university of groningen
faculty of economics and business

2021010-MARK

When Offline Stores Reduce Online Returns

June 2021

Christian F. Hirche
Tammo H.A. Bijmolt
Maarten J. Gijsenberg

SOM is the research institute of the Faculty of Economics & Business at the University of Groningen. SOM has six programmes:
- Economics, Econometrics and Finance
- Global Economics & Management
- Innovation & Organization
- Marketing
- Operations Management & Operations Research
- Organizational Behaviour

Research Institute SOM
Faculty of Economics & Business
University of Groningen

Visiting address:
Nettelbosje 2
9747 AE Groningen
The Netherlands

Postal address:
P.O. Box 800
9700 AV Groningen
The Netherlands

T +31 50 363 9090/7068/3815

www.rug.nl/feb/research

When Offline Stores Reduce Online Returns

Christian F. Hirche
University of Groningen, Faculty of Economics and Business, Department of Marketing

Tammo H.A. Bijmolt
University of Groningen, Faculty of Economics and Business, Department of Marketing

Maarten J. Gijsenberg
University of Groningen, Faculty of Economics and Business, Department of Marketing
m.j.gijsenberg@rug.nl

When Offline Stores Reduce Online Returns

Christian F. Hirche[a]

Tammo H.A Bijmolt[b]

Maarten J. Gijsenberg[c]

---

[a] Department of Marketing – University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. C.F.Hirche@rug.nl

[b] Department of Marketing – University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. T.H.A.Bijmolt@rug.nl

[c] Department of Marketing – University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. M.J.Gijsenberg@rug.nl
(corresponding author)

Abstract

Online shopping continues to grow in popularity, entailing growth in product purchases, but also in product returns. At the same time, offline stores remain present and might exert an influence on customers' online purchase and return behavior. To optimize profitability and allocate resources adequately, managers need to know whether and how offline stores influence online shopping behavior of customers. In an empirical study, the authors address this issue using data from a large Dutch shoe retailer and a database of the shoe retailer's store locations. They use a spatial model to estimate the influence of proximate retail stores on customers' online shopping behavior, while controlling for spatial and customer heterogeneity. Overall online shopping behavior is decomposed into the number of purchased shopping baskets and the number of purchased and returned products per shopping basket. For the latter, distinction is made between products of different price levels, discounted versus undiscounted, and uni-size versus multi-size products. Results show that the retailer's offline stores do not significantly affect online product purchases. However, they can decrease the number of returned products per shopping basket, depending on product characteristics. While returns of lower-risk products (i.e., uni-size, lower-price, or discounted) are not affected, returns of higher-risk products (i.e., multi-size, higher-price, or undiscounted) are decreased with as much as 13.91% to 16.26 %. This study thus shows that it is crucial to both consider the advantageous role of offline stores when investigating online purchase and return behavior, and to take between-product variation into account.

Keywords: Multichannel customer behavior; Customer purchase behavior; Online retailing; Product returns; Spatial regression

Declarations

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest/Competing interests

There are no conflicts of interest associated with the paper.

Availability of data and material

The used data is proprietary data and is not publicly available.

Code availability

The used software (R, Stan) is publicly available. The custom code for the analyses is not available.

Authors' contributions

All authors have contributed to the creation of this manuscript.

Ethics approval

Not applicable

Consent to participate

All authors have consented to participate in the creation of the manuscript.

Consent for publication

All authors have approved the manuscript and consent with its submission for publication.

1. Introduction

In recent years, the retail industry has undergone a profound transformation. Driven by customers, who increasingly shop online, retailers have been expanding to the Internet. For the traditional offline retail giant Wal-Mart, e-commerce has become a strategic pillar and the websites of once traditional offline retailers H&M, Asos, and Zara are now among the most-visited e-commerce sites (Deloitte, 2018; Zaczkiewicz, 2018). At the same time, former online-only retailers have embraced an omni-channel strategy and have begun to open or cooperate with offline stores. Examples are Warby Parker, Bonobos (Satell, 2017; Soper, 2018) and, most notably, Amazon, with a growing array of own physical stores in different product categories (e.g. “Amazon Books” bookstores, “Amazon Go” convenience stores, “Whole Foods” stores) and its cooperation with Kohl’s to allow for in-store returns at Kohl’s department stores (Chen, 2019). Ultimately, retailers and “e-tailers” might converge into a common, omni-channel type of business model (CB Insights, 2018; Thompson, 2017).

One crucial aspect in this transition, to both retailers and e-tailers, is the position of physical, offline, stores in their omni-channel portfolio. On the one hand, traditional offline retailers are under rising cost pressure and many physical retail stores have been closed in the past, a development for which the term “retail apocalypse” has been coined (Taylor, 2019). On the other hand, the opening of and cooperation with physical stores by former online-only retailers poses the question, which benefit lies in physical retail stores. Benefits might not only lie in the actual revenue incurred by the offline stores, but also in changing customer behavior with regards to the online channel.

One of the aspects in which traditional retailers who have expanded to the internet in the past differ from e-tailers, is the fact that they already have a strong physical offline presence. Question then becomes whether this offline presence constitutes a competitive advantage relative to the e-tailers with regard to online sales, or whether it is merely a cost that should be reduced: would opening or closing stores harm or help the *online* business of these traditional retailers?

In this paper, we investigate what effects physical stores of traditional offline retailers have on the customer behavior in their online channel. Online customer behavior pertains to two critical determinants of profitability: (1) the number of product purchases, and (2) the number of product returns (Minnema et al., 2018). In other words, what would happen to purchases and returns if store presence would be increased or decreased? Regarding the former, there might be negative cannibalization effects between the online and offline channel but also positive signaling effects. Regarding the latter, stores might increase returns by making returns easier and cheaper, but they might also decrease returns by allowing customers to ascertain a satisfactory product fit before ordering. Both might be differently affected by offline stores, depending on the performance risk associated with the product. Nearby offline stores might change customer purchase or return behavior only for products which need personal fitting and not products which do not, or only for cheaper products and not for more expensive products.

Prior research has generally focused on channel additions, i.e., extending an existing offline channel by an online channel (e.g., Li, 2021; van Nierop et al., 2011; Weltevreden, 2007) or extending an existing online channel by an offline channel (Avery et al., 2012; Pauwels & Neslin, 2015; Wang & Goldfarb, 2017) and found negative (e.g., van Nierop et al., 2011), neutral (e.g., Pauwels & Neslin, 2015), as well as positive (e.g., Avery et al., 2012) channel interactions. First, we aim to explain diverging findings by integrating risk-related product characteristics into the analysis and by investigating whether such characteristics change the influence of the offline channel on the online channel. Second, we propose and use a novel model specification, a hierarchical Spatial Durbin Error Poisson model, that allows us to investigate the effects of store *presence* in all national regions instead of store *openings* in retailer-selected regions, circumventing the issue of regional selection bias and possibly observing mainly short-term effects. In sum, we contribute to prior literature by:

- investigating the influence of offline store presence on online purchases – in terms of how often and how much at a time customers order – as well as returns,
- integrating how risk-related product characteristics moderate the influence of offline stores presence, and
- proposing a novel spatial model specification to analyze the effect, avoiding regional selection bias and taking account of the spatial structure of the data.

In our analysis, we use data of 10,000 online customers of a large retailer in the fashion industry, offering a range of products with different characteristics with regard to price and personal fit, over a period of 6 years. Data from this industry are particularly suited, as it is the consumer industry in which e-commerce is most prevalent, with, for example, two thirds of customers in the European Union shopping online for such products (Eurostat, 2018). As the percentage of e-commerce steadily grows in other product categories, issues that play a role in the fashion sector today might extend to other categories in the future. In our model section, we develop an extended version of the Spatial Durbin Error Model. This allows us to model the effect of stores on customer behavior while taking into account of customer heterogeneity, regional heterogeneity and the (extra-)regional influence of stores, i.e., store influence as a function of distance to the customer.

Results of our analyses show that the offline channel does not harm the sales in the online channel. Moreover, depending on the product characteristics, it helps the retailer by reducing product returns. A simulation study shows that reductions in product returns due to increased offline store availability can be as high as 13.91% to 16.26 %, depending on the type of product.

