



Impact of online channel use on customer revenues and costs to serve: Considering product portfolios and self-selection

Sonja Gensler ^{a,*}, Peter Leeflang ^{a,b,1}, Bernd Skiera ^{c,2}

^a Department of Marketing, University of Groningen, Post Box 800, 9700 AV Groningen, The Netherlands

^b LUISS Guido Carli, Rome, Italy

^c Department of Marketing, University of Frankfurt, Grüneburgplatz 1, 60323 Frankfurt am Main, Germany

ARTICLE INFO

Article history:

First received in 5, March 2010 and was under review for 9 months

Available online 8 March 2012

Area Editor: Koen H. Pauwels

Keywords:

Channel management
Self-selection effect
Product portfolio
Matching method
Online marketing
Distribution

ABSTRACT

Developing a strategy for online channels requires knowledge of the effects of customers' online use on their revenue and cost to serve, which ultimately influence customer profitability. The authors theoretically discuss and empirically examine these effects. An empirical study of retail banking customers reveals that online use improves customer profitability by increasing customer revenue and decreasing cost to serve. Moreover, the revenue effects of online use are substantially larger than the cost-to-serve effects, although the effects of online use on customer revenue and cost to serve vary by product portfolio. Self-selection effects also emerge and can be even greater than online use effects. Ignoring self-selection effects thus can lead to poor managerial decision-making.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Channels are an integral part of any firm's customer management strategy, and many firms have established online channels, often in an attempt to improve customer profitability (Gensler, Dekimpe, & Skiera, 2007; Hitt & Frei, 2002). Popular press reports imply that there is a positive association between customers' online use and customer profitability; an e-commerce executive at Bank of America asserted, for example, that the firm's 12.6 million online banking customers are 27% more profitable than its offline customers (Tedeschi, 2005). Such statements likely encourage managers to push customers to online channels. There are several reasons to believe, however, that simple comparisons of online and offline customer profitability can provide only limited insight into the effects of online use.

First, differences in profitability between online and offline customers may reflect self-selection effects, which exist if online and offline customers differ in characteristics such as age. Differences in customer profitability thus would be driven not by customers' online use but mainly by differences in these customer characteristics. Accordingly, firms must account for self-selection effects when investigating whether customers' online use affects key metrics such as customer profitability and whether managing customers' use of channels is even useful.

Second, firms need to identify the primary function of their online channels, and they need to separate the effects of customers' online use into the specific ways in which it influences customer revenue and cost to serve. Online channels can develop and enhance customer relationships if their use improves customer revenue (Blattberg, Kim, & Neslin, 2008). If online channels allow for decreasing the cost to serving customers, they are particularly suited to managing customer relationships that have limited cross- or up-selling opportunities. Thus understanding the revenue and cost-to-serve effects of online use suggests which strategy to pursue to ensure effective customer relationships.

Third, firms must determine whether the effects of customers' online use vary depending on the mix of products used by the customers (i.e. product portfolio). If the effects of online use vary by customer product portfolio, a better understanding of the differential effects of customers' online use would offer key insights into which customers the firm should target when aiming to actively manage customers' use of channels.

Accordingly, we aim to determine the effects of customers' online use on customer revenue and cost to serve by taking into account self-selection effects and the moderating impact of the customer's product portfolio. We further study the effect of customers' online use on customer profitability empirically. To determine the extent of self-selection effects, we use data pertaining to approximately 87,000 customers of a large European retail bank. Our results offer three main contributions.

First, our study reveals *both* the revenue and cost effects of customers' online use, enabling us to assess which effect is the primary driver of differences in customer profitability. In contrast, most previous studies (see Table 1) have focused on either customer profitability or

* Corresponding author. Tel.: +31 50 363 9086; fax: +31 50 363 3720.

E-mail addresses: S.Gensler@rug.nl (S. Gensler), p.s.h.leeflang@rug.nl (P. Leeflang), skiera@skiera.de (B. Skiera).

¹ Tel.: +39 50 363 3696; fax: +39 50 363 3720.

² Tel.: +49 69 798 34649; fax: +49 69 798 35001.

Table 1

Studies that associate online use with customer profitability, customer revenue, and cost to serve.

Study	Variables of interest			Moderating effect of product portfolio	Control for self-selection effects
	Customer profitability	Customer revenue	Cost to serve		
Ansari et al. (2008)		✓			Type of instrumental variables approach
Campbell and Frei (2010)	✓		✓		Propensity score matching
Hitt and Frei (2002)	✓				Covariate matching
Kumar and Venkatesan (2005) ^a		✓			No explicit correction
Pauwels et al. (2011)		✓			No explicit correction
Thomas and Sullivan (2005) ^a		✓			No explicit correction
Venkatesan, Kumar, and Ravishanker (2007) ^a	✓				No explicit correction
Xue et al. (2011)	✓				Propensity score matching
Xue et al. (2007)	✓				OLS using observed characteristics
This study	✓	✓	✓	✓	Hybrid matching

OLS = ordinary least squares.

^a Investigates the effect of multichannel use, which usually indicates that customers use both offline and online channels.

customer revenue in relation to online use. Campbell and Frei (2010) provide insights into the profitability and cost effects of online use, but the way in which they present their results does not allow any inference regarding the revenue effects of online use. Consequently, our knowledge about how to employ online channels to manage customer relationships remains quite limited, as also noted by Neslin and Shankar (2009).

Second, we investigate for the first time the moderating effects of the customer's product portfolio on revenue effects and cost effects of customers' online use. Although some studies note product-specific effects of online use (see Table 1; Hitt & Frei, 2002; Pauwels, Leeflang, Teerling, & Huizingh, 2011; Thomas & Sullivan, 2005; Xue, Hitt, & Chen, 2011), research has yet to examine the effects of online use when customers combine products with different characteristics. Such knowledge is necessary if firms are to manage their customers' use of channels effectively.

Third, we demonstrate the importance of accounting for self-selection effects. Only a few previous studies have considered self-selection effects at all (Table 1). We control for self-selection effects through hybrid matching. By introducing the method of hybrid matching into the marketing literature, we also demonstrate the viability of this approach.

The remainder of this article proceeds as follows. First, we discuss the expected effects of online use on customer revenue and on cost to serve. Next, we explain the importance of controlling for self-selection effects and describe different methods for doing so. We then describe our data, detail our methodology and present the empirical results. We end with a discussion of our findings, managerial implications, research limitations, and suggestions for further research.

2. Conceptual development

By employing an online channel, managers aim to increase customer profitability by improving customer revenue and decreasing

cost to serve. Online channels can increase customer revenue and decrease cost to serve if the use of online channels leads customers to alter their behavior (Degeratu, Rangaswamy, & Wu, 2000). Customer behavior is in this study reflected by customers' demand for one or more products and the number of transactions that customers undertake in each channel. Product demand influences customer revenue, which is defined as the customer's product demand multiplied by the contribution margin earned by the firm on a particular product. Meanwhile, the cost to serve a customer is the product of the number of transactions a customer undertakes in a channel multiplied by the channel-specific costs.

In the following section, we consider the incremental effect of online use on customer behavior, in line with previous studies, because most online customers use traditional offline channels as well (e.g., Chu, Chintagunta, & Cebollada, 2008). Online customers are thus defined as customers who use an online channel for at least some transactions, even if they do not use that channel exclusively.

