the risk of negative effects on switching behavior and search time. Online information search also can have multiple effects. Hoque and Lohse (1999) show that consumers may make different or more informed decisions in accordance with information they find online. Not only do consumers who use the Internet gain quality and efficiency improvements for their decision making (e.g., Alba & Lynch, 1997; Mick & Fournier, 1998), but firms may benefit from this effect as well, because consumers might search less for information offline. Ratchford, Lee, and Talukdar (2003) test this claim, using cross-sectional survey data for the automotive industry, and find that consumers gain efficiency, increased information, and bargaining power from an Internet channel, while car dealers save on the costs of salespeople's time.

Almost all studies use cross-sectional data from surveys. In contrast, we follow individual customers over time and observe how the introduction and use of the informational website affects their actual shopping and spending behavior in multiple categories. We thus are able to combine data regarding actual search behavior in one channel with actual purchase behavior in another for a specific company.

3. Empirical setting

We collected data from individual customers of a large, well-known, national retail department store in the Netherlands. The retailer introduced an informational website during the observation period. This website did not provide any transactional capabilities or link to any transactional website. The department store has 58 outlets in all major urban areas in the country, and each outlet contains a broad range of product categories, including clothing, sports, furniture, and so on.

3.1. Informational website

The informational website is a theme-oriented site that supports offline activities to increase the likelihood of store purchase. It provides customers with information about lifestyle issues related to the various product categories of the store, specific products offered in the stores, promotions, and the organization itself. Our data encompass six product categories: ladies' and men's fashions, children's products, accessories, living (interior design) products, and sports.

The data collection used a multimethod approach that involved the marketing activities of the focal and competitive companies, as well as observational data obtained with a longitudinal, quasi-experimental design (Cook & Campbell, 1979). The data refer to both website behavior *and* purchase behavior of customers of the focal firm. This firm participates in a national joint loyalty program of 21 partner firms in the Netherlands. Customers collect credits by purchasing from these different firms, which range from retail stores to banks to gasoline stations. In turn, they may exchange these credits to receive discounts on products sold by the member firms or theatre or airline tickets. This popular program was established in the early 1990s, and at the time of the data collection, more than half of all Dutch households were members. Only members of the loyalty program could use the website; they also had to register to be able to access it, using their loyalty card number. About 83% of the users registered immediately after the website's introduction, and the remaining users did so later.

We use observations from a panel of 8,615 customers, including both those who never used the website and frequent site visitors. Because the customers on the panel present their loyalty cards when they pay for their in-store purchases, we can obtain data about the offline buying behavior of the same customers. The data are available for 25 months (January 2000–May 2002), which span 14 months before the website introduction to 11 months after and refer to shopping trips in which customers purchased at least one product. In our subsequent analysis, we only include customers who made at least two purchases before and two purchases after the introduction of the website.

We aggregate purchase behavior to a monthly level because of the infrequency of department store buying behavior by individual customers. The total number of non-zero buying observations in the data set is 118,537. Because the time span of the panel is more than two years, we recognize the potential for panel attrition, which might limit the generalizability of our results. However, a very high percentage (92%) of the store's customers remained on the panel for the full study period. This low attrition rate seems reasonable, considering the type of store (department store with a large assortment), few competitors (department stores), the high number of participating firms in the loyalty program (21), and the popularity of this joint loyalty card program.

Through our quasi-experimental design (Campbell & Stanley, 1963), we collected pretest and posttest measures from the same panel of consumers. The experiment therefore featured a natural setting, such that real customers used a real website and made real store visits, during which they spent their own money. The assignment of consumers to groups is not random, as would be the case for laboratory experiments. To test our expectation that the registered website users may have different characteristics than non-website users, we computed socio-demographic statistics for both groups. On average, registered website users are slightly younger and more educated, and significantly more men appear in the website user group than in the non-site user group. These findings match the general Internet population at the time (2000–2002) of more educated, younger, male customers (e.g., Burke, 2002). We include these specific socio-demographics as control variables in the model.

In Figure 1, we illustrate the development of the average monthly number of store visits and the average amount spent per shopping trips of registered website users and non-users over time. The introduction of the website occurred in period 15. After the website introduction, users of the site visit the store less and spend slightly less than do non-users.

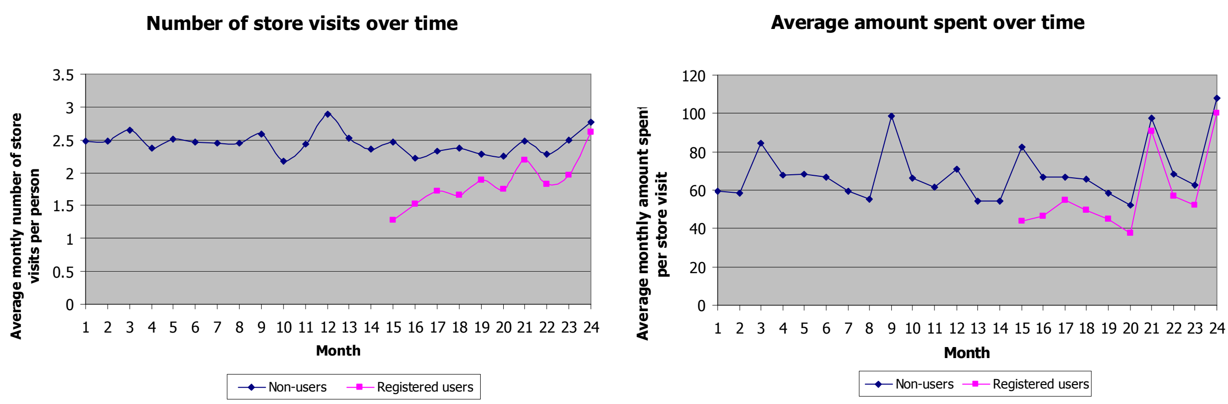

Fig. 1 Comparison of site users and non-site users over time for shopping trips and amount spent.
Note: Site introduction was in month 15.[5]

We display descriptive statistics for the data set in Table 1. When comparing the pre-introduction behavior of non-users with the pre-activity behavior of (eventually) registered users, we find that the number of shopping trips per period differs significantly, but the effect size is rather small (2.55 versus 2.46). We find significant differences in the amount spent per shopping trip between

[5] It is possible to create a similar graph using a static classification, in which users are labeled as such throughout the data set, instead of only after they become active. This approach would enable us to plot store visits and spending for both users and non-users *before* the introduction of the website as well. From this (unreported) graph, we determine that the pre-introduction patterns of customers who become active users at some point in time and of non-users are very similar. Therefore, no selection effect is evident; users and non-users exhibit comparable behavior before the website introduction.

groups for the ladies' fashion, children, accessories, and sports categories (superscript a). For both users and non-users, the average number of shopping trips per month drops in the period after site implementation, though for registered users, the drop is substantially larger: 0.59 for site users versus 0.11 fewer visits per month for non-site users (superscripts a and b). A difference-in-differences test shows that the cross-difference [(2.55 – 2.43) – (2.46 – 1.87) =] -0.48 is significant.[6] For registered users, we also observe a decrease in the average amount of money spent per shopping trip for all six categories (superscript b). For non-users, we observe no significant differences in spending in any categories. In subsequent analyses, we also investigate whether website use or other variables might cause these observed patterns.