2. Theoretical Background

Online and the offline channels compete with each other along the whole customer journey, i.e. during pre-purchase, purchase and post-purchase (Verhoef et al., 2015). During the customer journey, customers can interact with multiple channels (Bijmolt et al., 2019). For example, many customers visit stores – pre-purchase – and subsequently purchase online (so-called show-rooming), or visit a website – pre-purchase – and subsequently purchase the product in a store (so-called web-rooming). However, channel choices are dependent of each other as channels "affect one another because of lock-in effects, channel inertia, and cross-channel synergies" (Lemon & Verhoef, 2016). Therefore, the availability of one channel influences customer behavior across the whole customer journey and across channels. Concerning the consequences of offline channel availability on purchases and returns in the online channel, there might be multiple, competing processes, which we will describe in this section.

2.1. Purchases

There are two mechanisms of how the availability of one channel can influence purchases in another channel, and this both in terms of the number of ordered baskets as in terms of number of products per basket. The first mechanism consists of channel presence leading to channel substitution. The second mechanism, in turn, consists of channel presence having a global promotional effect. Both point to a different effect direction, resulting in two opposing hypotheses.

Customers choose channels based on characteristics of the channels, and the ability to fulfil customer preferences. Rohm and Swaminathan (2004) identified six motives for shopping: shopping convenience, information seeking, immediate possession, social interaction, the retail shopping experience and variety seeking. Different shopping channels are suited to a different extent to comply with these motives: Online shopping may score higher on convenience and variety seeking while offline shopping wins on immediate possession and the retail shopping experience (King, 2018; Rekuc, 2018). The ranking of the motives and thereby the ranking of the channels depends on personal preference. For example, variety seeking customers might prefer shopping online and social interaction-favoring customers might prefer going to a retail store (Neslin & Shankar, 2009). When their preferred channel (e.g., offline) is not available, customers need to choose the next-best option: forgoing shopping or shopping at another channel (e.g., online). In this case, online channel purchases will be higher due to the absence of the offline channel. The opposite holds when the preferred channel *is* available. Then, customers do not need to make compromises. As a consequence, the offline channel will get purchases from customers who otherwise would have chosen to shop at the online channel or abstain. Online channel purchases are therefore expected to be lower in presence of the offline channel. Since this could translate to both less frequent orders, i.e., less ordered shopping baskets, and/or smaller orders, i.e., less purchased products per shopping basket, we hypothesize the following:

H1a: Offline channel presence decreases the number of ordered shopping baskets.

H2a: Offline channel presence decreases the number of purchased products per shopping basket.

On the other hand, stores can also have a promoting function (Mitra & Fay, 2010) due to their mere presence, and can thus convert non-shoppers to shoppers. Customers can be primed in the need recognition phase by seeing stores of a brand to visit the web shop of this brand. Moreover, customers can also change channels in a later stage, as is exemplified by the phenomenon of show-rooming. Here, customers deliberately visit a store to choose a product without purchasing it. Instead, they change channels for the actual purchase. In general, the existence of the offline channel can thus draw customers to make purchases in the online channel. Again, this might result in both more frequent orders, i.e., more ordered shopping baskets, and/or larger orders, i.e., more purchased products per shopping basket. Therefore, we hypothesize the following:

H1b: Offline channel presence increases the number of ordered shopping baskets.

H2b: Offline channel presence increases the number of purchased products per shopping basket.

2.2. Returns

The presence of an offline channel can also influence the number of product returns. As before, we focus on the effect on the online channel, i.e., products returned that were purchased in the online channel. Again, two mechanisms with opposite directionality are at play, resulting in two opposing hypotheses.

On one side, the offline channel might decrease returns due to customers' reaction to perceived risk (Kushwaha & Shankar, 2013). According to Mitchell (1992), social risk, financial risk, physical risk, performance risk, time risk and psychological risk can influence consumer behavior. For fashion products – the type of products that we focus on – performance risk is pronounced (Gu & Tayi, 2015). For such products, performance is to a great extent determined by product fit, i.e., the match between customer preferences and product properties (Anderson et al., 2009). The amount of uncertainty concerning product fit depends on channel choice, as offline stores allow for much better inspection of physical products than an online website. Customers can anticipate this increased uncertainty and the resulting increased likelihood of returning and adapt their purchase behavior accordingly (Janakiraman & Ordóñez, 2012). In particular, when offline stores are available, customers can shift the first stages of their decision-making process to offline stores. That is, they would engage in show-rooming to reduce fit uncertainty. As a result, subsequent returns are likely lower. In addition, customers can shift the whole customer journey to the stores, becoming "one-stop shoppers" in the offline channel (Neslin & Shankar, 2009). Regarding the return rate, results would be the same: customers would return less in the online channel. Therefore, we hypothesize the following:

H3a: Offline channel presence decreases the number of returned products per shopping basket.

On the other side, returns might increase due to decreased "cost" of returning. It has been shown that the effort it takes to return a product influences the likelihood that customers actually return a product (Kim & Wansink, 2012). The decision to return a product is a result of a cost/benefit calculation. Even though a product might not ideally fit the customer's preferences, returning the product is associated with a cost. The cost of returning the product must be lower than the shortcoming in utility of the product in order for a return to be worthwhile (Anderson et al., 2009; Petersen & Kumar, 2015). When there is no offline channel available, customers must re-package the product in a parcel, label it, drive to a post office and, sometimes, pay shipping fees. In addition, a refund will usually be offered only after the firm received the product back and checked its condition. On the other hand, when stores *are* available, customers can simply return the product there, thereby avoiding shipping fees, investment in time and effort, and packaging. In consequence, returning costs for the customer are lower when the offline channel is available. Therefore, the cost/benefit calculation is likely more often in favor of returning and, thus, customers might also return more, leading to the following hypothesis:

H3b: Offline channel presence increases the number of returned products per shopping basket.

2.3. The Role of Risk

Purchasing a product is generally associated with risk. As argued above, fashion products suffer from performance risk due to uncertainty regarding the fit between product properties and customer preferences (Anderson et al., 2009; Gu & Tayi, 2015). While all products have some risk, the relative amount of risk differs, however.

In particular, whether or not a product fits, i.e., its fit uncertainty, depends on the type of product: A first type of fashion products is highly dependent on the customer's physiology. Such products are generally offered in multiple sizes (e.g., XS, M, XXL) and, by consequence, pre-purchase fit uncertainty is higher. Conversely, a second type of fashion products is mostly independent of the customer's physiology. Such products are therefore offered only in one size (e.g., bags and watches) and, by consequence, pre-purchase fit uncertainty is lower. Despite technological advances in the online channel (e.g., zoomable products images, Minnema et al., 2016), customers still need to visit a store and check their personal physiology in combination with the product to eliminate fit uncertainty. When offline stores are available, customers can therefore reduce fit uncertainty by visiting a store before ordering online or instead of ordering online. When customers visit a store before ordering online, fit uncertainty and thereby returns are likely reduced and when customers visit a store instead of ordering online, both purchases and returns are likely reduced. In sum, we expect offline store availability to have a negative effect on purchases and returns for the online channel for multi-size products. For purchases and returns of uni-size products, we expect effects to be limited or even absent.

In addition, performance risk of a product also depends on customers' expectations. Customers have higher quality expectations for products that are relatively more expensive (Grewal, 1995). For products that are perceived as less expensive (i.e., due to a low price relative to competitors or having a price discount), customers might be more willing to accept slight deficits in the product. Consequently, perceived performance risk is lower for such products. Therefore, customers will feel less the need to reduce the risk by interacting with the offline channel instead of or before interacting with online channel. In consequence, we expect availability of offline stores to have a negative effect on both purchases and returns in the online channel for undiscounted and relatively higher-priced products. For purchases and returns of discounted and relatively lower-priced products, we expect effects to be limited or even absent. In sum, we expect risk level with regards to fit uncertainty, price level and discount to moderate the effect of offline channel presence on both purchases and returns (Figure 1). Therefore, we hypothesize that:

H4a-b-c: The impact of offline channel presence on the *number of purchased products per shopping basket* is more negative/less positive for (a) products with higher fit uncertainty (i.e., multi-size products), (b) higher-priced products, and (c) undiscounted products compared to products with lower fit uncertainty, lower-priced products, and discounted products, respectively.