2.1. Effect of customers' online use on customer revenue

Online use can increase the customer's product demand for several reasons. When searching for product information, customers who use online channels generally perceive that they have greater information control than do customers using offline channels. These perceived measures of control include the selection of which information is presented, for how long, and which information follows from it (Ariely, 2000). Greater information control online thus improves the customer's ability to understand information relevant to the customer's choices, which is a critical determinant of decision making (Weathers, Sharma, & Wood, 2007). Furthermore, online channels often provide interactive decision tools that help customers to picture themselves using the product, a feature that should increase their purchase likelihoods (Huang, Lurie, & Mitra, 2009; Schlosser, 2003). Online channels also offer greater convenience and accessibility than do offline channels (Brynjolfsson, Hu, & Smith, 2003; Montoya-Weiss, Voss, & Grewal, 2003); customers do not have to consider opening hours or wait in checkout lines. Because online use should increase the customer's product demand, it should also enhance customer revenue (Hitt & Frei, 2002). We expect, in turn, online customer revenue, which combines online and offline revenue, to be higher than revenue earned from customers who use offline channels exclusively.

H1. Online use increases customer revenue.

The verification of this hypothesis is particularly important because previous research into the association between online use and customer revenue frequently has ignored self-selection effects (Table 1).

2.2. Effect of customers' online use on cost to serve

Beyond increasing customer revenue, firms invest in online channels because the costs per transaction are much lower than with offline channels, decreasing the cost to serve customers (Campbell & Frei, 2010). Online channels also reduce the cost from a customer's perspective (e.g., no travel, no waits) and should improve overall customer efficiency by lowering the marginal cost of transactions (Bitner, Brown, & Meuter, 2000), assuming customers substitute offline transactions with online transactions (Xue et al., 2011).

The reduction in marginal costs from the customer's perspective may, however, increase the total number of transactions. That is, online customers may engage in more transactions because they expend fewer resources on any single transaction (Xue et al., 2011). Whether the firm's total cost to serve customers increases or decreases, therefore, depends on how customers allocate their transactions across channels. More transactions can reduce costs only if customers substitute some of their costly offline transactions with less costly online

transactions; without such a substitution effect, the cost to serve increases. Because online use reduces the marginal cost of a transaction to the customer and increases the customer's efficiency, we expect customers to substitute online transactions for offline transactions. The substitution effect, combined with the lower cost of online transactions from the firm's perspective, may reduce the cost to serve an online customer. We propose the following:

H2. Online use lowers the cost to serve a customer.

Conventional wisdom suggests that online use decreases the cost to serve, yet empirical evidence is scarce (Table 1). Campbell and Frei (2010) provide the only investigation of online cost to serve, and they find a positive effect: online use increases cost to serve. Our test of H2 thus adds another data point to the discussion of how online use affects the cost to serve a customer.

2.3. Moderating effects of the customer's product portfolio

The size of the effects of online use may depend on a customer's product portfolio (i.e., the products of a firm that the customer uses) because some products tend to be more closely related to a particular channel than others (Inman, Shankar, & Ferraro, 2004; Liang & Huang, 1998). In this context, frequency of use and perceived product risk represent important product characteristics for matching products to channels (Gutiérrez, Izquierdo, & Cabezudo, 2010; Inman et al., 2004). Specifically, frequently used products are likely to benefit more from online use (Danaher, Wilson, & Davis, 2003; Xue, Hitt, & Harker, 2007) because customers search for more information about frequently used products than they do for less frequently used products (Bhatnagar & Ghose, 2004). The increase in search behavior online may result in greater product demand. Furthermore, frequently used products may benefit more from the accessibility and convenience of online channels – that is, online channels should make it easier to use frequently used products, which also should increase the customer's product demand.

Yet many customers associate greater risk with online channels rather than offline channels (Forsythe & Shi, 2003). Research further shows that risk in the channel, in turn, positively influences product-related risk (Gutiérrez et al., 2010). For example, customers may be more afraid of making a mistake when using an online channel (Forsythe & Shi, 2003), often because personalized advice (e.g., from a sales representative) – which can reduce consumers' perceptions of product-related risk (Alba et al., 1997) – is not available online. Additional information on a website (e.g., third-party evaluations, consumer feedback) may help reduce product-related risk, but such information tends to be regarded as less valuable than personalized advice (Weathers et al., 2007). High-risk products thus may appear more closely related to offline than online channels, and high levels of risk could attenuate the positive effect of online use on the customer's product demand (Smith & Sivakumar, 2004). However, we do not expect a negative effect of online use on the customer's product demand for high-risk products because online customers continue to use offline channels and have access to personalized advice. With all these arguments, we propose that the effect of online use on the customer's product demand, and ultimately on customer revenue, depends on the composition of a customer's product portfolio:

H3. The composition of the customer's product portfolio moderates the positive effects of online use on customer revenue.

The cost to serve also depends on product portfolios. For frequently used products, online channels can reduce transaction costs considerably; for example, online channels can eliminate time spent resolving service issues. Customers, therefore, should substitute

online transactions for offline transactions to improve their own efficiency (Campbell & Frei, 2010), with the added effect of decreasing firms' cost to serve.

Again though, this improved efficiency might lead customers to undertake more transactions, especially if they are making decisions about high-risk products. For customers who use high-risk products, an online channel provides continuous access to detailed product information, which may result in increased information monitoring and more active product management (Campbell & Frei, 2010). For example, online brokerage accounts encourage more transactions because customers perceive an opportunity to take personal control of their investments and improve their returns by actively managing their investments (Strader & Ramaswami, 2004). An increasing number of transactions limit the firm's opportunity to reduce cost to serve. We therefore propose that the composition of a customer's product portfolio moderates the effect of online channel use on cost to serve.

H4. The composition of the customer's product portfolio moderates the negative effects of online use on cost to serve.

2.4. Effect of customers' online use on customer profitability

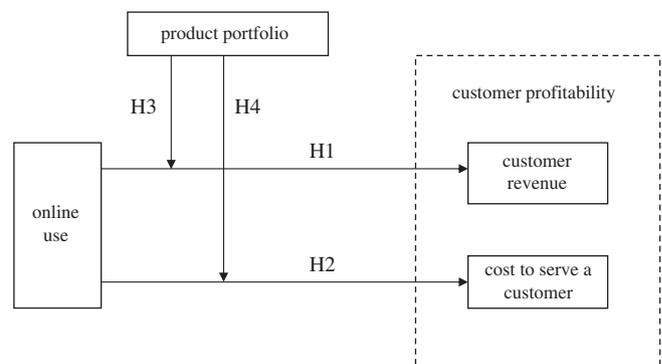
Customer profitability, in a service setting such as retail banking, is a function of customer revenue, cost to serve, and risk costs – the latter being a cost measure determined by the bank that reflects the customer's risk of default in payment (Skiera, Bermes, & Horn, 2011). To compute a customer's risk cost, the bank takes into account customer characteristics such as age and history with the bank.

The influence of online use on customer profitability should depend on its revenue (H1) and cost-to-serve (H2) effects. We predict a positive effect of online channel use on customer profitability because it should improve customer revenue and reduce cost to serve. However, we refrain from formulating a separate hypothesis for this effect on customer profitability because if it exists, it stems from the effects detailed in H1 and H2. Instead, we summarize our conceptual framework and proposed relationships in Fig. 1.

3. Main effects and self-selection controls

3.1. Formal definition of the effects of online use

To determine the effects of online use on customer revenue and cost to serve, we examine the specific effects of online use on product demand and number of transactions that customers undertake, i.e., factors that can translate into changes in customer revenue and cost to serve. The basic problem in identifying these effects is that we can observe both product demand and transactions in either the



Note: This figure only shows the relationships of focal interest. In our empirical application, we control for additional effects not depicted in this figure.