[6] For this test, which is executed using a regression of the number of shopping trips on the two group variables (users and before/after) and their interaction, we use the static classification described in the previous footnote. The dynamic classification would result in a perfect correlation between the user-variable and the interaction. As part of our robustness analysis, we checked for differences between the static and dynamic classification and produced the same table using the static classification, as described in footnote 5. The conclusions are very similar. The most striking difference is the number of pages online. Instead of 6.32, as reported in Table 1, this number decreases to 3.61 when using the static classification. The number in Table 1 is much higher because it includes only the periods when users are active.

Table 1.
Descriptive statistics (average and standard deviation across months), N = 118,537

Non-users	**Before website introduction**		**After website introduction**	
	Mean	St. dev.	Mean	St. dev.
Number of shopping trips	2.55[ab]	1.99	2.43[ab]	1.99
Amount spent (€) on				
Ladies' fashion	15.32[a]	30.39	16.15[a]	44.28
Men's fashion	6.84	20.93	6.27	20.77
Children's products	9.14[a]	23.66	9.84[a]	32.28
Accessories	9.23[a]	19.79	9.01[a]	20.23
Living	9.68	41.29	10.02[a]	47.92
Sports	6.23[a]	23.22	5.64[a]	23.78

Registered users	**Before becoming active on the website**		**After becoming active on the website**	
	Mean	St. dev.	Mean	St. dev.
Number of shopping trips	2.46[ab]	1.91	1.87[ab]	1.90
Amount spent (€) on				
Ladies' fashion	13.32[ab]	28.90	11.92[ab]	32.53
Men's fashion	6.74[b]	21.84	5.93[b]	26.10
Children's products	7.22[ab]	21.30	5.92[ab]	20.82
Accessories	8.02[ab]	19.66	7.16[ab]	22.49
Living	9.51[b]	48.79	7.78[ab]	45.39
Sports	5.37[ab]	20.06	4.60[ab]	25.09
Number of pages online			6.32	15.35

[a]Significant difference (95% confidence) between registered users and non-users (bottom versus top panel), based on a two-sided, independent sample t-test.
[b]Significant difference (95% confidence) between before and after site introduction (left versus right panel), based on a two-sided, independent sample t-test.

3.2. Explanatory variables

We determine the effect of the introduction and use of the informational website on offline spending through a numerical specification of models that contain multiple explanatory variables, such as the number of website visits, offline promotional activities by the focal and competitive department stores, individual customers characteristics, and a time trend. We provide a detailed description of the variables in Table 2.

Table 2.
Variables available and used in the models.

Variable	Number	Variable description	Notation [a]
Online behavior variables			
Website visits	1	Overall number of website visits by individual *i* in period *t*.	W_{it}
Website category visits	2	The number of website pages visited by individual *i* in period *t* in category *c*.	W_{itc}
Website user control variable			
Dummy after introduction, non-user	3	This dummy variable equals 1 after introduction of the website for non-users, and 0 otherwise. More formally: $D_{it} = 1$ if $t \geq 15$ and individual *i* is never a website user in our data set, 0 otherwise. This variable is included to account for changes in the behavior of non-users after the site introduction. We thus can control for changes in the marketplace or store environment that may influence buying behavior.	D_{it}
Own promotions			
Time-specific dummy, for three major offline promotional activities	4	During the holiday shopping season in November and December.	P_{1t}
	5	A general promotion discount for all categories in the store.	P_{2t}
	6	A non-price promotion for fashion categories.	P_{3t}
Competitive promotions			
Time-specific dummies, for four major offline competitive promotional activities by the department store's two main competitors		*Competitor 1*:	
	7	Introduced a webstore in 2000.	CP_{1t}
	8	Used major tv advertisements to announce a new loyalty program in 2001.	CP_{2t}
	9	Used advertisements with extensive promotions late 2001.	CP_{3t}
		Competitor 2:	
	10	Introduced of a new door-to-door magazine distributed to every household early 2000.	CP_{4t}
Other			
Trend variable	11	The (log of the) number of months since site introduction, to control for trends in consumer behavior.	$ln(t)$
Individual-specific customer characteristics	12	The distance the customer must travel to the nearest outlet of the department store.	$Distance_i$
	13	Age.	Age_i
	14	Gender (0 = male, 1 = female).	$Gender_i$
	15	Education dummy equal to 1 if the customer has a college education or higher.	$HighEduc_i$

[a] Subscripts indicate whether the variable is specific to individual *i*, time *t*, and/or category *c*.

3.3. Estimation and holdout sample

To keep the estimation of the model manageable, we draw a random sample of 436 customers (5% of the 8,615 customers in the data set) responsible for 5,685 purchase observations. This sample contains both registered users (209) and non-users (227), all of whom were active in the store or on the website for at least two periods before and two periods after the website introduction. We split the data randomly into an estimation sample of 4,572 observations and a holdout sample of 1,113 (20%) observations.

4. Methodology

4.1. Model specification

We are interested in determining the effects of the introduction and use of an informational website on components of shopping behavior and therefore apply a decomposition. We decompose the total amount of money spent by individual i during month t (M_{it}) in shopping trips (V_{it}) in which he or she purchases at least one product and the total amount of money spent (M_{itc}) in all categories, indexed with c. Specifically:

$$M_{it} = V_{it} * \sum_{c} \frac{M_{itc}}{V_{it}} \qquad , i = 1, ..., I; c = 1, ..., C; t = 1, ..., T_i, \tag{1}$$

where

I = number of individual customers,

C = number of product categories, and

T_i = number of months in which individual i makes a shopping trip to the store.

We focus our analyses at the product category level instead of overall monetary value to determine (1) category-specific informational website and (2) category-specific webpage effects.

4.1.1. Modeling the number of visits to the store.

The number of store visits V_{it}, constitutes count data. For this type of data, the Poisson regression model is widely used (see Wooldridge, 1997) and appropriate. The probability that individual i at period t engages in v_{it} shopping trips is

$$\Pr(V_{it} = v_{it}) = \frac{e^{-\lambda_{it}} \lambda_{it}^{v_{it}}}{v_{it}!}, \qquad (2)$$

where λ_{it} reflects the expected number of shopping trips for individual i in month t, as explained by the regressors contained in Z_{it}:

$$\ln \lambda_{it} = \theta Z_{it}. \qquad (3)$$

The vector Z_{it} contains the non-category specific explanatory variables of Table 2—that is, all variables except for W_{itc}, the category-specific number of web pages visited. This latter variable appears in the total amount of money spent (M_{itc}) component of the model. The parameter vector θ describes the effects of these explanatory variables.