H5a-b-c: The impact of offline channel presence on the *number of returned products per shopping basket* is more negative/less positive for (a) products with higher fit uncertainty (i.e., multi-size products), (b) higher-priced products, and (c) undiscounted products compared to products with lower fit uncertainty, lower-priced products, and discounted products, respectively.

Fig. 1 Research framework

3. Data

3.1. Sample

To assess the impact of offline channel availability on the customer purchase and return behavior in the online channel, we conduct an empirical analysis with data from a large Dutch retailer in the fashion segment that sells shoes, bags, belts, socks, wallets and similar products. We have a sample of 10,000 online customers who ordered at least one product during a period from September, 8th, 2008 to October, 5th, 2014. In total, customers ordered more than 20,000 products (in about 15,000 shopping trips or baskets) and returned more than 3,500 products in this period. Generally, for each product order, we have detailed information on the product that was ordered and the customer who ordered it. In a small number of cases, product or customer information is missing and we had to remove the associated orders from the dataset (see section: Invalid Cases and Missing Variables).

Using the zip-code of the customer addresses, we can situate each customer in small, contiguous zip-code region. This allows us to examine the influence of the retailer's offline stores on customers' online shopping behavior while taking proximity into account. The number of stores is rather stable over time, ranging from 405 in 2005 to 430 in 2017, with a peak of 460 in 2012. We gather the location of the retailer's stores from a database of store locations (*schoenenwinkelsoverzicht.nl*, assessed in June, 2017).

Using the store location data, we calculate the count of stores within each zip-code region plus the count of stores beyond that region and apply spatial decay for the latter (we provide more information on the so-called spatial decay in the Model section) and use these as our explanatory variables. The average size of a region is 10.900 km^2 (sd = 20.276 km^2).

3.2. Variables

For our analysis, we use the following dependent variables: the number of ordered shopping baskets per customer, the number of purchased products per shopping basket, and the number of returned products per shopping basket. Decomposing the total number of purchases and returns into ordered shopping baskets and purchases/returns per shopping baskets has two benefits: First, it allows to separately assess (1) how many times customers order, and (2) how much they purchase and return per order. Second, it allows to add order-level control variables to the analysis of (2).

We expect store availability to have a different influence for products of different price levels, discounts and sizes. Thus, we separately analyze purchases and returns of three pairs of complementing subsets of products, namely the subsets of products with either (1) higher or lower fit uncertainty, (2) with higher or lower price, and (3) without or with a discount. For distinguishing higher- from lower-priced products, we use a within-category median split of undiscounted product prices. Using a within-category split instead of an across-category spit is based on the notion that cheap categories can still contain relatively expensive – and thus risky – products and vice versa. In total, we thus have seven sets of products: one set with all products, and three times two complementary sets. For each set, we analyze purchases and returns, resulting in 14 outcome variables (see Table 1).

Table 1 Product purchase and return variables

K	Variables	Description
k=0	$purchases_{i,c,r,0}$ $returns_{i,c,r,0}$	Number of overall purchased/returned products[a]
k=1	$purchases_{i,c,r,1}$ $returns_{i,c,r,1}$	Number of purchased/returned multi-size products[a]
k=2	$purchases_{i,c,r,2}$ $returns_{i,c,r,2}$	Number of purchased/returned uni-size products[a]
k=3	$purchases_{i,c,r,3}$ $returns_{i,c,r,3}$	Number of purchased/returned products priced above the median price in their category[a]
k=4	$purchases_{i,c,r,4}$ $returns_{i,c,r,4}$	Number of purchased/returned products priced below the median price in their category[a]
k=5	$purchases_{i,c,r,5}$ $returns_{i,c,r,5}$	Number of purchased/returned products without discount[a]
k=6	$purchases_{i,c,r,6}$ $returns_{i,c,r,6}$	Number of purchased/returned products with discount[a]

[a]in shopping basket i of customer c in region r

As control variables, we use demographic information, month and year of the order and additional regional information like welfare, urbanization, car ownership, and post office locations. The latter originate from Statistics Netherlands (CBS), the Netherlands Institute for Social Research (SCP) and the Dutch post (PostNL), respectively. For post offices, we reverse geo-coded their geographical coordinates to postal code regions using the ArcGIS services API (ESRI, 2016) and included a count with spatial decay as a control variable, similar to our approach of including retail store locations.

3.3. Invalid Cases and Missing Variables

In some cases, the data contained incomplete or unusable information on ordered products or customers. First, we removed the product orders for which product characteristics were missing as we cannot estimate the influence of product characteristic-based moderators. Second, we removed product orders, where customers could not be assigned to a region or regional statistical information was not available. In total, we removed 57 full customers (from 10,000 to 9,943, i.e., 0.57%), 705 entire shopping baskets (from 15,122 to 14,426, i.e., 4.60%), and 1171 specific product purchases (from 21,360 to 20,189, i.e., 5.48%).

Data on gender and age is only available for a portion of the data (percentage observed: 49.3% for gender, 46.3% for age). Deleting all cases with missing data would dramatically shrink the sample size and hence statistical power (Graham, 2009). Therefore, for age, we add an additional dummy variable to indicate missingness and set the missing data value itself to the mean of non-missing data. For gender, we only add an additional dummy variable to indicate missingness.

3.4. Descriptive Overview

In Table 2, we present a description of all variables, that we include in our analysis. In total, we observe 14,426 ordered shopping baskets by 9,943 customers. A customer orders on average 1.451 shopping baskets (sd = 1.044, range 1-18). A shopping basket contains on average 1.399 purchased products (sd = 0.826, range 1-17) and 0.251 returned products (sd = 0.649, range 0-10). Customer identified predominantly as female (86.63% of the observed gender) and were 40.524 years on average (sd = 11.314). The average number of stores of our retailer per region is 0.121 (sd = 0.594, range 0-12).

Table 2 Variables used in the purchase and returns models

Variable	Description
Customer behavior	
$baskets_{c,r}$	Number of ordered shopping baskets by customer c in region r
$purchases_{i,c,r,k}$	Number of purchased products in shopping basket i of customer c in region r with product characteristic k
$returns_{i,c,r,k}$	Number of returned products in shopping basket i of customer c in region r with product characteristic k
Regional characteristics	
$(W\ x\ stores)_r$	Number of stores in region r and its surrounding regions with spatial decay
$(W\ x\ posts)_r$	Number of post offices in region r and its surrounding regions with spatial decay
$welfare_r$	welfare indicator[a] for region r
$urbanization_r$	urbanization indicator[a] for region r
$cars_r$	cars per capita[a] in region r
Order characteristics	
$year_i$	Year in which the order i was made (0=2008, …, 6=2014)
$month_{2,i}$, …, $month_{12,i}$	11 indicators for the month in which order i was made (baseline: January)
Customer characteristics	
age_c	Age of customer c in years[a]
$gender_c$	Customer c is female (1) or male (0)

[a]mean-centred

4. Methodology

We employ three hierarchical spatial regression models to estimate the effects of store availability. First, we estimate a model for the number of shopping baskets ordered by a customer over the period of observation, dependent on the number of stores in proximity – in addition to several other control variables. We account for distance by pre-multiplying a weighting matrix W to the vectors *stores* and *posts.* This operation adds to the two variables the spatially decayed (i.e., diminished) values of neighboring regions. By that, a higher number of stores in one region increases also the number of stores in proximate regions, but to a lesser extent the higher the distance between both regions. Furthermore, we add a regional error term for unobserved regional heterogeneity. Since proximate regions are more similar to each other than distant regions, we permit spatial auto-correlation in the regional error term. In the subsection "Spatial Durbin Error Model", we describe the spatial aspect of our model in more detail. We assume a zero-truncated Poisson distribution for the dependent variable, because all customers in the sample order at least once. The resulting, first model is shown in Equation (1):

$$baskets_{c,r} \sim ztPoisson\left(exp\left(\mu_{baskets_{c,r}}\right)\right)$$
$$\mu_{baskets_{c,r}} = \alpha_1 + \beta_1 age_c + \beta_2 gender_c$$
$$+\theta_1 welfare_r + \theta_2 urbanisation_r + \theta_3 cars_r \quad (1)$$
$$+\gamma_1 (W \times stores)_r + \gamma_2 (W \times posts)_r$$
$$+u_r$$
$$u_r = (I_J - \lambda_k W_u)^{-1}\varepsilon_\mathrm{r}, \varepsilon_\mathrm{r} \sim N(0, \tau_k)$$