Fig. 1. Conceptual framework.

online or the offline scenario. For example, we can observe the number of online transactions by an online customer only if he or she uses the online channel.

Specifically, we use the average treatment on the treated effect (ATT) to evaluate the effects of online use on product demand and number of transactions (Heckman & Navarro-Lozano, 2004; Zhao, 2004). The ATT provides a basis for assessing whether online use influences product demand and the number of transactions completed among customers who actually use the online channel:

$$ATT_k = E_i(y_{i,k}^1 | d_i = 1) - E_i(y_{i,k}^0 | d_i = 1) \quad \forall k \in K, \quad (1)$$

where d_i represents a binary variable that equals 1 if customer i is online and 0 otherwise. Meanwhile, $(y_{i,k}^1 | d_i = 1)$ is the value of the outcome variable k (e.g., number of transactions completed) for online customer i , and $E_i(\dots)$ is the expected value of k for all online customers. The counterfactual outcome in Eq. (1) that requires estimation is $E_i(y_{i,k}^0 | d_i = 1)$ or the expected value of outcome variable k for all online customers if they were offline.

To determine this counterfactual outcome in Eq. (1), we could use the observed mean values for all offline customers, $(E_j(y_{j,k}^0 | d_j = 0))$. Yet online customers also may differ in their characteristics from offline customers j , and the same characteristics could influence the customer's product demand as well as the number of transactions. Simple mean comparisons, then, cannot represent the effects of online use because self-selection effects (SE) are confounded with the effects of online use. More formally,

$$[E_i(y_{i,k}^1 | d_i = 1) - E_j(y_{j,k}^0 | d_j = 0)] = ATT_k + SE_k \quad \forall k \in K. \quad (2)$$

In Eq. (2), the left-hand side represents the observed difference in means between online and offline customers. SE_k is the self-selection effect for outcome variable k .

3.2. Methods to control for self-selection effects

The most popular approaches to control for self-selection effects are matching, instrumental variables (IV) and control function methods (Heckman & Navarro-Lozano, 2004). The latter two methods control for selection on unobserved characteristics and rely on instrumental variables,³ which must satisfy two conditions: They must be exogenous, and they must be highly correlated with the binary treatment variable (i.e., customers' online use) (Stock, Wright, & Yogo, 2002). If we do not observe variables that meet these conditions, the estimated effects can be highly biased; the biases induced by such weak instrumental variables can be so great that even a simple ordinary least squares (OLS) regression that used observed customer characteristics to control for selection could perform better (Woglom, 2001).

Matching methods do not rely on instrumental variables but instead eliminate self-selection effects by comparing online and offline customers with similar observed characteristics. Such methods thus rely on the selection of observed customer characteristics to build matched samples; the selection should be based on theory and previous empirical findings. If researchers fail to simultaneously include all observed customer characteristics that affect the outcome (e.g., checking account balances) and treatment decision (i.e., to use the online channel), the estimated effects will be biased. Matching methods thus have high data demands. Moreover, the observed customer characteristics must be sufficient to ensure that the outcome variables are independent of the treatment and conditional on customer characteristics – that is, the observed customer characteristics

must address the conditional independence assumption (Rosenbaum & Rubin, 1983). Furthermore, the methods fail if the observed customer characteristics are perfect predictors of the treatment decision because no matching partners can then be identified. Matching methods thus assume, given a set of observed characteristics, that some unspecified randomization is capable of allocating customers to online use (Heckman & Navarro-Lozano, 2004).

There are a variety of matching methods available to build matched samples. *Covariate matching* builds samples on the basis of observed characteristics (Zhao, 2004). Its limitation is that when there are many characteristics that could drive self-selection effects, it becomes infeasible to match customers directly. The Mahalanobis distance can map multiple characteristics into a single measure that expresses the gap between any two customers (Rosenbaum, Ross, & Silber, 2007). However, this approach requires customer characteristics measured on metric scales. Such difficulties have limited the use of covariate matching in the past.

An alternative approach is *propensity score matching* (Dehejia, 2005; Mithas, Krishnan, & Fornell, 2005). The propensity score is in our case the conditional probability that a customer with a specific vector of observed characteristics uses the online channel. This probability can be estimated with a logit or probit model. With the propensity score, we ensure that the distribution of characteristics in the two groups (i.e., online versus offline) is the same (Rosenbaum & Rubin, 1983). Yet propensity score matching cannot guarantee that the matched online and offline customers are directly comparable in all of their characteristics (e.g., age).

Hybrid matching compares online and offline customers according to both the propensity score and certain selected customer characteristics (Rosenbaum & Rubin, 1985), ensuring that matched customers are directly comparable with respect to these characteristics.

In our empirical study, we use customer transaction data and customer characteristics that are typically stored in customer databases (e.g., age, length of relationship). These customer characteristics are not exogenous and thus are not strong instrumental variables. We therefore focus on the use of matching methods and show that IV and control function methods are less appropriate in such a setting (Section 5.4). Specifically, we use hybrid matching to determine the counterfactual outcomes in Eq. (1). We expect that in our setting, hybrid matching may improve the accuracy of the estimated effects because hybrid matching ensures that online and offline customers correspond on selected characteristics in addition to their probability of using the online channel. We subsequently compare the predictive performance of propensity score and hybrid matching and find that the latter yields slightly higher predictive performance (Table 6).

4. Empirical study

4.1. Data

We used transaction data from a random sample of approximately 87,000 private clients of a large European retail bank over a three-month period. Of these clients, 38.4% were enrolled in online banking, although only 1.9% actively used it.⁴ For our purposes, a customer actively uses the online channel if he or she conducts at least two transactions with the firm through the online channel during our three-month observation period.

In Table 2, we provide an overview of the variables. For each customer, we collected information about the monthly number of transactions completed with the retail bank. We also determined the number of transactions each customer completed in each channel. Furthermore, we had information about each customer's product

³ Strictly speaking, the control function method does not require instrumental variables, but Puhani (2000) shows that failing to take instrumental variables into account results in poor performance.

⁴ Customers often are enrolled by bank employees without their active request. Such passive enrollment may result in no use of online banking, so we focus explicitly on customers who actually use online banking.

Table 2
Overview of variables, operationalizations, and summary statistics.

Variable	Operationalization
Number of transactions	Average monthly number of transactions per customer
Number of checking accounts	Number of checking accounts per customer
Number of savings accounts	Number of savings accounts per customer
Number of brokerage accounts	Number of brokerage accounts per customer
Number of credit cards	Number of credit cards per customer
Balance in checking accounts	Average monthly balance of checking accounts in EUR per customer
Balance in savings accounts	Average monthly balance of savings accounts in EUR per customer
Balance in brokerage accounts	Average monthly balance of brokerage accounts in EUR per customer
Balance on credit cards	Average monthly balance of credit card in EUR per customer
Relative balance on credit cards	Average monthly balance of credit card in EUR divided by average monthly balance of checking accounts in EUR per customer
Age	Age of customer in years
Length of relationship	Length of customer's relationship with the bank in months
Joint account	Dummy variable indicating whether customer has a joint checking account or not
Customer revenue	Average monthly customer revenue in EUR = \sum (balance product p * net contribution margin for balance product p) + \sum (number of products p * net contribution margin for product p)
Cost to serve customer	Average monthly cost to serve customer in EUR = \sum average monthly number of transactions in channel c * cost per transaction in channel c
Risk cost	Average monthly cost related to cost of shortfalls in payment (provided by the firm for each customer)
Customer profitability	Average monthly customer profitability in EUR = customer revenue – cost to serve customer – risk cost

demand, denoted by the balances and number of accounts associated with the following products on a monthly basis: checking account, savings account, brokerage account, and credit card.⁵ The number of accounts indicates how many accounts of a particular product a customer uses.