4.1.2. Modeling the amount spent.

The second component of Equation (1) is $\frac{M_{itc}}{V_{it}}$. For brevity, we call this variable Y_{itc}. It is a truncated variable, because we never observe values below 0.[7] Regular regression is not appropriate; it would lead to biased estimates. Instead, we use the type-II tobit model (Amemiya, 1985; Bucklin & Sismeiro, 2003; Chib, 1992; Fox, Montgomery, & Lodish, 2004). Because we deal with multiple product categories that may correlate, we adopt the multivariate type-II tobit, without correlation between the two stages, which is also known as a two-part model (for a similar approach, see Fox et al., 2004). The multivariate type-II tobit model consists of two stages. In the first stage, the model explains whether a customer buys in a particular category (purchase incidence), which is essentially a multivariate probit model (MVP; Chib & Greenberg 1998). In the second stage, it models the actual amount of money spent, given that the customer buys. This part of the model amounts to a truncated regression formulation (Franses & Paap,

[7] We observe a substantial number of zeros in the six category-spending variables. With 4,572 in-sample observations, we have $4{,}572 \times 6$ spending values, but they are positive in only 26% of the cases.

2001). For the two stages, we use the same symbols for dependent ($Y_{itc}^{(\cdot)}$) and independent ($X_{it}^{(\cdot)}$ and W_{itc}) variables and indicate the pertinent stage with superscripts (1 or 2).

For each individual *i*, we denote *purchase incidence* in category *c* in month *t* as $Y_{itc}^{(1)}$. This variable equals 1 if the customer buys in category *c* and 0 if not. The MVP part of the type-II tobit model has the following structure:

$$\begin{aligned} Y_{itc}^{(1)} &= 1 \qquad \text{if } Y_{itc}^{*(1)} > 0 \\ Y_{itc}^{(1)} &= 0 \qquad \text{if } Y_{itc}^{*(1)} \leq 0 \end{aligned} \tag{4}$$

where

$$Y_{itc}^{*(1)} = \alpha_c^{(1)} X_{it}^{(1)} + \alpha_{ic}^{(1)} W_{itc} + \gamma_{ic}^{(1)} + \varepsilon_{itc}^{(1)}. \tag{5}$$

Here $Y_{itc}^{*(1)}$ is the latent utility for individual *i* of buying from category *c* in period *t*. If the utility $Y_{itc}^{*(1)}$ is greater than 0, the individual purchases. The vector $X_{it}^{(1)}$ contains all variables specified in Table 2, with the exception of the online behavior variables W_{it} and W_{itc}. The parameter $\alpha_c^{(1)}$ describes the average effect of these variables; the parameter $\alpha_{ic}^{(1)}$ refers to the individual-specific effect of the number of website visits, W_{itc}.[8] Finally, $\gamma_{ic}^{(1)}$ is the unobserved random intercept, and the error term $\varepsilon_{itc}^{(1)}$ follows a multivariate normal distribution, with mean 0 and covariance matrix $\Sigma^{(1)}$.

In the second part of the type-II tobit model, *the amount spent* by individual *i* in category *c* in month *t*, $Y_{itc}^{(2)}$, is modeled as

$$Y_{itc}^{(2)} = \begin{cases} Y_{itc}^{*(2)} & \text{if } Y_{itc}^{(1)} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

---

[8] In the online appendix, where the solution to the endogeneity (see section 4.2) is provided, we label this variable $W_{itc}^{(2)}$, because the number of website visits is modeled explicitly in a second, parallel tobit model, which also has two stages.

where

$$Y_{itc}^{*(2)} = \alpha_c^{(2)} X_{it}^{(2)} + \alpha_{ic}^{(2)} W_{itc} + \gamma_{ic}^{(2)} + \varepsilon_{itc}^{(2)} . \tag{7}$$

Although not strictly necessary in the type-II tobit model, the vector $X_{it}^{(2)}$ contains exactly the same variables as $X_{it}^{(1)}$. The parameter $\alpha_c^{(2)}$ describes the average effects of these explanatory variables, the parameter $\alpha_{ic}^{(2)}$ refers to the individual-specific effects of the number of website visits, and $\gamma_{ic}^{(2)}$ is the unobserved random intercept. The error term $\varepsilon_{itc}^{(2)}$ follows a multivariate normal distribution, with mean 0 and covariance matrix $\Sigma^{(2)}$.

Throughout, we concede that the effects of the explanatory variables on the decision to spend and spending levels may differ (i.e., parameters estimated for $X_{itc}^{(1)}$ and $X_{itc}^{(2)}$ can differ). So for example, promotional activities may have an insignificant effect on purchase incidence ($Y_{itc}^{(1)}$) but significantly influence the amount spent ($Y_{itc}^{(2)}$). The unobserved individual-specific heterogeneity parameters $\alpha_{ic}^{(\cdot)}$ and $\gamma_{ic}^{(\cdot)}$ in both stages are draws from a multivariate normal distribution with mean 0 and variance matrix $Q^{(\cdot)}$ (e.g., Allenby & Rossi, 1999).

The multivariate error distributions for $\varepsilon_{itc}^{(1)}$ and $\varepsilon_{itc}^{(2)}$ allow for information from one category to influence the conditional predictions of other categories. We expect contemporaneous correlations of the disturbances across product categories, because excess expenditures in one category may result in either less spending in other categories (substitution) or complementary sales.

4.2. Endogeneity

Consumers who spend more in the store also might end up spending more time on the website. If we ignore this endogeneity, we might overestimate the effect of the website on the number of visits.[9] In technical terms, endogeneity means that the number of website visits could be correlated with the unobserved, individual-specific component driving store visits (Poisson model) or spending (type-II tobit model). In the online appendix, we describe how we deal with and correct for this potential endogeneity. For the Poisson model, it amounts to an instrumental variable-type approach, and for the type-II tobit model, we use a Bayesian approach, such that we use a second equation for each stage to model the number of website visits, then link these two equations with unobserved heterogeneity.

4.3. Estimation

We estimate a model that explains the number of shopping trips using maximum likelihood (Greene, 2002). The category-specific multivariate type-II tobit model for money spent is estimated for the six product categories simultaneously using the Markov chain Monte Carlo (MCMC) methodology. Similar to Fox et al. (2004), we obtain posterior results by implementing the Gibbs sampling technique (Geman & Geman, 1984) with data augmentation (Tanner & Wong, 1987).[10] To ensure parameter convergence, we use five different starting points for the MCMC algorithm and compute the potential scale reduction, as defined by Gelman, Carlin, Stern, and Rubin (1995). The closer this value is to 1, the better the convergence. After running the chains for a sufficient number of iterations, we achieved satisfactory convergence, with scale reduction values approaching 1. Because of its greater number of parameters, the model that takes endogeneity into account appeared to converge rather slowly, such that we had to use more burn-in iterations. We use 20,000 draws for the burn-in and 10,000 as final draws. We use every 10th of the final draws for inferences, because thinning the series of draws reduces autocorrelation and storage capacity demand (Gelman et al. 1995).

---

[9] This description refers to the Poisson model for the number of store visits, but it is analogous for the type-II tobit model for spending.
[10] The conditional posterior distributions are available on request.

To detect possible cross-category correlations and still ensure identification, we applied minimal restrictions to the covariance matrices. For identification purposes, the diagonal elements of $\Sigma^{(1)}$ must be set to 1 (Manchanda, Ansari, & Gupta, 1999). This correlation matrix provides information about cross-category effects in purchase incidence, such that a high positive correlation indicates that purchases in two product categories usually coincide. The covariance matrix $\Sigma^{(2)}$ is set to the identity matrix, and those of the unobserved heterogeneity $Q^{(1)}$ and $Q^{(2)}$ are set to diagonal matrices. For the parallel model that accounts for endogeneity (see the online appendix), we use the same settings as applied to the associated main equation of interest.