Second, we estimate a model for the number of purchased products per shopping basket, as shown in (2). We estimate the same model for all $purchases_{i,c,r,k}$ variables shown in Table 1. For $purchases_{i,c,r,0}$ we assume a zero-truncated Poisson distribution since all baskets contain at least one product of any sort and for the remaining orders variables we assume a regular Poisson distribution.

$$purchases_{i,c,r,0} \sim ztPoisson\left(exp\left(\mu_{purchases_{i,c,r,0}}\right)\right),$$
$$purchases_{i,c,r,k} \sim Poisson\left(exp\left(\mu_{purchases_{i,c,r,k}}\right)\right) \ \forall k \in \{2..7\},$$
$$\mu_{purchases_{i,c,r,k}} = \alpha_{1,k} + \beta_{1,k} year_i + \sum_{m=2}^{12} \beta_{2.m,k} month_{m,i}$$
$$+\beta_{3,k} age_c + \beta_{4,k} gender_c$$
$$+\theta_{1,k} welfare_r + \theta_{2,k} urbanisation_r + \theta_{3,k} cars_r \quad (2)$$
$$+\gamma_{1,k} (W \times stores)_r + \gamma_{2,k} (W \times posts)_r$$
$$+\varepsilon_{c,k} + u_{r,k},$$
$$\varepsilon_\mathrm{c,k} \sim N(0, \sigma_\mathrm{k}), \quad u_{r,k} = (I_J - \lambda_k W_u)^{-1}\varepsilon_\mathrm{r,k}, \varepsilon_\mathrm{r,k} \sim N(0, \tau_k).$$

Third, we estimate a model for the number of returned products per shopping basket, as shown in Equation (3). We estimate the same model separately for all $returns_{i,k}$ variables shown in Table 1 and we always assume a regular Poisson distribution for the dependent variable. Similar to before, we control for regional and customer heterogeneity as well as year and month:

$$returns_{i,k} \sim Poisson(\exp\left(\mu_{returns_{i,k}}\right))$$
$$\mu_{returns_{i,k}} = \alpha_\mathrm{1,k} + \beta_\mathrm{1,k} purchases_i$$
$$+\beta_\mathrm{2,k} year_i + \sum_{m=2}^{12} \beta_{3.m,k} month_{m,i}$$
$$+\beta_\mathrm{4,k} age_c + \beta_\mathrm{5,k} gender_i \quad (3)$$
$$+\theta_\mathrm{1,k} welfare_r + \theta_\mathrm{2,k} urbanisation_r + \theta_\mathrm{3,k} cars_r$$
$$+\gamma_\mathrm{1,k} (W \times stores)_r + \gamma_\mathrm{2,k} (W \times posts)_r$$
$$+\varepsilon_{c,k} + u_{r,k},$$
$$\varepsilon_\mathrm{c,k} \sim N(0, \sigma_\mathrm{k}), \quad u_{r,k} = (I_J - \lambda_k W_u)^{-1}\varepsilon_\mathrm{r,k}, \varepsilon_\mathrm{r,k} \sim N(0, \tau_k).$$

In sum, we use separate models to identify whether a customer (a) orders more often (i.e., more shopping baskets), (b) purchases more products at a time (i.e., more products per shopping basket), and/or (c) returns more products per shopping basket. This allows us to assess consumer behavior in more detail than only estimating a model of total product purchases/returns and conflating order frequency and quantity.

4.1. Spatial Durbin Error Model

In our models, we use both spatial predictors and a spatial error term. The use of spatial predictors follows the idea that "responses by individuals are [...] correlated in such a manner that individuals near one another in the space generate similar outcomes" (Bradlow et al., 2005, p. 268). More formally, an observed outcome of unit *A* does not only depend on characteristics of entity *A*, but also on characteristics of neighboring entity *B*, *C* and so on. This is called *spatial dependence*. Spatial models reflect that process by incorporating the influence of neighboring entities into the regression equation, weighted by a metric of their distances. The use of a spatial error term follows the idea that spatially situated observations are influenced more or less by the same unobservable factors (e.g., infrastructure, culture), depending on their distance to each other. This would normally lead to heteroskedasticity in the error term, but can be dealt with by a spatial model by explicitly modelling autocorrelated errors. Using both spatial predictors and a spatial error term is termed the spatial Durbin error model (Halleck Vega & Elhorst, 2015). In its most basic form, it looks as follows:

$$
\begin{aligned}
& \mathrm{Y} = X_A \beta + \mathrm{W} X_B \theta + \mathrm{u}, \# \\
\text{with: } & u = \lambda W u + \varepsilon, \\
& \varepsilon \sim N(0, \sigma^2 I_n)
\end{aligned}
$$

The dependent variable Y_i of observation i is influenced by the independent variables in X_A belonging to the same observation i and a linear combination of independent variables in X_B of possibly all *other* observations, through pre-multiplying the data vector or matrix with the square weighting matrix W. How neighbors influence each other is thus defined by the matrix W. The weighting matrix W is also used to model the autocorrelation between error terms ($u = \lambda W u$).

We modify this model to estimate both own-region and cross-region effects with θ by including the diagonal in W. Furthermore, we restrict the parameter range of λ to a maximum of one by scaling the matrix W used in the formula to calculate *u* (LeSage & Pace, 2009). The resulting matrices are

$$
W \text{ with } W_{i,j} = \begin{cases} e^{-\frac{2}{10} D_{i,j}}, & e^{-\frac{2}{10} D_{i,j}} > 10\%, \\ 0, & otherwise \end{cases}
$$

$$
W_u = \frac{1}{v - 1} (W - I)
$$

with $D_{i,j}$ being the distance between i and j and v the largest eigenvalue of W. We assume an exponential decay with a cut-off at an influence of less than 10%, which corresponds to $D_{i,j} > 11.5km$. A cut-off value is necessary to generate a sparser matrix and make computation more efficient. For $D_{i,j}$, we use the linear distance between the centroids of each zip-code region.

The spatial Durbin error model is the preferred spatial model when modelling local spillover effects (Elhorst, 2014; LeSage, 2014). However, the model shown above has some limitations we need to address. First, the model is non-hierarchical which makes it unsuitable for most marketing applications. We therefore add a second level to the model, thereby building on the work by Lacombe and McIntyre (2016). Such second hierarchy level allows to express the nesting of multiple customers within one spatial region. As shown above, we coalesce regional effect within- and beyond-region effects so that we have one parameter estimate per regional influential factor. Second, to be able to estimate effects on discrete (instead of continuous) outcomes, we use a Poisson (or zero-truncated Poisson) distribution instead of a Normal distribution. Since multivariate count distributions have a range of issues with regards to theoretical feasibility, computability and interpretability – with no generally agreed-upon solution (see e.g., Inouye et al., 2017), we model the influences on each dependent variable separately using a (zero-truncated) univariate Poisson distribution for each of the resulting hierarchical Spatial Durbin Error Poisson models.

4.2. Estimation

We use a Bayesian approach to obtain the estimates for the models. In particular, we use the No-U-Turn Markov Chain Monte Carlo sampling (NUTS) provided by the Stan software package (Stan Development Team, 2017). We use uninformative (uniform) priors for all parameters. The error structure of the model requires to compute the inverse of a function of the spatial weighting matrix W at each sampling step. The high number of regions and the resulting high dimensionality of the spatial weight matrix, however, renders this computationally unfeasible. Instead, we approximate the inverse by a Neumann series. When estimating, all relevant parameters of all models converged, with their so-called potential scale reduction factor (PSRF or $\hat{R}$) being smaller than 1.1. We also tested the model by simulating data and comparing the true to the estimated parameters. Here as well, results were satisfying. More details on simulation, model estimation and convergence can be found in the Appendix.