To avoid biases caused by individual fluctuations in balances, we generally used the average value for the three-month observation period. Hence, we obtained one observation for each variable for each customer. We eliminated outliers, following the method proposed by Hadi (1994). In addition, we observed customer age, length of the customer's relationship with the firm, and whether the customer's checking account is a joint account. We used internal information about net contribution margins to compute the (net) revenue earned from each customer. We also gathered information about the cost to serve based on the number of transactions completed within a channel and the associated channel-specific costs. Finally, we calculated each customer's profitability based on customer revenue, cost to serve, and risk cost.

⁵ Information about credit cards is available only if customers also own a checking account. Moreover, the credit card balance is affected by the checking account balance because balances on the credit card are paid from the checking account. To address this interdependence in further analyses, we consider the relative credit card balance by dividing the credit card balance by the checking account balance.

To assess the predictive performance of the hybrid matching method, we collected information about online customers during a three-month period that started 6 months before our focal observation period. We thus were able to obtain information about online customers at two different points in time: the observation period ($t=0$), and an earlier period ($t=-6$ months). At $t=-6$, there were 108 current online customers who were still offline; their use of the online channel started later. These 108 customers represent the holdout sample; we assume that no effects other than online use cause any differences in each person's product demand or number of transactions completed between $t=0$ and $t=-6$. This assumption is defensible, considering the short time frame. We do not include the customers in the holdout sample in our estimates of the effects of online use.

4.2. Implementation of the hybrid matching method

We estimate each customer's propensity to use the online channel as a function of the following: the customer's age (AGE); the length of the customer's relationship with the firm (LOR); the customer's ownership of checking accounts (CHECK), savings accounts (SAV), brokerage accounts (BROK) and credit cards (CREDIT); and whether the checking account is a joint account (JOINT). The ownership variables indicate whether a customer uses a particular product. We use these variables because previous studies show that online customers tend to be younger than offline customers and have shorter relationships with the firm (Campbell & Frei, 2010; Degeratu et al., 2000). Moreover, when a customer uses a specific product, she or he also may be more or less likely to use an online channel. Owners of checking accounts, for example, may find it more convenient to move online than owners of savings accounts, because the online channel offers greater convenience for managing checking accounts than for managing savings accounts, which require little upkeep.⁶ Owners of joint accounts may realize greater cost savings from online banking, making them more prone to use it than other account holders would be (Campbell & Frei, 2010; Hitt & Frei, 2002).

In Table 3, we list the estimated effects of customer characteristics on propensity to use the online channel. As expected, customers who are younger, have shorter relationships with the firm, and have a joint checking account are more likely to use online banking. Customers who own a checking account or credit card are also more likely to use the online channel, whereas owners of savings and brokerage accounts are less likely. However, the latter two effects are not significant.

Next, we use the propensity score to build matched samples of online and offline customers who are comparable in their characteristics. The percentage reduction in bias is an important metric with which to evaluate whether the two groups are more comparable after matching (Table 3). To compute the percentage reduction in bias, we compare the difference in means for characteristics measured on a metric scale (or relative frequencies for characteristics measured on a nominal scale), after matching versus before matching (Rosenbaum & Rubin, 1985):

$$\text{reduction_bias}_m = \left(1 - \frac{|\bar{x}_{d_i=1,m}^{\text{after}} - \bar{x}_{d_j=0,m}^{\text{after}}|}{|\bar{x}_{d_i=1,m}^{\text{before}} - \bar{x}_{d_j=0,m}^{\text{before}}|} \right) \cdot 100 \quad \forall m \in M, \quad (3)$$

where $\bar{x}_{d_i=1,m}^{\text{after}}$ and $\bar{x}_{d_j=0,m}^{\text{after}}$ are the mean values (relative frequencies) for the customer characteristic m for online and offline customers

⁶ We thank an anonymous reviewer for providing this argument.

Table 3
Parameters of customer characteristics in propensity score model (logit model, dependent variable: online use).

Customer characteristic	Parameter	t-Value	Reduction in bias (%) ^a
Age (years)	-.04	-8.51	55.36
Length of relationship (month)	-.02	-19.94	43.94
Joint account	.32	3.31	-82.14
Checking account ownership	5.35	5.34	74.20
Credit card ownership	.48	5.74	34.59
Brokerage account ownership	-.00	-.02	6.94
Savings account ownership	-.03	-.59	35.14
Intercept	-7.22	-7.16	
Sample size = 86,754; log-likelihood = -7023.04, pseudo R ² = .131			

Bold values are significant at the 5% level.

^a Computation based on Eq. (3).

after matching, and $\bar{x}_{d_i=1,m}^{before}$ and $\bar{x}_{d_i=0,m}^{before}$ are the mean values (relative frequencies) for the same customer characteristic for online and offline customers before matching. The percentage reduction in bias thus indicates whether the comparability of the two groups improves after matching (Rosenbaum & Rubin, 1984).

Overall, the reduction in bias is substantial for most characteristics (Table 3), such that online and offline customers become more similar with respect to their observed characteristics after matching. For example, the percentage reduction in bias is 55.36% for the customer characteristic 'age'. Only for the joint account characteristic does the metric decrease, such that the two groups become less comparable. However, we retain this characteristic in estimation of the propensity score because the joint account characteristic has a significant effect and because the difference in the relative frequency before and after matching is rather small (0.010 versus 0.017).

To ensure that the matched online and offline customers are actually comparable, we employ a common support restriction. This restriction eliminates all customers who do not lie within the region of common support (Heckman, Ichimura, & Todd, 1997), a construct defined as the overlap between the propensity score distributions of online and offline customers. This approach excludes customers with a propensity score smaller (larger) than the minimum (maximum) value of the propensity score of the region of common support. In our study, 77.6% of the offline customers and 100% of the online customers were within the region of common support, providing us a sufficient number of matching partners for each online customer (Zhao, 2004).

Studies in a variety of settings indicate that online and offline customers differ substantially in age and length of relationship with the firm (e.g., Hitt & Frei, 2002; Shankar, Smith, & Rangaswamy, 2003). Therefore, we included these characteristics as well as the propensity score when estimating the effects of online use. In other words, matched customers are explicitly comparable with respect to age and length of relationship with the firm. To identify matching partners, we use the Mahalanobis distance (M) (Rosenbaum et al., 2007) and compute the effects of interest with the following specified version of Eq. (1):

$$ATT_k = E_i \left((y_{ik}^1)_{d_i=1} \middle| \hat{p} \left(\begin{matrix} AGE_i, LOR_i, JOINT_i, CHECK_i, \\ SAV_i, CREDIT_i, BROK_i \end{matrix} \right), AGE_i, LOR_i \right) - E_j \left((y_{jk}^0)_{d_j=0} \middle| \hat{p} \left(\begin{matrix} AGE_j, LOR_j, JOINT_j, CHECK_j, \\ SAV_j, CREDIT_j, BROK_j \end{matrix} \right), AGE_j, LOR_j \right) \forall k \in K, \tag{4}$$

where $E_i (y_{ik}^1 | d_i = 1, \dots)$ equals the mean value of outcome variable k across all online customers, and $E_j (y_{jk}^0 | d_j = 0, \dots)$ equals the mean value of outcome variable k across all *matched* offline customers (counterfactual outcome).