5. Findings

5.1. Shopping trips

The estimation results of the parameter vector θ of the Poisson model (see Equation (3)), which explains the number of shopping trips, appear in Table 3. We find a negative effect of the number of website visits (W_{it}) on the number of store visits ($\theta_1 = -0.302$, $p < 0.001$); for an individual customer, an informational website is expected to cause a reduction in offline store visits. Combined with the after introduction dummy for non-users, this result is striking. The dummy indicates that after correcting for possible other variables, after the introduction of the website, non-users visit the store more than do the registered users ($\theta_3 = 0.178$, $p<0.001$). This effect is also reflected in the number of shopping trips in Table 1 if we compare registered users with non-users.

Whereas the number of store visits is positively influenced by the holiday shopping season ($\theta_4 = 0.164$, $p = 0.002$) and the general promotion ($\theta_5 = 0.106$, $p = 0.035$), the fashion promotion does not have a significant effect on the number of shopping trips ($p = 0.180$).[11] The effects of the first two promotional activities by competitor 1 are negative ($\theta_7 = -0.177$, $p = 0.015$ and $\theta_8 = -0.241$, $p = 0.006$). The negative coefficient ($\theta_{11} = -0.137$, $p = 0.003$) for the log of the time since introduction illustrates a trend effect that can be explained partly by a macro-economic decline.[12] The distance to the closest store has a negative influence on the number of shopping trips ($\theta_{12} = -0.018$, $p < 0.001$), indicating that the farther customers live, the fewer store visits they make. Older people visit the store more frequently than do younger people, though this effect is barely significant ($\theta_{13} = 0.003$, $p = 0.067$). We find no difference in the number of shopping trips between men and women ($\theta_{14} = -0.023$, $p = 0.470$). Finally, we find a positive relation between people with at least a college education and the number of shopping trips ($\theta_{15} = 0.057$, $p = 0.047$).

As a robustness check, we run a gamma regression on the count data. This model overcomes a limitation of the Poisson model, which assumes the mean and variance are the same, because the gamma regression uses two separate parameters. We find the same negative effects of website visits. The other parameters change slightly in size, but the directions persist.

[11] We run a gamma regression to check if this result holds across different specifications and find non-significant effects of the fashion promotion.
[12] We also test a model that uses year dummies. We obtain a significant negative effect of the dummy for the first year and an insignificant effect for the second year.

Table 3

Poisson parameter estimates for the number of store visits (N = 4,572) (equation 3).

Number[a]	Variable	Coefficient θ	*p*-value[c]
	Intercept	1.155	<0.001
1	Number of website visits	-0.302	<0.001
3	Dummy after introduction, non-user[b]	0.178	<0.001
	Own Promotions		
4	Holiday shopping season dummy	0.164	0.002
5	General promotion dummy	0.106	0.035
6	Fashion promotion dummy	0.079	0.180
	Competitive Promotions		
7	Competitor 1: starts webstore	-0.177	0.015
8	Competitor 1: major TV advertisement	-0.241	0.006
9	Competitor 1 advertisements in 2001	0.002	0.948
10	Competitor 2: introduction door-to-door magazine	0.002	0.971
	Other		
11	(Log) months since introduction	-0.137	0.003
12	Distance to closest store in miles	-0.018	<0.001
13	Age in years	0.003	0.067
14	Gender (0 = male, 1 = female)	-0.023	0.470
15	Higher education (0 = no, 1 = yes)	0.057	0.047

[a] This number corresponds to the numbers in the second column in Table 2. Because the Poisson model is not category-specific, variable number 2, the category-specific number of web pages (W_{itc}), is missing in this model.

[b] This variable equals 0 before the introduction of the informational website; after introduction, it equals 1 for non-users and 0 for registered users.

[c] This column shows *p*-values for two-sided t-tests.

5.2. Amount spent per trip per category

Most customer characteristics have no significant effects on the purchase incidence component of the simultaneously estimated multivariate type-II tobit model. Pooling tests reveal that we should opt for models with unique parameters for all categories. Therefore, to determine the final model, we choose the model that offers the best predictions for the holdout sample.

We estimate the model that accounts for endogeneity (see Section 4.2 and the online appendix). In the type-II tobit model, we find no evidence for endogeneity. Apparently, there is no correlation between the error term and the number of website visits in the main equations (equation (5) and (7)), which is what would happen if unobserved factors ended up in the error term. Because we do not find evidence of endogeneity, we discuss the results for the type-II tobit model that does not take endogeneity into account here.

In Table 4, we provide the estimation results for the parameters $\alpha_c^{(1)}$ (purchase incidence, equation (5)) and $\alpha_c^{(2)}$ (amount spent, equation (7)). The "Purchase Incidence" columns contain parameters that specify the effects of the variables that determine whether someone purchases in a particular category. The "Amount" columns indicate the effect of each variable on the amount of money spent per shopping trip in a specific category, given that an individual customer buys from this category. We evaluate the performance of the Purchase Incidence stage of the type-II tobit model by computing a hit rate that indicates the percentage of observations for which we can predict the correct value. In the estimation sample, we achieve a hit rate of 72%; for the holdout sample, this value equals 66%. To evaluate the fit of the full tobit model, the common practice is to use McKelvey and Zavoina's (1975) pseudo R^2 measure (Veall & Zimmermann, 1994). Across categories, our model achieves an average out-of-sample R^2 value of 0.77. We obtain the lowest and highest values for accessories (0.68) and sports (0.92), respectively.

Table 4.

Parameter estimates for the multivariate type-II tobit model: Purchase incidence (equation 5) and amount spent (equation 7) (N = 4,572)