5. Results

We first report the influence of offline stores on the number of shopping baskets customers order online. Then, we present the influence of stores on the number of purchased and returned products per shopping basket, when including all products. Finally, we report the influence of stores on purchases and returns per shopping basket for complementing subsets of products, namely based on price level, discounting, or size.

5.1. Ordered Shopping Baskets

Table 3 presents the estimation results for the model explaining the number of ordered shopping baskets per customer. The availability of offline stores of the retailer does not help or harm the online channel, as it does not significantly influence the number of ordered shopping baskets (-0.046, 95% credibility interval (CI): -0.127, 0.032). Thus, we find no empirical support for H1a or H1b.

Regarding demographic control variables, customers with a higher age order more (0.003, CI: 0.000, 0.007) and customers who do not disclose their age or gender order less (-0.250, CI: -0.319, -0.180; and -1.215, CI: -1.312, -1.113 respectively). Regional control variables are generally insignificant at the 5% level with the exception of car ownership (0.213, CI: 0.019, 0.390). There is significant heterogeneity at the regional level (0.572, CI: 0.527, 0.619) and significant spatial auto-correlation (0.335, CI: 0.010, 0.857).

Table 3 Regional influence on the number of ordered shopping baskets

DV	**$baskets_c$**		
Intercept	-0.002	(-0.061, 0.056)	
$(W \times stores)_r$	-0.046	(-0.127, 0.032)	
Regional control variables			
$(W \times posts)_r$	0.066	(-0.013, 0.146)	
$welfare_r$	-0.000	(-0.038, 0.039)	
$urbanization_r$	-0.008	(-0.048, 0.033)	
$cars_r$	0.223	(0.019, 0.390)	*
Customer control variables			
$gender\_man_c$	-0.065	(-0.160, 0.027)	
$gender\_missing_c$	-1.215	(-1.312, -1.113)	***
age_c	0.003	(0.000, 0.007)	*
$age\_missing_c$	-0.250	(-0.319, -0.180)	***
Random effects and spatial autocorrelation			
τ (random effect region)	0.572	(0.527, 0.619)	***
λ (spatial auto-correlation)	0.335	(0.010, 0.857)	***

Numbers in brackets indicate the 95% credible interval (CI) and stars indicate whether the 95% (*), 99% (**), 99.9 (***) CI excludes zero.

5.2. Purchases and Returns per Shopping Basket

Table 4 presents the estimation results for the models explaining the number of purchases and returns per shopping basket. Again, we find no evidence that the availability of offline stores helps or harms the online channel in terms of purchased products per shopping basket, as it does not significantly influence the number of purchased products per shopping basket (-0.039, CI: -0.101, 0.024). Therefore, we find no support for H2a or H2b. However, the availability of offline stores has a beneficial effect on returns per shopping basket as it significantly decreases the number of returned products per shopping basket (-0.127, CI: -0.216, -0.038), which is empirical evidence for H3a (and against H3b).

Regarding control variables, men return less than women (-0.511, CI: -0.682, -0.346) and customers who do not disclose their gender have both fewer purchases and returns (-0.511, CI: -0.682, -0.346; and -0.320, CI: -0.434, -0.206 respectively). Conversely, customers who do not disclose their age return more (0.158, CI: 0.040, 0.275). Both purchases and returns increase over time (0.197, CI: 0.165, 0.230; and 0.075, CI: 0.034, 0.115 respectively). Several regional control variables have a significant effect, as well: post office availability increases purchases (0.068, CI: 0.012, 0.126), car ownership increases purchases (0.222, CI: 0.065, 0.375) and decreases returns (-0.437, CI: -0.749, -0.143), and urbanization decreases returns (-0.061, CI: -0.106, -0.016). For both purchases and returns, there is significant spatial auto-correlation (0.492, CI: 0.030, 0.973; and 0.585, CI: 0.041, 0.986 respectively), and significant heterogeneity on the customer (0.736, CI: 0.698, 0.776; and 0.785, CI: 0.723, 0.850, respectively) and regional level (0.056, CI: 0.003, 0.144; and 0.196, CI: 0.034, 0.325, respectively).

Table 4 Regional influences on the general number of purchases and returns (per shopping basket) in the online channel

DV	purchases$_i$			returns$_i$		
Intercept	-2.000	(-2.193, -1.817)	***	-2.879	(-3.111, -2.652)	***
(W x stores)$_r$	-0.039	(-0.101, 0.024)		-0.127	(-0.216, -0.038)	**
Regional control variables						
(W x posts)$_r$	0.068	(0.012, 0.126)	*	-0.059	(-0.140, 0.022)	
welfare$_r$	-0.027	(-0.058, 0.002)		-0.015	(-0.056, 0.025)	
urbanisation$_r$	0.012	(-0.021, 0.045)		-0.061	(-0.106, -0.016)	*
cars$_r$	0.222	(0.065, 0.375)	**	-0.437	(-0.749, -0.143)	**
Customer control variables						
gender_man$_c$	-0.024	(-0.136, 0.088)		-0.511	(-0.682, -0.346)	***
gender_missing$_c$	-0.244	(-0.333, -0.157)	***	-0.320	(-0.434, -0.206)	***
age$_c$	0.000	(-0.004, 0.005)		0.003	(-0.003, 0.009)	
age_missing$_c$	0.033	(-0.056, 0.128)		0.158	(0.040, 0.275)	**
Order control variables						
orders$_i$		n.a.		0.572	(0.543, 0.600)	***
year$_i$	0.197	(0.165, 0.230)	***	0.075	(0.034, 0.115)	***
month$_{i,2}$	0.037	(-0.149, 0.222)		-0.135	(-0.345, 0.073)	
month$_{i,3}$	0.383	(0.235, 0.536)	***	0.058	(-0.128, 0.241)	
month$_{i,4}$	0.487	(0.339, 0.633)	***	0.028	(-0.148, 0.201)	
month$_{i,5}$	0.459	(0.316, 0.602)	***	-0.000	(-0.179, 0.178)	
month$_{i,6}$	0.496	(0.349, 0.645)	***	0.109	(-0.073, 0.287)	
month$_{i,7}$	0.687	(0.548, 0.833)	***	-0.271	(-0.453, -0.082)	**
month$_{i,8}$	0.479	(0.332, 0.628)	***	-0.233	(-0.428, -0.041)	*
month$_{i,9}$	0.315	(0.171, 0.464)	***	-0.438	(-0.628, -0.249)	***
month$_{i,10}$	0.427	(0.278, 0.575)	***	-0.238	(-0.429, -0.050)	**
month$_{i,11}$	0.130	(-0.047, 0.298)		-0.073	(-0.280, 0.122)	
month$_{i,12}$	0.406	(0.241, 0.562)	***	-0.151	(-0.356, 0.044)	
Random effects and spatial autocorrelation						
σ (random effect customer)	0.736	(0.698, 0.776)	***	0.785	(0.723, 0.850)	***
τ (random effect region)	0.056	(0.003, 0.144)	***	0.196	(0.034, 0.325)	***
λ (spatial auto-correlation)	0.492	(0.030, 0.973)	***	0.585	(0.041, 0.986)	***

Numbers in brackets indicate the 95% credible interval (CI) and stars indicate whether the 95% (*), 99% (**), 99.9 (***) CI excludes zero.

5.3. Influence of Risk

Table 5 presents the estimation results for the models explaining the number of purchases and returns per shopping basket for the subsets of products with multiple sizes (multi-size) and the with one size (uni-size). The availability of offline stores does not significantly harm purchases in the online channel of products of either subset (-0.015, CI: -0.041, 0.012; and 0.126, CI: -0.034, 0.281 respectively). Thus, we find no support for H4a. However, the availability of offline stores does have a beneficial influence on product returns as it reduces the number of returns per shopping basket of multi-size products (-0.123, CI: -0.212, -0.037). In line with our reasoning, it has no significant influence on returns of uni-size products (-0.365, CI: -1.147, 0.395). Thus, we find support for H5a.