To estimate the effects of online use on the customer's product demand and number of transactions completed, we implement a

Gaussian kernel algorithm that matches each offline customer with a specific online customer and assigns a weight to each match that reflects the distance between members of the pair (for details, see the Appendix). This algorithm is effective for samples with large numbers of offline customers who possess characteristics similar to those of the online customer (Smith & Todd, 2005), as we find to be the case in our data.⁷

4.3. Assessment of the hybrid matching method

To test the robustness of the estimated effects, we use propensity score stratification and determine if the effects of online use vary across customers. We designate five strata using the estimated propensity scores; the strata differ with respect to the estimated likelihood of using the online channel (Dehejia & Wahba, 2002; Mithas & Krishnan, 2009). After estimating the effects of online use on the customer's product demand and number of transactions completed for each stratum, we apply an analysis of variance (ANOVA) to evaluate if the effects of online use differ across strata. We use each online customer's observed value on an outcome variable k and its corresponding counterfactual outcome; if we find significant differences across strata, heterogeneity exists across customers, and the average effects of online use measured across all customers would be biased.

We also assess the predictive performance of the hybrid matching method using information from the holdout sample. That is, we compare actual and predicted values for the customers in the holdout sample in terms of number and balance of checking accounts, number of transactions completed, and cost to serve. We focus on product demand for checking accounts because too few customers in our holdout sample use other products. To assess the predictive performance, we calculate the absolute percentage error (APE):

$$APE_k = \frac{\left| E_i (y_{i,k,t=0}^1 | d_i = 1) - [E_i (y_{i,k,t=-6}^0 | d_i = 1) + ATT_k] \right|}{E_i (y_{i,k,t=0}^0 | d_i = 1)} \quad \forall k \in K, \tag{5}$$

where $(E_i (y_{i,k,t=-6}^0 | d_i = 1) + ATT_k)$ represents the predicted value of variable k (e.g., balance of checking accounts) for the holdout sample, equal to the sum of the observed mean value for k in the holdout sample at $t = -6$ (i.e., when customers were still offline) and the estimated effect of online use on k . The actual value of k is the observed mean value in the holdout sample at $t = 0$ ($E_i (y_{i,k,t=0}^1 | d_i = 1)$).

With this information, we compare the hybrid matching results against the results of propensity score matching, OLS regression, and the IV and control function methods to evaluate predictive performance. For the IV and control function methods, we use customer age and length of relationship as instruments. Although customer age and length of relationship with the bank are important drivers of online use, we would expect both characteristics to be weak instrumental variables because they are correlated with the customer's product demand and number of transactions completed. As a result, their predictive performance is expected to be poor.

5. Results

Customers' online use first affects the customer's product demand and number of transactions completed, and then it affects customer revenue and cost to serve. We will therefore detail the effects of online use on product demand and number of transactions completed before

⁷ We also implemented the one-nearest and four-nearest neighbor algorithm and compared their predictive performance. The Gaussian kernel algorithm offers the highest predictive performance. These results are available from the first author on request. We discuss how we assess predictive performance in the subsequent section.

extending our discussion to the effects on customer revenue and cost to serve and to the moderating effect of product portfolios.

5.1. Effect of customers' online use on product demand and number of transactions completed

We employ Eq. (4) to assess the treatment effects and provide the results in Table 4. Online use has a significant positive effect on the number of checking accounts used by a customer ($ATT=0.01$, $p=0.02$).⁸ However, we find no significant effects of online use on the number of brokerage accounts ($ATT=0.00$, $p=0.24$) or on the number of credit cards ($ATT=0.00$, $p=0.68$) that the customer holds. The effects of online use on the customer's balances are not significant at the 5% level. We also find that online use affects the number of transactions completed positively ($ATT=2.50$, $p=0.00$): On average, online customers conduct 2.50 more transactions per month than do their offline counterparts.

If we ignore self-selection effects and simply compare product demand and number of transactions completed for online and offline customers, we find that online customers use significantly more checking accounts ($\Delta=0.18$, $p=0.00$) and credit cards ($\Delta=0.02$, $p=0.00$) but have fewer brokerage accounts ($\Delta=-0.02$, $p=0.00$) (Table 4).⁹ Moreover, online customers have significantly less money in their checking accounts ($\Delta=-201.38$ EUR, $p=0.02$) and savings accounts ($\Delta=-883.33$ EUR, $p=0.00$) than offline customers. Finally, online customers conduct more transactions with the bank ($\Delta=3.68$, $p=0.00$). The differences between the observed mean difference and ATT thus indicate substantial self-selection effects. For example, the 0.18 difference in number of checking accounts held by online versus offline customers appears to be caused mainly by self-selection. With Eq. (2), we find the self-selection effect to be 0.17 ($p=0.00$), while the effect of online use is only 0.01. The observed difference in means is thus driven by self-selection more than by online use.

5.2. Effect of customers' online use on customer revenue and cost to serve

To assess the effect of online use on customer revenue, we multiply the individual observed values of the online customer's product demand and estimated counterfactual outcome by the corresponding net contribution margins for each product. We sum these outputs across products to ascertain actual and estimated customer revenues. Finally, we compute the mean difference between actual and estimated customer revenues to derive the effect of online use (Eq. (4)).

The information on cost to serve is based on the number of transactions conducted in a particular channel multiplied by the channel-specific cost. We use the mean difference in actual and estimated costs to serve to derive the effect of online use.

In support of H1, online use improves customer revenue significantly ($ATT=0.18$, $p=0.03$). Online customers generate, on average, 0.18 EUR in additional revenue per month (Table 4). If we ignore self-selection effects, the difference in means is -0.99 EUR per month, implying there exists a *negative* self-selection effect of -1.17 EUR per month (Eq. (2)). The size and sign of the self-selection effect underscore the importance of controlling for such effects.

Cost to serve decreases by 0.12 EUR per month ($p=0.00$), as we predicted in H2. This result suggests that online customers substitute offline transactions with online transactions; the number of transactions completed increases by 2.50 transactions per month on average but cost to serve decreases. When comparing only the costs to serve of online and offline customers, we find a difference in means of

⁸ To evaluate whether the differences in means for the matched samples are significant, we use a paired sample *t*-test (Rosenbaum & Rubin, 1985).

⁹ The results are based on simple mean comparison tests using *t*-tests for independent samples (Rosenbaum & Rubin, 1985).

Table 4

Estimated treatment effects of using online banking on the customer's product demand, number of transactions completed, revenue, cost to serve, and profitability.

	Observed mean difference: online versus offline (<i>p</i> -value)	Treatment effect of using online channel: ATT (<i>p</i> -value)
Product demand		
Number of checking accounts	.18 (.00)	.01 (.02)
Number of brokerage accounts	-.02 (.00)	.00 (.24)
Number of credit cards	.02 (.00)	.00 (.68)
Balance checking accounts (EUR)	-201.38 (.02)	8.94 (.79)
Balance savings accounts (EUR)	-883.33 (.00)	-174.39 (.20)
Balance brokerage accounts (EUR)	-1056.83 (.28)	140.49 (.81)
Relative balance credit cards	0.68 (.31)	.59 (.49)
Number of transactions completed	3.68 (.00)	2.50 (.00)
Customer revenue (EUR)	-.99 (.00)	.18 (.03)
Costs to serve customer (EUR)	-.02 (.74)	-.12 (.00)
Customer profitability (EUR)	-1.12 (.00)	.28 (.00)

Sample size = 86,754; bold values are significant at the 5% level, according to two-sided *t*-tests for independent samples (observed mean difference) and paired samples (ATT). ATT = average treatment on the treated effect.