		Ladies' Fashion		**Men's Fashion**		**Children**		**Accessories**		**Living**		**Sports**	
Nr[a]	**Variable**	Purch. Inc.	Amount	Purch. Inc.	Amount	Purch. Inc.	Amount	Purch. Inc.	Amount	Purch. Inc.	Amount	Purch. Inc.	Amount
	intercept	**-0.241** *	**-0.238** **	**-0.973** ***	**-0.917** ***	**-0.614** ***	**-0.635** ***	-0.178	**-0.168** *	**-0.565** ***	**-0.493** ***	**-0.670** ***	**-0.624** ***
2	number of website pages visited in category-related website section	**-0.441** ***	**-0.105** **	**-0.550** ***	-0.078	-0.238	-0.062	**-0.329** ***	**-0.072** *	**-0.465** ***	-0.103	**-0.825** ***	**-0.191** ***
3	dummy after introduction, non-user	**0.171** *	**0.222** ***	0.099	0.139	0.110	0.121	-0.003	0.032	-0.008	-0.016	0.056	0.072
4	Holiday shopping season	-0.022	-0.027	-0.023	-0.019	-0.005	-0.013	**0.270** ***	**0.256** ***	**0.205** ***	**0.191** ***	**-0.415** ***	**-0.373** ***
5	General promotion	**0.183** ***	**0.166** ***	**0.230** ***	**0.223** ***	**0.257** ***	**0.258** ***	**0.328** ***	**0.316** ***	**0.277** ***	**0.248** ***	0.001	0.001
6	Fashion promotion	**0.147** **	**0.142** **	**0.251** ***	**0.249** ***	**0.137** *	**0.138** **	**0.175** ***	**0.172** ***	-0.041	-0.044	-0.080	-0.083
7	competitor 1: starts webstore	**-0.195** ***	**-0.199** ***	**-0.157** **	**-0.156** **	-0.033	-0.023	-0.054	-0.060	-0.006	-0.019	-0.051	-0.055
8	competitor 1: major TV advertisement	-0.194	-0.188	**-0.363** **	**-0.351** **	-0.088	-0.070	0.089	0.062	**0.254** *	**0.263** *	-0.111	-0.077
9	competitor 1 advertisements in 2001	**-0.091** **	**-0.086** **	-0.002	-0.002	**0.170** ***	**0.172** ***	0.045	0.045	**0.075** *	0.067	0.041	0.029
10	competitor 2: introduction door-to-door magazine	-0.010	0.008	0.003	-0.003	0.007	0.007	0.002	0.000	-0.006	-0.003	-0.004	-0.001
11	months since intro log	0.077	0.065	0.013	0.007	**-0.357** ***	**-0.362** ***	**-0.188** **	**-0.184** **	**-0.212** **	**-0.197** ***	**-0.166** *	**-0.142** *
12	distance to closest store in miles	**-0.017** *	**-0.011** **	-0.009	-0.003	-0.008	-0.007	**-0.028** ***	**-0.024** ***	**-0.030** ***	**-0.022** ***	**-0.017** *	-0.007
13	gender (0=male, 1=female)	0.084	**0.130** **	**-0.313** ***	**-0.245** ***	-0.021	0.019	-0.031	0.000	-0.045	-0.041	**-0.208** *	**-0.161** **
14	education (0=low/middle, 1=college or higher)	-0.080	-0.057	0.072	0.058	-0.109	-0.102	0.050	0.059	-0.034	-0.044	-0.096	-0.077

[a] This number corresponds to the numbers in the second column in Table 2. Because the type-II tobit model is category-specific, we include variable 2, the category-specific number of web pages (W_{itc}), and leave out variable 1, the overall number of website visits by the individual (W_{it}).

Significant parameters are in bold (90%).

* Zero is not contained in the 90% HPD (highest posterior density) interval.

** Zero is not contained in the 95% HPD interval.

*** Zero is not contained in the 99% HPD interval.

From Table 4, we can draw two main conclusions. First, website visits significantly decrease purchase incidence (yes/no) for all product categories except children's products. The partial relation between website visits and the average amount of money spent in each product category in the store indicates negative signs for three of the six categories: ladies' fashion, accessories, and sports. Second, the parameters associated with the dummy for non-users after introduction indicate that their spending after the introduction does not change and even increases in one category (ladies' fashion). That is, the decrease in spending is not a general phenomenon after the introduction of a website: *It only takes place for users of the website.*

We find positive effects for the general and fashion promotions for most categories, though the holiday shopping season increases spending in only a few categories. During the holiday season, other, more specialized retail outlets may be more competitive than the department store for categories such as fashion and children. The fashion promotion has a positive effect on purchase incidence and amount spent in the fashion categories, as well as in the children and accessories categories.

We observe mixed effects of competitive actions. We also find a general negative trend (the log of the number of months since the introduction of the website) effect on purchase incidence and amount of money spent, with the two fashion categories as exceptions.

We also have included various customer characteristics in the model. The results show that the distance to the store has a significant negative effect on the purchase decision and amount spent in several categories. Moreover, women spend more in ladies' fashion but less in the men's fashion and sport categories. We find no significant effect of education.

With the multivariate type-II tobit model, we can investigate which categories are correlated. The coefficients that indicate contemporary correlation suggest cross-category or co-occurrence effects (Manchanda et al., 1999) for the ladies' fashion category with men's fashion, children's products, accessories, and sports. In addition, the men's fashion category exhibits co-occurrence effects with accessories and sports. All these cross-category effects suggest that the categories are complementary.

5.3. Individual-specific website effect parameters

The main finding of our preceding analyses is the negative effect of website visits on purchase incidence in category *c* (*c* = 1, ..., 6) and the amount spent. This conclusion does not hold for all customers in our data set though; for some of them, website visits have a positive effect. The unobserved heterogeneity parameters of the type-II tobit model ($\alpha_{ic}^{(1)}$ and $\alpha_{ic}^{(2)}$) identify these customers. The percentage of website users that experience a negative effect from using the website is substantially larger than the percentage of those who experience a positive effect. This finding is particularly true for the parameters that describe the effect of website visits on money spent; very few registered site users (0–2.45%) reveal positive effects. The percentage of users with negative website visit effects instead varies across categories, from 74% to as much as 98%, as we show in Table 5.

Table 5.
Percentage of website users with significant (95% confidence) negative or positive website effects for purchase incidence and spending

	Effects on Purchase Incidence		Effects on Amount Spent	
Category	Negative	Positive	Negative	Positive
Ladies' fashion	2.45%	1.78%	91.76%	2.00%
Men's fashion	2.23%	1.34%	97.55%	0.89%
Children	3.12%	1.78%	84.86%	1.56%
Accessories	2.90%	2.67%	74.16%	2.45%
Living	3.12%	1.34%	98.44%	0.00%
Sport	5.57%	2.00%	96.66%	0.89%

We compare the characteristics and behavior of the positively and negatively affected users using two-sided, independent samples t-tests, with a confidence level of 95%. The positively affected users spend significantly more money than those who are negatively affected, in particular in the children and living categories, and engage in more shopping trips. The analyses also reveal that registered users who are affected positively have conducted significantly more website visits and look at significantly more web pages.

Finally, we compare the flow (Csikszentmihalyi, 1990; Hoffman & Novak 1996) that website users experience while browsing the site,[13] on the basis of a survey conducted in May 2002 among registered users. Across all categories, users that exhibit negative website effects experience a significantly lower flow than those who experience a positive website effect (2.33 versus 2.62, two-sided t-test, $p = 0.081$). For the effects on the amount spent, we come to a similar conclusion (2.38 versus 2.66, two-sided t-test, $p = 0.002$). These comparisons require cautious interpretation but also add face validity to the individual-specific effects in our model.

6. Conclusions

6.1. Findings

For most customers, the introduction and use of a retailer's informational website has negative effects on the number of shopping trips they take, their decision to buy in a particular category, and the amount of money they spend across all six categories that we analyzed. Although we have no conclusive evidence, we consider several factors that might contribute to these negative results. First, customers exhibit more planned shopping behavior as consequence of their access to and use of more information. Second, the information on a website can be easily

[13] For website usage, Hoffman and Novak (1996) define flow as a state characterized by a seamless sequence of responses that are facilitated by machine interactivity, intrinsic enjoyment, a loss of self-consciousness, and self-reinforcement. This state enhances attitudes toward a website (Mathwick & Rigdon 2004) and behaviors such as depth of search and repeat visits (Hoffman & Novak 1996). We measure flow with a five-point scale that features items such as, "I often forget my immediate surroundings," "I often do not realize the duration of my Web visit," and "Time seems to fly by," which we then combine into one construct.

compared with information from competitive stores. Third, the quality of the website might also explain the negative effect. We discuss each of these explanations in more detail.