Table 5 Regional influences on online purchasing and returning (per shopping basket) for uni-size products vs. multi-size products

	uni-size products		multi-size products	
DV	**purchases$_i$**	**returns$_i$**	**purchases$_i$**	**returns$_i$**
Intercept	-3.572 ***	-8.908 ***	-0.074	-2.902 ***
(W x stores)$_r$	0.126	-0.365	-0.015	-0.123 **
Regional control variables				
(W x posts)$_r$	-0.150 *	-0.234	0.028 *	-0.060
welfare$_r$	0.062	-0.119	-0.011	-0.013
urbanisation$_r$	-0.150 ***	0.141	0.010	-0.068 **
cars$_r$	-0.010	-0.716	0.087 *	-0.475 ***
Customer control variables				
gender_man$_c$	-0.956 ***	-0.893 *	0.016	-0.513 ***
gender_missing$_c$	-0.010		-0.082 ***	-0.319 ***
age$_c$	-0.003	-0.053 *	0.000	0.004
age_missing$_c$	-0.067	0.571	0.014	0.151 *
Order control variables				
purchases$_i$	n.a.	3.045 ***	n.a.	0.596 ***
year$_i$	0.013	-0.133	0.057 ***	0.077 ***
month$_{i,2}$	0.243	-1.701	-0.022	-0.124
month$_{i,3}$	-0.105	0.577	0.111 **	0.048
month$_{i,4}$	-0.611 **	0.505	0.152 ***	0.007
month$_{i,5}$	0.087	0.502	0.127 ***	-0.016
month$_{i,6}$	-0.116	0.750	0.147 ***	0.088
month$_{i,7}$	-0.285	-0.439	0.249 ***	-0.285 **
month$_{i,8}$	0.052	0.621	0.148 ***	-0.252 *
month$_{i,9}$	-0.081	0.260	0.090 *	-0.460 ***
month$_{i,10}$	-0.418 *	-0.005	0.124 ***	-0.243 *
month$_{i,11}$	0.321	0.070	0.019	-0.077
month$_{i,12}$	0.711 ***	-0.246	0.056	-0.133
Random effects and spatial autocorrelation				
σ (random effect customer)	1.254 ***	2.149 ***	0.010 ***	0.780 ***
τ (random effect region)	0.117 ***	0.627 ***	0.010 ***	0.203 ***
λ (spatial auto-correlation)	0.506 ***	0.499 ***	0.500 ***	0.601 ***

Stars indicate whether the 95% (*), 99% (**), 99.9 (***) CI excludes zero. Missing and male gender dummies have been collapsed due to low case count for returns of uni-size products.

Table 6 shows the estimation results for the models explaining the number of purchases and returns per shopping basket for the subsets of lower- and higher-price products, respectively. As before, the availability of offline stores has no significant harmful effect for either set of products in terms of number of purchases (-0.018, CI: -0.069, 0.034; and -0.006, CI: -0.041, 0.031 respectively). Thus, we find no empirical evidence for H4b. However, the availability of offline stores again has a beneficial effect by decreasing returns per shopping basket in the case of more expensive products (-0.132, CI: -0.240, -0.029) and, as argued, it has no significant influence on returns of cheaper products (-0.127, CI: -0.283, 0.028). Thus, we find support for H5b.

Table 6 Regional influences on online purchasing and returning products (per shopping basket) priced below or above the median price

	products priced below median		**products priced above median**	
DV	**purchases$_i$**	**returns$_i$**	**purchases$_i$**	**returns$_i$**
Intercept	-1.034 ***	-4.035 ***	-0.547 ***	-3.414 ***
(W x stores)$_r$	-0.018	-0.127	-0.006	-0.132 *
Regional control variables				
(W x posts)$_r$	0.044	-0.060	0.002	-0.063
welfare$_r$	0.019	-0.029	-0.029 ***	0.011
urbanisation$_r$	0.052 ***	-0.052	-0.023 **	-0.055 *
cars$_r$	0.106	-0.375	0.113 **	-0.557 ***
Customer control variables				
gender_man$_c$	0.157 ***	-0.476 ***	-0.135 ***	-0.502 ***
gender_missing$_c$	-0.094 **	-0.355 ***	-0.058 *	-0.263 ***
age$_c$	-0.005 **	-0.001	0.003 *	0.004
age_missing$_c$	0.076	0.086	-0.027	0.183 **
Order control variables				
purchases$_i$	n.a.	0.824 ***	n.a.	0.895 ***
year$_i$	0.038 **	0.044	0.065 ***	0.078 **
month$_{i,2}$	-0.119	-0.247	0.070	-0.102
month$_{i,3}$	-0.250 ***	0.116	0.294 ***	-0.039
month$_{i,4}$	0.026	0.217	0.196 ***	-0.076
month$_{i,5}$	0.293 ***	0.413 **	-0.005	-0.208
month$_{i,6}$	0.471 ***	0.541 ***	-0.174 ***	-0.021
month$_{i,7}$	0.724 ***	0.181	-0.404 ***	-0.434 ***
month$_{i,8}$	0.428 ***	-0.104	-0.156 ***	-0.244 *
month$_{i,9}$	-1.201 ***	-1.816 ***	0.493 ***	-0.484 ***
month$_{i,10}$	-0.994 ***	-1.149 ***	0.488 ***	-0.245 *
month$_{i,11}$	-0.487 ***	-0.272	0.294 ***	-0.079
month$_{i,12}$	-0.150 *	-0.185	0.243 ***	-0.150
Random effects and spatial autocorrelation				
σ (random effect customer)	0.464 ***	1.088 ***	0.022 ***	0.742 ***
τ (random effect region)	0.114 ***	0.210 ***	0.025 ***	0.226 ***
λ (spatial auto-correlation)	0.599 ***	0.511 ***	0.516 ***	0.518 ***

Stars indicate whether the 95% (*), 99% (**), 99.9 (***) CI excludes zero.

Table 7 presents the estimation results for the models explaining the number of purchases and returns per shopping basket for the subsets of discounted and undiscounted products. The influence of offline store availability on purchases/returns of discounted and undiscounted products is in line with the effect of store availability on purchases/returns of cheaper and more expensive products. That is, the availability of offline stores has no significant influence on purchases of both discounted products and undiscounted products (0.001, CI: -0.060, 0.064; and -0.009, CI: -0.039, 0.022 respectively). Thus, we find no support for H4c. Similar to before, the availability of offline stores decreases returns of undiscounted products (-0.153, CI: -0.249, -0.057) while not affecting returns of discounted products (-0.068, CI: -0.237, 0.096), however. Thus, we find support for H5c.

Table 7 Regional influences on online purchasing and returning (per shopping basket) for undiscounted vs. discounted products