-0.02 , which is not significant ($p=0.74$). Again, controlling for self-selection effects is critical.

The overall effect of online use on customer profitability is positive (Table 4). Customer profitability improves by 0.28 EUR per month ($p=0.00$), 60% of which is due to the revenue effect. If we ignored self-selection effects, online customers would appear 1.12 EUR ($p=0.00$) less profitable per month (Table 4), with a *negative* self-selection effect of approximately -1.40 EUR per month. Once again, the size and sign of the self-selection effect underscore the importance of controlling for such effects.

5.3. Moderating effect of the customer's product portfolio

Theory suggests that the impact of customers' online use on revenue and cost to serve may vary depending on customers' product portfolio. We employ an ANOVA to test whether product portfolios moderate the effect of online use on customer revenue; in support of H3, they do (F -value = 7.60, $p=0.00$).¹⁰

In Table 5, we show the effects of online use on customer revenue across product portfolios, based on the portfolios actually used by online customers. The overall effect of online use on customer revenue reported in Table 4 is equal to the weighted average of the second column of Table 5 with the number of online customers using a product portfolio as weights. The overall effect of online use on cost to serve reported in Table 4 is equal to the weighted average of the fourth column of Table 5 with the number of online customers using a product portfolio as weights. The revenue effect is largest for a product portfolio that contains checking and brokerage accounts (2.84 EUR per month) – a surprising result, considering that brokerage accounts represent high-risk products (Wang, Keller, & Siegrist, 2011). The effect is significantly larger than for product portfolios consisting of the following: checking accounts alone (0.18 EUR per month, $p=0.00$); checking and savings accounts (-0.30 EUR per month, $p=0.00$); or the combination of checking accounts, savings accounts, and credit cards (-0.12 EUR per month, $p=0.00$). The second largest effect appears in the portfolio with checking accounts and credit cards (1.50 EUR per month). This effect is larger compared to the revenue effect on checking accounts only ($p=0.00$),

¹⁰ The ANOVA is based on online customers and considers portfolios only of products they actually used. We employ the estimated individual effect of online use on customer revenue (observed value – estimated counterfactual outcome) as the dependent variable and product portfolio as the independent variable. A similar approach indicates whether the product portfolio moderates the effect of online use on cost to serve.

Table 5
Moderating effect of customer's product portfolio on customer-revenue and cost-to-serve effects.

Portfolio	Estimated effect on customer revenue		Estimated effect on cost to serve	
	Mean	Standard error	Mean	Standard error
	Checking account	.18	.10	−.24
Checking + savings accounts	−.30	.17	.17	.09
Checking accounts + credit card	1.50	.33	−.29	.16
Checking + brokerage accounts	2.84	.73	−.62	.38
Checking + savings accounts + credit card	−.12	.48	.59	.23
Checking + savings + brokerage accounts	−.01	.77	−.02	.39

Sample size = 1627 (online customers).

checking and savings accounts ($p = 0.00$) and checking, savings accounts, and credit cards ($p = 0.08$).

Overall, the revenue effect of online use is less pronounced for customers who have savings accounts in their product portfolios; indeed, the effect can be negative. Savings accounts are also the product with the lowest frequency of use, confirming our prediction that the effect of online use is smaller for infrequently used products. In contrast, products that are used more frequently benefit from online use.

Again applying an ANOVA, we test for the moderating effect of a customer's product portfolio on cost to serve and find support for H4 (F -value = 5.82, $p = 0.00$). In Table 5, we report the mean values; the largest cost reductions occur for checking and brokerage accounts (−0.62 EUR per month), checking account and credit card (−0.29 EUR per month), and checking account only (−0.24 EUR per month). Although online use increases the number of transactions that occur in these portfolios, the costs to serve still decrease. It appears likely that customers substitute enough online transactions for offline transactions that the bank realizes significant cost savings.

Cost to serve is not reduced as much for customers with savings accounts. Online use actually increases cost to serve by 0.17 EUR per month when customers own checking and savings accounts and by 0.59 EUR per month when they additionally have credit cards. Customers with savings accounts substitute online transactions for offline transactions to a far lesser degree. Overall, customers with savings accounts in their product portfolio react less favorably to online use, and the substitution effect is significantly smaller for these customers than for customers without savings accounts in their product portfolios. Customers with savings accounts tend to make extensive use offline channels and seem hesitant about migrating to the online channel.

5.4. Further evaluation of the estimated effects

The effects of online use on product demand and number of transactions completed could vary across customers. In this case, the effects in our previous analyses would be biased if we failed to account for such heterogeneity. We thus employ propensity score stratification and an ANOVA to test whether the effects of online use differ across customers. Across strata, we find no significant differences between the estimated effects of online use on the customer's product demand or number of transactions completed. Overall, there is no substantial customer heterogeneity with respect to the effects of online use.

To compare the predictive performance of the hybrid matching method with that of propensity score matching, OLS regression, and the IV and control function methods (Table 6), we use information about online customers for the three-month period at $t = -6$ months and turn again to our holdout sample of 108 customers. As we expected, the IV and control function estimates are heavily

Table 6
Predictive performance of methods to control for self-selection effects (holdout sample).

Method	Average percentage error (APE, Eq. (5))				Mean APE
	Balance checking accounts	Number of checking accounts	Number of transactions completed	Cost to serve	
Hybrid matching method	4.92	5.70	24.59	18.61	13.45
Propensity score matching	3.76	5.07	27.62	23.43	14.97
OLS regression	6.13	7.05	22.24	29.13	16.14
Control function method (instrument: age)	14.37	422.04	110.34	86.25	158.25
Control function method (instrument: relationship length)	13.63	295.80	49.22	46.79	101.36
Instrumental variable (IV) (instrument: age)	27.35	799.79	.25	426.81	313.55
Instrumental variable (IV) (instrument: relationship length)	21.62	466.75	52.51	247.15	197.01

Size of holdout sample: 108 customers. We used other observed customer characteristics as instrumental variables for the control function and IV method, but these variables did not improve the performance of either method.

biased because we lack strong, valid instrumental variables.¹¹ The OLS regression (mean APE [MAPE] = 16.14%) performs much better than the IV (MAPE = 313.55% and 197.01%) or control function (MAPE = 158.25% and 101.36%) methods with weak instrumental variables. The control function method seems less affected by weak instrumental variables than the IV method, but its predictive performance is still very low.

Propensity score matching exhibits a MAPE of 14.97%. It performs very well with respect to the number and balance of checking accounts and presents the lowest APE for these two outcome variables. However, hybrid matching attains slightly better results with respect to predictive performance (MAPE = 13.45%).¹²

It is comforting to find that hybrid matching performs best in the holdout exercise. However, we acknowledge that we have only demonstrated adequate predictive performance. Further research is needed to examine the validity of the estimated effects.

6. Discussion and implications

This study deepens our understanding of the effects of customers' online use on customer revenue, cost to serve, and profitability as well as the moderating effect of the customer's product portfolio. Although we investigate a retail bank, the need to develop a strategy for online channels is not limited to this sector, and our study's insights have value in other settings as well.