Additional information leads to more planned behavior (Ajzen, 1991). As a consequence, consumers may reduce the number of visits to the store. With respect to the amount spent, being more informed about what they want might help consumers self-regulate better when they are in the store (Baumeister, 2002) and thus spend less money. While browsing the website, it is not possible to touch and feel the product, and such stimuli tend to induce impulse buying behavior (Peck & Childers, 2006; Rook, 1987; Rook & Hoch, 1985). According to Underhill (1999, p. 158), “almost all unplanned buying is a result of touching, hearing, smelling or tasting something on the premises of the store.” When website users cannot approach the product, it may reduce their impulse buying and thus their spending.

Shiv and Fedorikhin (1999) also indicate that in situations with scarce processing resources, more impulsive customers choose products on the basis of their spontaneous evoked affect rather than cognitions. Providing highly impulsive customers with an informational website, with which they can interact but *not transact*, puts them in a position from which they can engage sufficient processing resources, with more emphasis on cognition. Our findings indicate that for customers who use the website, offline purchase incidence drops, as does spending in the store. Therefore, they should have fewer opportunities to choose products in response to the affect the products evoked as they walk through the store. Therefore, website visits may reduce impulse buying behavior in the physical store.

Switching costs are very low in online environments. Competitors are just a click away, and psychological bonds are loose (e.g., Neslin et al., 2006). In this respect, we emphasize that one of the retailer’s main competitors opened a website several months before the introduction of the website by the focal department store. The informational website offers information that can be compared easily with the information provided by competitors that also have informational or even transactional websites. After obtaining information on the website of the focal company, consumers may have an impulse to buy (Rook, 1987), but they can do so only on the competitors' transactional websites. Unfortunately, we did not observe whether customers actually make this switch, but Huang et al. (2009) find that this behavior is common among people who browse (transactional) websites.

The extent to which customers appreciate the website also might help explain the behavior we observe. In the May 2002 survey we used to obtain the flow data, we also found, using five-point scales, that average customer satisfaction with the website was 3.43 and average satisfaction with the store was 3.78. The averages differ significantly from the neutral point of 3 ($p < 0.001$). However, because satisfaction figures are often skewed to the right (Peterson & Wilson, 1992), we refrain from concluding that they are positive ratings. The average customer satisfaction measures do not differ significantly between users who were positively versus negatively influenced ($p = 0.596$ and 0.242, respectively, two-sided t-tests).

6.2. Managerial implications

In this study, we find that on the firm level, multiple channels provide customers, but not necessarily firms, with benefits. There are some alternatives open to the retailer to cope with our empirical findings. First, for a small percentage of customers (less than 3% across the six categories), visiting the website has a positive impact on the amount spent in the store. If these customers contribute substantially to the firm's revenues, an exclusive website makes sense.

Second, managers could consider changing the content of their informational sites to obtain more positive effects. For example, if consumers tend to become more efficient buyers, site content that enables them to buy more efficiently (e.g., overviews of available merchandise, detailed product and price information) harms the firm more and should be excluded. The site instead should focus on information that has a positive effect, such as references to brand building efforts, the availability and arrival of new services, and any news that makes customers curious to visit the store. In this respect, products sold only by one retailer likely have different effects than products that are not unique and can be sold in virtually the same conditions elsewhere (Pauwels et al., 2011). More (experimental) research among a firm's consumers could provide specific insights, before the introduction of the website.

Third, managers may upgrade their informational websites to transactional websites or at least link informational to transactional websites. For example, Zara could link its informational website to a Bloomingdale's site where customers can make online purchases of its clothing. This option is possible only if customers are willing to buy online, the merchant is able to execute the site in a cost-effective way, and the negative effects of online transactional channels in previous studies are smaller than those for the informational site (e.g., Ansari et al., 2008; Gensler et al., 2007).

Fourth, it is useful to monitor the effects of a new channel on existing channels as soon as possible after its introduction. Monitoring satisfaction scores of the website is, in this respect, less useful than monitoring its effects on metrics such as the average number of shopping trips, revenues, amount spent per trip, and so on.

Fifth, this study indicates that a decomposition may provide insights into which components of consumer behavior are affected by which variables. Although we generally find the same results in the two components, it is not unimaginable that we might find results that differ between the two components, such that the reduction in the number of visits might be offset by higher spending per visit. Our model can detect such patterns.

It also is interesting to consider the type of website(s) the retailer currently hosts, several years after we performed our study and presented the results to it. The focal retailer (as of January 2011) maintains three websites, two informational and one transactional. One informational website is targeted at a specific consumer segment, and the other pertains to a specific service, namely, the restaurant chain owned by the department store. This service is something that no other department store in the Netherlands provides, making it a unique offer that could make the store more attractive to visit. In addition, both of the store's informational sites contain links to its transactional website.

6.3. Limitations and further research

For this study, we had data from only one (large) retailer. Thus, we need additional data to determine the extent to which our findings generalize to other contexts, including other product types (e.g., search versus experience goods), competitive positions (e.g., discounters versus high-end retailers), and assortments (e.g., merchants with specialized offerings versus those that offer one-stop shopping). The results from studies that focus on a more general (non–firm-specific) level indicate that consumers benefit from using multiple channels (e.g., Burke, 2002; Verhoef, Neslin, & Vroomen, 2007) and that firms' offline channels benefit from this behavior too. Our study, together with work by Van Baal and Dach (2005), Gensler et al. (2007), and Ansari et al. (2008), indicates this benefit for the firm may not hold. Further research should determine the circumstances and/or conditions in which both firms and customers benefit from a multichannel environment. Experimental settings also could offer more insights into which factors have positive and negative effects on offline behavior. In future research, if data allow, it would be interesting to compare informational websites with transactional websites (or a combination) and thereby consider their impact on both online and offline sales.

Many variables can affect cross-channel behavior; we consider mostly variables pertaining to individual behavior, either offline or online. No data were available for marketing instruments, such as regular prices, features, and displays. Consequently, our results are less suitable for forecasting purposes; rather, they are intended primarily to provide insights into how the use of an informational website influences buying behavior in the offline store. Also, we investigated the effects on existing customers only and studied a single website with specific content.

An extension could be to estimate heterogeneity parameters for both model components simultaneously, as Bucklin, Gupta, and Siddarth (1998) do. However, because our model already takes endogeneity into account, it would result in a very complex model, so we choose to leave this extension for further research.

Nevertheless, our study provides a preliminary clarification of the impact an informational website can have on the offline buying behavior of existing customers. The models we have developed can be applied easily to other situations in which both online and offline information about individual customers is available. In conclusion, our research demonstrates that the implementation of an informational website should be undertaken with great care.

Appendix

It is possible that consumers who visit the store more often, due to some unobserved factor, also end up visiting the website more often. In this case, the independent variable "number of web visits" is correlated with the error term in the main equation (equations 3, 5 and/or 7). If this endogeneity is ignored, we may overestimate the effect of the number of website visits on the number of store visits, or amount spent.