	discounted products		**undiscounted products**	
DV	**purchases$_i$**	**returns$_i$**	**purchases$_i$**	**returns$_i$**
Intercept	-4.059 ***	-5.814 ***	0.354 ***	-3.128 ***
(W x stores)$_r$	0.001	-0.068	-0.009	-0.153 **
Regional control variables				
(W x posts)$_r$	-0.029	-0.113	0.036 *	-0.042
welfare$_r$	-0.039 *	-0.018	0.003	-0.006
urbanisation$_r$	0.045 **	-0.094 *	-0.014	-0.055 *
cars$_r$	0.092	-0.607 *	0.074	-0.294
Customer control variables				
gender_man$_c$	-0.245 ***	-0.845 ***	0.071 *	-0.449 ***
gender_missing$_c$	-0.260 ***	-0.440 ***	0.008	-0.271 ***
age$_c$	0.001	0.008	-0.000	0.002
age_missing$_c$	0.055	0.203	-0.019	0.142 *
Order control variables				
purchases$_i$	n.a.	0.940 ***	n.a.	0.775 ***
year$_i$	0.405 ***	0.278 ***	-0.053 ***	0.055 **
month$_{i,2}$	0.489 ***	0.429	-0.114 *	-0.307 *
month$_{i,3}$	1.210 ***	0.642 **	-0.283 ***	0.009
month$_{i,4}$	0.980 ***	0.601 **	-0.099 *	0.033
month$_{i,5}$	1.076 ***	0.360	-0.161 ***	-0.047
month$_{i,6}$	0.493 ***	0.470 *	0.052	0.021
month$_{i,7}$	0.532 ***	-0.458	0.161 ***	-0.377 ***
month$_{i,8}$	0.378 ***	-0.106	0.093 *	-0.287 **
month$_{i,9}$	1.296 ***	-0.010	-0.356 ***	-0.406 ***
month$_{i,10}$	1.123 ***	0.042	-0.169 ***	-0.158
month$_{i,11}$	0.924 ***	0.437	-0.196 ***	-0.048
month$_{i,12}$	1.191 ***	0.295	-0.212 ***	-0.125
Random effects and spatial autocorrelation				
σ (random effect customer)	0.766 ***	1.200 ***	0.014 ***	0.736 ***
τ (random effect region)	0.069 ***	0.146 ***	0.013 ***	0.224 ***
λ (spatial auto-correlation)	0.500 ***	0.504 ***	0.501 ***	0.563 ***

Stars indicate whether the 95% (*), 99% (**), 99.9 (***) CI excludes zero.

5.4. Simulation of Effect Magnitude

Translating the coefficient estimates reported above to actual effects sizes is a non-trivial exercise, since the models we use to calculate the effects of store availability on orders and returns are non-linear. Therefore, in this section, we present the results of using partly simulated data to predict the outcome variable in order to find out what change in the number of stores leads to what reduction in returned products per shopping basket. We restrict the prediction to the models where the effect of stores was significant. Therefore, the predictions pertain to the influence of stores on the number of returned products per shopping basket – regarding all products as well as regarding multi-size, higher-priced, and undiscounted products.

The simulation is based on the actual observed data, with the exception of store availability. We base our simulation as closely as possible on the actual data in order to obtain estimates that reflect as closely as possible what would happen in a real-world situation, when changing nothing but store availability. That means we keep almost the entire database used in the analyses intact, but only modify the data on store availability. Next, we predict the outcome variables for each observation using the estimation results presented in the previous subsections. A more stylized example would allow for more freedom with regards to the assumed customer, product and regional structure but would also suffer from arbitrariness because many different plausible configurations are conceivable, thus lowering the external validity.

With regards to the store availability variable, we use four scenarios of reducing and four scenarios of increasing store availability – in addition to the as-is scenario of actual observed store availability. We created all scenarios based on what would be plausible from a managerial perspective, i.e., closure of the *least* promising stores versus opening of *most* promising stores based on the regional online client-to-store ratio. The underlying rationale is that stores in regions with high online client-to-store ratios will likely have the highest influence on online returns and therefore, it is most beneficial to open such stores first (and close them last). Conversely, stores in regions with low online client-to-store rations will likely have a small influence on online returns and therefore it is beneficial to close such stores first (and open them last). In order to facilitate comparisons, we use the same (absolute) changes in store numbers for opening and closure scenarios. Table 8 lists all scenarios and the associated changes in store availability in detail.

Table 8 Scenarios used in the simulation

Scenario	**Change in no. stores**	**Description**
Complete closure	-430	Closure of all stores
Minimal presence	-195	Close of all but one store per region
Moderate store reduction	-109	Closure of one store in each region which has a below-median client/store ratio
Modest store reduction	-51	Closure of one store in each region with a client/store ratio in the lowest quartile
As-is store availability	±0	Actual store availability as observed in the data
Modest store increase	+51	Opening of one store per region, ranked by high to low (prospective) client/store ratio until the same number of stores are opened which were closed in the “Modest store reduction” scenario
Moderate store increase	+109	Opening of one store per region, ranked by high to low (prospective) client/store ratio until the same number of stores are opened which were closed in the “Moderate store reduction” scenario
Large store increase	+195	Opening of one store per region, ranked by high to low (prospective) client/store ratio until the same number of stores are opened which were closed in the “Minimal presence” scenario
Doubling stores	+430	Opening of one store per region, ranked by high to low (prospective) client/store ratio until the same number of stores are opened which were closed in the “Complete closure” scenario

The results of the simulation show that store openings and closures both can have a large effect on product returns (Figure 2). The effect sizes greatly depend on how many stores are opened or closed. The most pronounced effects result from closing all stores and doubling stores, whereby returns increase by 15.28%-19.32% and decrease by 13.91%-16.26%, respectively. In general, for larger changes, store closures have a bigger effect on returns than store openings, while for small changes, store openings have a bigger effect. For example, a modest store increase reduces returns by 2.41%-2.81% while a modest store reduction only increases returns by 1.59%-1.94%. This asymmetric effect results from stores with likely greatest positive influence being opened first (in the opening scenarios) and being closed last (in the closure scenarios). The effect size varies slightly between product subsets with smallest effect sizes for multi-size products and largest effect sizes for undiscounted products. Figure 2 provides a detailed account of all effects for each scenario and product subset.

Fig. 2 Number of returned products per shopping basket for different product subsets and store availability

6. Discussion

6.1. Summary of Findings

In this paper, we studied the effect of physical, offline, retail stores on the number of product purchases and product returns by customers in the retailer's online channel. The results are partly in line with what we predicted (Table 9).

Table 9 Overview of the main findings

Hypothesis		Findings
H1a	Offline channel presence decreases the *number of ordered shopping baskets*.	Not supported
H1b	Offline channel presence increases the *number of ordered shopping baskets*.	Not supported
H2a	Offline channel presence decreases the *number of purchased products per shopping basket*.	Not supported
H2b	Offline channel presence increases the *number of purchased products per shopping basket*.	Not supported
H3a	Offline channel presence decreases the *number of returned products per shopping basket*.	Supported
H3b	Offline channel presence increases the *number of returned products per shopping basket*.	Not supported
H4a	The impact of offline channel presence on the *number of purchased products per shopping basket* is more negative/less positive for products with higher fit uncertainty compared to products with lower fit uncertainty	Not supported
H4b	The impact of offline channel presence on the *number of purchased products per shopping basket* is more negative/less positive for higher-priced products compared to lower-priced products	Not supported
H5a	The impact of offline channel presence on the *number of returned products per shopping basket* is more negative/less positive for products with higher fit uncertainty compared to products with lower fit uncertainty	Supported
H5b	The impact of offline channel presence on the *number of returned products per shopping basket* is more negative/less positive for higher-priced products compared to lower-priced products	Supported
H5c	The impact of offline channel presence on the *number of returned products per shopping basket* is more negative/less positive for undiscounted products compared to discounted products	Supported

Our key result is that offline retail stores are in no case harmful and in some cases actually beneficial to the online channel. More specifically, the presence of offline retail stores is not harmful to the online channel as it does not decrease the number of products purchased online: it does not significantly affect the number of purchased shopping baskets per customer nor the number of purchased products per shopping basket. Theory indicates two possible mechanisms through which physical, offline, stores could influence online product purchases: by detracting customers from the online channel due to channel substitution (e.g., Neslin & Shankar, 2009), and by attracting customers to the online channel due to cross-channel promotion (Mitra & Fay, 2010). Our findings suggest that both effects are rather small and/or stay in balance.

Besides physical stores not being harmful, their presence can in fact be helpful, as it significantly decreases the number of returned products per shopping basket. This overall pattern, in addition, appears to be most outspoken for products that show a higher perceived risk for customers. These findings are in line with what we theorized. Customers appear to anticipate the risk of needing to return and adapt their purchase behavior, especially for higher-risk products (Janakiraman & Ordóñez, 2012; Mitchell, 1992). They thus seem to engage in show-rooming for those higher-risk products in order to avoid product returns (Neslin & Shankar, 2009). Products with multiple size, a higher price or no discount have a higher perceived risk for customers than uni-size, lower-priced or discounted products. It should therefore not come as a surprise that for the former types of products, physical, offline, stores do reduce product returns whereas for the latter, lower risk, products, offline stores have no significant effect. We find no empirical evidence for an increase in product returns, which could have been fueled by the decreased hassle or cost of actually returning a product (Kim & Wansink, 2012).