We argue in particular that it is important to determine the revenue and cost effects of customers' online use when developing an online channel strategy. Previous studies have focused on either revenue or cost effects; we address both. Our results show that online use affects customer revenue positively, on average, as other authors have found (Kumar & Venkatesan, 2005; Thomas & Sullivan, 2005). Online use also decreases cost to serve, in contrast with the findings of Campbell and Frei (2010). The revenue effect of online use is approximately 50% greater than the effect on cost to serve (0.18 versus 0.12 EUR per month); in other words, increased revenue accounts for approximately 60% of the profitability effect. Managers may hope to increase customer profitability by reducing cost to serve, but our results indicate that moving customers online can also lift profitability by

¹¹ We also estimated the IV and control function models using both instrumental variables (age and length of relationship) at the same time, but the performance of the models declined.

¹² A detailed comparison of the results of hybrid matching versus propensity score matching indicated very similar estimated effects of online use. Only for the balance in the checking account do we find a significant positive effect of online use when propensity score matching is used and no significant effect with hybrid matching.

increasing customer revenue. Therefore, the online channel seems well suited for encouraging customer relationships, and managers need to develop effective online channel strategies to exploit these revenue benefits.

The effects of online use on customer revenue and cost to serve also vary across customer product portfolios, a finding that could provide a good basis for customer segmentation. Managers need differentiated strategies to motivate customers to use online channels, yet most evidence from the popular press suggests that firms use generic strategies that ignore differences in customer product portfolios. For example, Catherine Palmieri, managing director at Citibank, noted, “We’ve given away cash, we’ve had \$25,000 sweepstakes and other contests.... We do a lot of things in the branch where we get the sales force excited about getting the customers online” (Gurliacci, 2005).

Our results show, however, that customers with savings accounts are less attractive targets for moving online because their revenues are likely to decrease, and their costs to serve are likely to increase. Savings accounts are used relatively less frequently than other products in our study, and the frequency of use may be an important product characteristic for segmentation efforts. The convenience and accessibility of the online channel proves especially beneficial for frequently used products; managers should work to motivate customers to move to the online channel when those customers own frequently used products. On a firm level, firms that offer high-usage products may benefit more from actively managing customer channel use than do firms selling products that are used less frequently.

We also provide evidence of substantial self-selection effects. Ignoring these effects can lead to poor managerial decision making. For example, the observed mean difference for customer revenue is –0.99 EUR per month, but the estimated revenue effect of online use is .18 EUR per month, indicating a negative self-selection effect of –1.17 EUR per month. In other words, online customers appear to produce less revenue than offline customers. Yet online use improves customer revenue, so supporting the online channel can prove beneficial for the bank. Because self-selection effects account for a critical portion of observed differences in means between online and offline customers, managers must take these effects into account when assessing the impact of online use on customer metrics such as customer revenue and cost to serve.

Finally, we demonstrate an effective application of hybrid matching and propensity score matching that provided much higher predictive performance than the IV and control function methods. The latter two tactics seem inappropriate in the face of weak instrumental variables. Instead, matching methods are viable when controlling for self-selection effects, especially if only weak instrumental variables are available. These insights should transfer meaningfully to other situations in which managers attempt to quantify the effectiveness of their strategies and only weak instrumental variables are available.

7. Limitations and further research directions

This study suffers from several limitations that suggest avenues for further research. First, researchers should replicate our study with firms in other industries to investigate the generalizability of our findings, particularly our finding that the revenue effect of online use is substantially larger than the cost effect. Researchers also may wish to examine whether customer product portfolios provide actionable bases of segmentation when developing customer migration strategies in other settings, and which product characteristics are most important in determining whether moving customer to the online channel will increase profitability. Of further interest is whether product characteristics, such as frequency of use, moderate the effects of online use on customer revenue and cost to serve.

In addition, our empirical data set provides only limited information about individual customers. Additional characteristics that are

likely to be correlated with online use (e.g., attitudes toward Internet use) could be influential, although the good predictive performance of our results provides some support for their adequacy. Further studies could examine, in detail, which method of controlling for self-selection effects is most appropriate in certain settings.

A further limitation of our data is that they are cross-sectional; consequently, we cannot address the long-term effects of online use by examining the effects on customer retention. Some previous studies provide initial evidence that online use improves customer retention (e.g., Hitt & Frei, 2002; Shankar et al., 2003). Time-series data could reveal whether the effects of online use on the customer’s product demand and number of transactions completed change over time and whether the extent of the customer’s online channel use moderates the effects of online use. Investigating the dynamic effects of online use on customer behavior would be particularly interesting. Furthermore, it would be worthwhile to investigate competition and whether use of the online channel increases the firm’s share of the customer’s budget. If customers were routed to an online channel when buying or using a product, causality would be reversed. Using panel data and formal causality testing could address this issue. Reverse causality is not a problem in our study – the bank managers we consulted assured us that customers were not routed to the online channel – but it may be of concern in other applications.

Online marketing activities are also part of the incremental effect of online use on customer behavior. Additional research should disentangle the effects of online use and online marketing activities to provide more detailed insights in this realm.

Finally, moving customers across channels requires appropriate marketing instruments. An additional step for researchers would be to identify ways of encouraging customers to use online channels. Reinders, Dabholkar, and Frambach (2008), Ansari et al. (2008), and Thomas and Sullivan (2005) provide some insights, but more research is needed to reveal how firms can manage customers’ channel use actively and effectively.

Acknowledgment

We thank the E-Finance Lab at the House of Finance at Goethe-University and its partners for supporting this research.

Appendix A. Hybrid matching approach

We briefly outline the implementation of the hybrid matching method for estimating ATT using the Mahalanobis distance and a Gaussian kernel algorithm:

- 1 Randomly order online customers and offline customers.
- 2 For online customer i , determine similarity to all offline customers in the sample, according to the Mahalanobis distance (M), by considering the propensity score, customer’s age (AGE), and length of relationship (LOR).
- 3 Estimate the counterfactual outcome for every outcome variable for online customer i using a weighted average of the outcome variable of all offline customers. The weights assigned to each offline customer j depend on the similarity between online and offline customers, as determined by the Gaussian kernel algorithm. The bandwidth parameter τ is a function of the standard deviation σ of the similarity measure and sample size N , such that $\tau = 1.06 \cdot \sigma \cdot N^{-0.2}$ (Silverman, 1986).¹³ The weight of an offline customer j for an online customer i equals:

¹³ The standard deviation σ is the sample standard deviation after taking the common support restriction into account. Thus, it is estimated across all customers.

$$w_{ij} = \frac{K\left(\frac{M_{ij}}{\tau}\right)}{\sum_{j \in J} K\left(\frac{M_{ij}}{\tau}\right)} \quad \forall i \in I, j \in J, \quad (\text{A.1})$$

where $K(\cdot)$ is a normally distributed kernel function.

- 4 Remove online customer i from the list of online customers.
- 5 Does i equal $|I|$? If no, set i to $i + 1$ and go to Step 2. If yes, go to Step 6.
- 6 Compute the ATT for every outcome variable, using the average outcome for the online customers and the average estimated counterfactual outcome.