To account for the endogeneity in the Poisson model, we use the instrumental variable (IV) approach discussed in Mullahy (1997). Instead of including the number of website visits directly into equation (3), we first 'regress' this variable on instrumental variables. As instrument, we use the first differences of the number of web visits (W_{it}). This variable has a .46 correlation with the number of website visits (W_{it}) and a correlation of -.06 with the number of store visits (V_{it}). The predicted variable from the instrumental variables regression is therefore indeed a suitable instrument to replace the original number of website visits. The estimates reported in the main text are based on this IV-approach.

The Bayesian estimation of the tobit-2 model allows us to model endogeneity as follows. We introduce an additional model for category-specific website visits. Just like the amount spent, the number of web visits is a truncated variable, because it never takes on values below 0. It has a similar amount of observations where the number of web visits is equal to zero. Therefore, we again employ the tobit-2 framework for relating the number of website visits to the amount spent. Below, we describe the approach for the first stage of the tobit model (the decision to visit the category on the website). For the second stage the approach is analogous and can be obtained by replacing the superscript (1) with a (2).

Parallel to the equation for the decision to spend $Y_{itc}^{(1)}$, we add an equation for the decision to visit the website. We denote the decision to visit the website by $W_{itc}^{(1)}$, which takes on a value of 1 when the individual decides to visit the website and 0 otherwise. Like $\gamma_{ic}^{(1)}$ in the equation for $Y_{itc}^{*(1)}$ (equation 5), the equation for $W_{itc}^{*(1)}$, the unobserved utility of visiting the website, contains an individual-specific unobserved random intercept, denoted by $\delta_{ic}^{(1)}$. More formally, for the number of website visits, we add the following equation for the latent utility of visiting the website in category *c*.

$$W_{itc}^{*(1)} = \beta_c^{(1)} R_{itc}^{(1)} + \beta_{ic}^{(1)} Y_{i,t-1,c}^{*(1)} + \delta_{ic}^{(1)} + \eta_{itc}^{(1)} \quad . \tag{A1}$$

The vector $R_{itc}^{(1)}$ contains individual-specific explanatory variables, such as age, gender and education, so that observed individual heterogeneity is taken into account, resulting in a more precise estimate of the unobserved heterogeneity. The parameter $\beta_c^{(1)}$ captures the influence of these explanatory variables on the website visits in the various product categories in month *t*. The parameter $\beta_{ic}^{(1)}$ reflects the individual-specific effect of the previous decision to spend on the decision whether to visit the website. The parameter $\delta_{ic}^{(1)}$ is an individual-specific random intercept. The error term $\eta_{itc}^{(1)}$ follows a multivariate normal distribution, with mean 0 and covariance matrix $\Omega^{(1)}$.

To introduce the correlation between the unobserved random intercepts of the spending and website visits models, we add $\rho_c^{(1)}\delta_{ic}^{(1)}$ to equation (5) . The parameter $\rho_c^{(1)}$ captures the category-specific correlation of the random intercepts $\gamma_{ic}^{(1)}$ and $\delta_{ic}^{(1)}$. This equation now looks as follows.

$$Y_{itc}^{*(1)} = \alpha_c^{(1)} X_{itc}^{(1)} + \alpha_{ic}^{(1)} W_{itc}^{(2)} + \gamma_{ic}^{(1)} + \rho_c^{(1)} \delta_{ic}^{(1)} + \varepsilon_{itc}^{(1)}, \tag{A2}$$

where $W_{itc}^{(2)}$ denotes the number of website visits by individual *i* in website category *c* in period *t*. The individual-specific random intercept $\gamma_{ic}^{(1)} + \rho_c^{(1)}\delta_{ic}^{(1)}$ captures to what extent an individual is inclined to decide to spend in the store. The unobserved random intercept $\delta_{ic}^{(1)}$ now appears in both equation (A1) and (A2). The parameter vector ρ_c captures the potential correlation of the number of website visits, described in equation (A1) with the unobserved individual-specific component driving the store spending, described in (A2). It is this correlation that is the root of the endogeneity problem. If $\rho_c^{(1)}$ is unequal to zero, endogeneity is present for the decision to spend in category *c*. We refer to Van Nierop, Fok and Franses (2008) for a similar approach. For the spending stage of the tobit-2 model, the equations are formed analogously, by replacing the superscript (1) with (2).

As an example for stage 2, consider an individual who often visits a certain category *c* on the website, more often than one would expect based on the total model in equation (A1, for stage $^{(2)}$). Therefore, this individual has a positive value for $\delta_{ic}^{(2)}$. If for this category the value of $\rho_c^{(2)}$ is positive, equation (A2$^{(2)}$) indicates that we expect this person to also spend more in category *c*, owing to the positive value of $\rho_c^{(2)}\delta_{ic}^{(2)}$. Conversely, if for a particular category $\rho_c^{(2)}$ appears to be insignificant, we may conclude that there is no endogeneity for this category, with regards to the amount spent and amount of web visits engaged in by the individual.

References

Accenture (2007), http://newsroom.accenture.com/article_display.cfm?article_id=4529, accessed July 6, 2010.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179-211.

Alba, J., & Lynch, J. (1997). Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing*, 61(3), 38-54.

Allenby, G. M., & Rossi, P.E. (1999). Marketing models of consumer heterogeneity. *Journal of Econometrics*, 89, 57-78.

Amemiya, T. (1985). *Advanced Econometrics*. Oxford, UK: B Blackwell.

Ansari, A., Mela, C., & Neslin, S. (2008). Customer channel migration. *Journal of Marketing Research*. 45(1), 60-76.

Bakos, Y., & Brynjolfsson, E. (2000). Bundling and competition on the Internet. *Marketing Science*, 19(1), 63-82.

Baumeister, R.F. (2002). Yielding to temptation: Self-control failure, impulsive purchasing, and consumer behavior. *Journal of Consumer Research*, 28 (4), 670-676.

Biyalogorsky, E., & Naik, P. (2003). Clicks and mortar: The effect of online activities on offline sales. *Marketing Letters*, 14, 21-32.

Bucklin, R.E., & Sismeiro, C. (2003). A model of Web site browsing behavior estimated on clickstream data. *Journal of Marketing Research*, 40(3), 249-267.

Bucklin, R.E., S. Gupta & S. Siddarth (1998), Determining segmentation in sales response across consumer purchase behaviors, *Journal of Marketing Research*, 35 (May), 189-197.

Burke, R. R. (2002). Technology and the customer interface: What consumers want in the physical and virtual store. *Journal of the Academy of Marketing Science,* 30(4), 411-432.

Campbell, D. T., & Stanley, J.C. (1963). *Experimental and Quasi Experimental Designs for Research*. Boston: Houghton Mifflin.

Chib, S. (1992). Bayes inference in the tobit censored regression model. *Journal of Econometrics*, 51(1-2), 79-99.

Chib, S., & Greenberg E. (1998). Analysis of multivariate probit models. *Biometrika,* 85 (2), 347-361.

Coelho, F., Easingwood, C., & Coelho, A. (2003). Exploratory evidence of channel performance in single vs. multiple channel strategies. *International Journal of Retail and Distribution Management*, 31(11), 561-573.

Cook, T. D., & Campbell, D.T. (1979). *Quasi-Experimentation: Design and Analysis for Field Settings*. Boston: Houghton Mifflin.

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row

Deleersnyder, B., Geyskens, I., Gielens, K., & Dekimpe, M. (2002). How cannibalistic is the Internet channel? *International Journal of Research in Marketing*, 19(4), 337-348.

Dholakia, R. R., Zhao, M., & Dholakia, N. (2005). Multichannel retailing: A case study of early experiences. *Journal of Interactive Marketing*, 19(2), 62 – 74.

Fox, E.J., Montgomery, A.L., & Lodish, L.M. (2004). Consumer shopping and spending across retail formats. *Journal of Business,* 77(2), 25-60.

Franses, P.H., & Paap, R. (2001). *Quantitative Models in Marketing Research.* Cambridge: Cambridge University Press.

Gelman, A., Carlin J.B., Stern, H.S., & Rubin, D.B. (1995). *Bayesian Data Analysis*, New York: Chapman and Hall.

Geman, S., & Geman, D. (1984). Stochastic relaxations, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence,* 6, 721-741.

Gensler, S., Dekimpe, M.G., & Skiera, B. (2007). Evaluating channel performance in multi-channel environments. *Journal of Retailing and Consumer Services*, 14(1), 17-23.

Geyskens, I., Gielens, K., & Dekimpe, M.G. (2002). The market valuation of Internet channel additions. *Journal of Marketing*, 66(2), 102-119.

Greene, W. (2002). *Econometric Analysis*. Englewood Cliffs, NJ: Prentice Hall International.

Hoffman, D. L., & Novak, T.P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing*, 60(3), 50-68.

Hoffman, D. L., Novak, T.P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM*, 42(4), 80-85.

Hoque, A. Y., & Lohse, G.L. (1999). An information search cost perspective for designing interfaces for electronic commerce. *Journal of Marketing Research*, 36(3), 387-394.

Huang, P., Lurie, N., & Mitra, S (2009). Searching for experience on the Web: An empirical examination of consumer behavior for search and experience goods. *Journal of Marketing*, Mar 2009, 73, 55-69.

Krillion (2008), http://www.krillion.com/xAV-news-20080324_etailing_survey, accessed May 31, 2010.

Kushwaha, T. L., & Shankar, V. (2007). Optimal multichannel allocation of marketing efforts. *MSI Report*, 07-207, 2007.

Lee, R. P., & Grewal, R. (2004). Strategic responses to new technologies and their impact on firm performance. *Journal of Marketing*, 68(4), 157-171.

Manchanda, P., Ansari, A., & Gupta, S. (1999). The shopping basket: A model for multicategory purchase incidence decisions. *Marketing Science*, 18(2), 95-115.

Mathwick, C., & Rigdon, E. (2004). Play, flow and the online search experience. *Journal of Consumer Research*, 31(September), 324-32.

McKelvey, R., & Zavoina, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, 4, 103-120.

Mendelsohn, T., Johnson, C.A., & Meyer, S. (2006). Understanding US cross-channel shoppers. Forrester Research, April 19.

Mick, D. G., & Fournier, S. (1998). Paradoxes of technology: Consumer cognizance, emotions and coping strategies. *Journal of Consumer Research*, 25(2), 123-143.

Moe, W. W., & Fader, P.S. (2004). Capturing evolving visit behavior in clickstream data. *Journal of Interactive Marketing*, 18(1), 5-19.

Mullahy, J. (1997), Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior. *The Review of Economics and Statistics*, 79(4): 586-593.

Neslin, S. A., Grewal, D., Leghorn, R., Shankar, V., Teerling, M.L., Thomas. J.S., & Verhoef, P.C. (2006). Challenges and opportunities in multichannel customer management. *Journal of Services Research,* 9(2) [Special Issue on Customer Relationship Management], 1-18.

Neslin, S.A, & Shankar, V. (2009). Key issues in multichannel customer management: Current knowledge and future directions. *Journal of Interactive Marketing,* 23(1), 70-81.

Pauwels, K., Leeflang, P.S.H., Teerling, M.L., & Huizing, K.R.E. (2011). Does online information drive offline revenues? Only for specific products and consumer segments! *Journal of Retailing*, forthcoming.

Peck, J. and T. L. Childers (2006). If I touch it I have to have it: Individual and environmental influences on impulse purchasing. *Journal of Business Research*, 59(6), 765-769.

Ratchford, B. T., Lee, M., & Talukdar, D. (2003). The impact of the Internet on information search for automobiles. *Journal of Marketing Research*, 40(2), 193-209.

Peterson, R.A. & Wilson, W.R. (1992). Measuring customer satisfaction: Fact and artifact. *Journal of the Academy of Marketing Science*, 20 (1), 61-71.

Rook, D.W. (1987). The buying impulse. *Journal of Consumer Research*, 14(2), 189-199.

Roy, S., & Ghose, S. (2006). Internet adoption as a two-stage transition. Converting Internet non-users to Internet users and to online buyers. *International Journal of Market Research*, 48(3), 321-349.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, 26(3), 214-224.

Sismeiro, C., & Bucklin, R.E. (2004). Modeling purchase behavior at an e-commerce website: A task-completion approach. *Journal of Marketing Research*, 41(3), 306-323.

Tanner, M.A., & Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398), 528-540.

Teo, T., S.H., & Y. Pian. (2004). A model for Web adoption. *Information & Management*, 41 (4), 457-468.

Underhill, P. (1999). *Why We Buy: The Science of Shopping*. New York: Simon and Schuster; 1999.

Van Baal, S., & Dach, C. (2005). Free riding and customer retention across retailers' channels. *Journal of Interactive Marketing*, 19(2), 75-85.

Van Nierop, E., D. Fok, and P.H. Franses (2008). Interaction between Shelf Layout and Marketing Effectiveness and its Impact on Optimizing Shelf Arrangements. *Marketing Science*, 27 (6): 1065-1084.

Veall, M.R., & K.F. Zimmermann (1994). Goodness of fit measures in the tobit model. *Oxford Bulletin of Economics & Statistics*, 56(4), 485-499.

Verhoef, P.C., Neslin, S.A., &Vroomen, B. (2007). Multichannel customer management: Understanding the research-shopper phenomenon. *International Journal of Research in Marketing*, 24(2), 129-148.

Viswanathan, S., Kuruzovich, J., Gosain, S., & Agarwal, R. (2007). Online infomediaries and price discrimination: Evidence from the automotive retailing sector. *Journal of Marketing*, 71(3), 89-107.

Wallace, D. W., Giese, J.L., & Johnson, J.L. (2004). Customer retailer loyalty in the context of multiple channel strategies. *Journal of Retailing*, 80(4), 249 - 263.

Weathers, D., Sharma, S., & Wood, S. (2007). Effects of online communication practices on consumer perceptions of performance uncertainty for search and experience goods. *Journal of Retailing*, 83 (4), 393-401.

Wooldridge J. (1997). Quasi-likelihood methods for count data. In M. Pesaran and P. Schmidt, eds. *Handbook of Applied Econometrics, Vol. II: Microeconomics*. London: Blackwell Publishers.

Wu, F., Mahajan, V., & Balasubramanian, S. (2003). An analysis of e-business adoption and its impact on business performance. *Journal of the Academy of Marketing Science,* 31 (4), 425-447.

Wu, J., & Rangaswamy, A. (2003). A fuzzy set of search and consideration with an application to an online market. *Marketing Science*, 22(3), 441-434.