References

- Alba, J., Lynch, J., Weitz, B., Janiszewski, C., Lutz, R., Sawyer, A., et al. (1997). Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *Journal of Marketing*, 61(3), 35–53.
- Ansari, A., Mela, C., & Neslin, S. (2008). Customer channel migration. *Journal of Marketing Research*, 45(1), 60–76.
- Ariely, D. (2000). Controlling information flow: Effects on consumers' decision making and preference. *Journal of Consumer Research*, 27(2), 233–248.
- Bhatnagar, A., & Ghose, S. (2004). Online information search termination patterns across product categories and consumer demographics. *Journal of Retailing*, 80(3), 221–228.
- Bitner, M. J., Brown, S., & Meuter, M. (2000). Technology infusion in service encounters. *Journal of the Academy of Marketing Science*, 28(1), 138–149.
- Blattberg, R., Kim, B.-W., & Neslin, S. (2008). *Database marketing – Analyzing and managing customers*. New York: Springer Science + Business Media.
- Brynjolfsson, E., Hu, Y., & Smith, M. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science*, 49(11), 1580–1596.
- Campbell, D., & Frei, F. (2010). The cost structure, customer profitability, and retention implications of self-service distribution channels: Evidence from customer behavior in an online banking channel. *Management Science*, 56(1), 4–24.
- Chu, J., Chintagunta, P., & Cebollada, J. (2008). A comparison of within-household price sensitivity across online and offline channels. *Marketing Science*, 27(2), 283–299.
- Danaher, P., Wilson, L., & Davis, R. (2003). A comparison of online and offline consumer brand loyalty. *Marketing Science*, 22(4), 461–476.
- Degeratu, A., Rangaswamy, A., & Wu, J. (2000). Consumer choice behaviour in online and traditional supermarkets: The effects of brand name, price and other search attributes. *International Journal of Research in Marketing*, 17(1), 55–78.
- Dehejia, R. (2005). Practical propensity score matching: A reply to Smith and Todd. *Journal of Econometrics*, 125(1–2), 355–364.
- Dehejia, R., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1), 151–161.
- Forsythe, S., & Shi, B. (2003). Consumer patronage and risk perceptions in internet shopping. *Journal of Business Research*, 56(11), 867–875.
- 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.
- Gurliacci, D. (2005). Banks Create Incentives to Encourage Online Banking. Westchester County Business Journal, July 11, 29, available at: <http://www.allbusiness.com/banking-finance/banking-lending-credit-services-cash/10572721-1.html>, last update: September 1, 2011.
- Gutiérrez, S., Izquierdo, C., & Cabezedo, R. (2010). Product and channel-related risk and involvement in online contexts. *Electronic Commerce Research and Applications*, 9(3), 263–273.
- Hadi, A. S. (1994). A modification of a method for the detection of outliers in multivariate samples. *Journal of the Royal Statistical Society: Series B: Methodological*, 56(2), 393–396.
- Heckman, J., & Navarro-Lozano, S. (2004). Using matching, instrumental variables, and control functions to estimate economic choice models. *The Review of Economics and Statistics*, 86(1), 30–57.
- Heckman, J., Ichimura, H., & Todd, P. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4), 605–654.
- Hitt, L. M., & Frei, F. X. (2002). Do better customers utilize electronic distribution channels? The case of PC banking. *Management Science*, 48(6), 732–748.
- 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*, 73(2), 55–69.
- Inman, J., Shankar, V., & Ferraro, R. (2004). The roles of channel-category associations and geodemographics in channel patronage. *Journal of Marketing*, 68(2), 51–71.
- Kumar, V., & Venkatesan, R. (2005). Who are the multichannel shoppers and how do they perform? Correlates of multichannel shopping behavior. *Journal of Interactive Marketing*, 19(2), 44–62.
- Liang, T., & Huang, J. (1998). An empirical study on consumer acceptance of products in electronic markets: A transaction cost model. *Decision Support Systems*, 24(1), 29–43.
- Mithas, S., & Krishnan, M. S. (2009). From association to causation via a potential outcomes approach. *Information Systems Research*, 20(2), 295–313.
- Mithas, S., Krishnan, M. S., & Fornell, C. (2005). Why do customer relationship management applications affect customer satisfaction? *Journal of Marketing*, 69(4), 201–209.
- Montoya-Weiss, M., Voss, G., & Grewal, D. (2003). Determinants of online channel use and overall satisfaction with a relational, multichannel service provider. *Journal of the Academy of Marketing Science*, 31(4), 448–458.
- Neslin, S., & 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., & Huizingh, E. (2011). Does online information drive offline revenues? Only for specific products and consumer segments! *Journal of Retailing*, 87(1), 1–17.
- Puhani, P. (2000). The Heckman correction for sample selection and its critique. *Journal of Economic Surveys*, 14(1), 53–68.
- Reinders, M., Dabholkar, P., & Frambach, R. (2008). Consequences of forcing consumers to use technology-based self-service. *Journal of Service Research*, 11(2), 107–123.
- Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.
- Rosenbaum, P., & Rubin, D. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387), 516–524.
- Rosenbaum, P., & Rubin, D. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1), 33–38.
- Rosenbaum, P., Ross, R., & Silber, J. (2007). Minimum distance matched sampling with fine balance in an observational study of treatment for ovarian cancer. *Journal of the American Statistical Association*, 102(477), 75–83.
- Schlosser, A. (2003). Experiencing products in the virtual world: the role of goal and imagery in influencing attitudes versus purchase intentions. *Journal of Consumer Research*, 30(2), 184–198.
- Shankar, V., Smith, A. K., & Rangaswamy, A. (2003). Customer satisfaction and loyalty in online and offline environments. *International Journal of Research in Marketing*, 20(2), 153–175.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. New York: Chapman & Hall.
- Skiera, B., Bermes, M., & Horn, L. (2011). Customer equity sustainability ratio: A new metric for assessing a firm's future orientation. *Journal of Marketing*, 75(3), 118–131.
- Smith, D., & Sivakumar, K. (2004). Flow and internet shopping behavior—A conceptual model and research propositions. *Journal of Business Research*, 57(10), 1199–1208.
- Smith, J., & Todd, P. (2005). Does matching overcome Lalonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(1/2), 305–353.
- Stock, J., Wright, J., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics*, 20(4), 518–529.
- Strader, T., & Ramaswami, S. (2004). Investor perceptions of traditional and online channels. *Communications of the ACM*, 47(7), 73–76.
- Tedeschi, B. (2005). E-commerce report; to attract more internet customers, some banks are adding services available on their web sites to their A.T.M.'s. *The New York Times*, March 7, available at: <http://query.nytimes.com/gst/fullpage.html?res=9806E5DD1E3DF93A35750CA9639C8B63&sec=&spn=&pagewanted=all>, last update: August 17, 2011.
- Thomas, J., & Sullivan, U. (2005). Managing marketing communications with multichannel customers. *Journal of Marketing*, 69(4), 239–251.
- Venkatesan, R., Kumar, V., & Ravishanker, N. (2007). Multichannel shopping: causes and consequences. *Journal of Marketing*, 71(2), 114–132.
- Wang, M., Keller, C., & Siegrist, M. (2011). The less you know, the more you are afraid of: a survey on risk perceptions of investment products. *Journal of Behavioral Finance*, 12(1), 9–19.
- 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.
- Woglom, G. (2001). More results on the exact small sample properties of the instrumental variable estimator. *Econometrica*, 69(5), 1381–1389.
- Xue, M., Hitt, L., & Harker, P. (2007). Customer efficiency, channel usage, and firm performance in retail banking. *Manufacturing & Service Operations Management*, 9(4), 535–558.
- Xue, M., Hitt, L., & Chen, P. (2011). The determinants and outcomes of internet banking adoption. *Management Science*, 57(2), 291–307.
- Zhao, Z. (2004). Using matching to estimate treatment effects: Data requirements, matching metrics, and Monte Carlo evidence. *The Review of Economics and Statistics*, 86(1), 91–107.