Findings of prior research regarding the influence of offline stores on the online channel are mixed. Pauwels and Neslin (2015) find that adding physical stores does not cannibalize purchases in the online channel. Avery et al. (2012), in turn, find that adding physical stores does not influence purchases in the online channel in the short run, but does increase them in the long run. Our findings are consistent with the findings by Pauwels and Neslin (2015), as we find no evidence for a difference in online purchases due to a higher offline store presence. A reason for the difference in findings compared to the study by Avery et al. (2012) might be that the retailer in our study is large and well-known so that stores did not increase brand awareness, which was suggested as an explanation for the positive effect in their study. If there would be a positive effect of offline stores on online orders, it would in fact strengthen the notion that stores are helpful and not harmful to the online channel.

Regarding product returns, Pauwels and Neslin (2015) find that adding physical stores does increase the total return frequency. Their findings, however, are across all channels, thereby not distinguishing between returns of products purchased online vs offline. As a consequence, it is not possible to directly compare our study with their study. Besides, one explanation of the difference might be the higher number of nearby stores in our study. When the retailer has an offline channel but stores are very distant, stores might be attractive enough for customers to use for returning but not attractive enough to use for casual show-rooming, thus fostering returns instead of averting them.

Our study is the first to investigate the influence of the offline channel on purchases and returns in the online channel which distinguishes between effects for products with different risk perceptions for the customer. Our findings show that the investigated risk-perception driving product characteristics are a decisive factor for offline- on online-channel influence, and that neglecting them masks an important aspect in the relationship between both channels. Furthermore, we developed a statistical model that takes the regional nature of stores into account and showed that the estimation of a spatial model in this case is feasible, even for a large number of regions. By that, we contribute to prior research on multi-channel interactions both substantially as well as methodologically.

6.2. Managerial Implications

Over the course of the last decades, customers begun to shop more and more online, which led to the rise of e-commerce companies such as Amazon and the creation of online channels by traditional retailers such as Wal-Mart. The increasing importance of the online channel poses the question whether the offline channel is still beneficial to maintain (or further develop) – or whether it harms the retailer's performance. Our findings indicate no signs of cannibalization of the online channel by the offline channel. Instead, physical, offline, stores decrease returns and thereby help to increase the performance of the online channel, which often suffers from high return rates (Ofek et al., 2011).

The usefulness of physical stores, however, depends on the retailer's profile and strategy. The retailer's assortment determines how much overall reduction of product returns in the online channel can be expected. The higher the risk of non-fit for the customer or the more expensive the products on display are, the more returns are reduced. That means retailers, which sell higher price products, products without many discounts and products which have personal fit requirements, stand to benefit the most from physical retail stores.

Furthermore, our simulation shows that effects of small increases and decreases in the number of stores are asymmetric, with an increase in stores leading to a stronger reduction in returns compared to the increase in returns that results from a similar reduction in the number of stores. This indicates that retailers likely benefit from accounting for online customer locations when setting store locations. While a small reduction in the number of – from the point of view of the online channel: unappealing – stores, as a consequence, may not harm the retailer much with regard to product returns, a small increase in the number of – from the point of view of the online channel: appealing – stores will help more. Besides, increasing store availability shows diminishing returns, i.e., adding a few stores has a higher average per-store effect on product returns than adding many stores. That means that retailers stand to gain most, per-store, by the first additional store openings.

Overall, the decision to which degree physical stores should be maintained depends on many factors. Maintaining stores comes at a substantial cost and managers need to take the actual sales and cost of the offline channel into account. To have an unbiased whole picture, however, managers should also take into account the beneficial effect of physical stores on returns in the online channel, depending on store profile.

6.3. Limitations and Further Research Directions

In our research, we use online data from a fashion retailer. According to Eurostat (2018), clothes and sports goods are the most popular product category for online purchases. Household goods, various media products (books, films, games) and electronic equipment are the next popular ones. All these categories differ in their perceived risk for the customer. Therefore, the effect of physical stores on purchases and returns in the online channel might be different. For example, household goods such as furniture usually have a high price and so the effect of physical stores on returns in the online channel might be even more pronounced. A future study could examine the effect of physical stores on purchases and returns for other product categories. In addition, as we only have data on the online channel, we limit our study to investigate the influence of physical stores on *online* purchases and returns. A future study could integrate offline customer behavior and examine the effect of stores on the combined online and offline purchases and returns.

In addition, we focus our analysis on the within-firm influence of the offline channel on online shopping behavior since firms can directly influence the existence of their own stores. Nevertheless, there could be cross-firm cross-channel influences. Usually, higher switching costs, lock-in effects – sometimes voluntarily set up such as loyalty programs (Bijmolt & Verhoef, 2017) – and assortment differences impede customers from switching easily from one firm to another (Brynjolfsson et al., 2013; Lemon & Verhoef, 2016). In spite of this, one could imagine e.g., a cross-firm category promotion effect. In sum, a future study could investigate the influence of competitor store presence on online purchase and return behavior.

Stores might also have a more or less pronounced promotional and trust-building effect, depending on the retailer. Due to this, we theorized that physical stores could actually increase online sales. While we found no significant effect, this might be due to the fact that our data comes from a large and well-known retailer. In contrast, for smaller and less well-known retailers, this effect might play a role in fostering online purchases

(Avery et al., 2012). A future study could investigate different retailers and find out under which conditions such a cross-channel promotion effect exists.

7. References

Anderson, E. T., Hansen, K., & Simester, D. (2009). The Option Value of Returns: Theory and Empirical Evidence. *Marketing Science*, *28*(3), 405–423. https://doi.org/10.1287/mksc.1080.0430

Avery, J., Steenburgh, T. J., Deighton, J., & Caravella, M. (2012). Adding Bricks to Clicks: Predicting the Patterns of Cross-Channel Elasticities Over Time. *Journal of Marketing*, *76*(3), 96–111. https://doi.org/10.1509/jm.09.0081

Bijmolt, T. H. A., Broekhuis, M., de Leeuw, S., Hirche, C., Rooderkerk, R. P., Sousa, R., & Zhu, S. X. (2019). Challenges at the marketing–operations interface in omni-channel retail environments. *Journal of Business Research*, S014829631930699X. https://doi.org/10.1016/j.jbusres.2019.11.034

Bijmolt, T. H. A., & Verhoef, P. C. (2017). Loyalty Programs: Current Insights, Research Challenges, and Emerging Trends. In B. Wierenga & R. van der Lans (Eds.), *Handbook of Marketing Decision Models* (pp. 143–165). Springer International Publishing.

Bradlow, E. T., Bronnenberg, B., Russell, G. J., Arora, N., Bell, D. R., Duvvuri, S. D., Ter Hofstede, F., Sismeiro, C., Thomadsen, R., & Yang, S. (2005). Spatial models in marketing. *Marketing Letters*, *16*(3–4), 267–278. https://doi.org/10.1007/s11002-005-5891-3

Brooks, S., Gelman, A., Jones, G. L., & Meng, X.-L. (2011). Handbook of Markov Chain Monte Carlo. In *Handbook of Modern Statistical Methods*. Chapman & Hall/CRC.

Brynjolfsson, E., Hu, Y. J., & Rahman, M. S. (2013). Competing in the Age of Omnichannel Retailing. *MIT Sloan Management Review*.

CB Insights. (2018). *We Analyzed 6 E-Commerce Brands Breaking Into Brick-And-Mortar To Understand The Secrets To Their Success—Here's What We Learned*. https://www.cbinsights.com/research/online-offline-retail-ecommerce-brick-and-mortar/

Chen, C. (2019, July 2). All Kohl's stores are now accepting Amazon returns—Here's how it'll work and what Kohl's hopes to win from this convenient service. *Business Insider*. https://www.businessinsider.com/amazon-returns-at-kohls-how-does-it-work

Deloitte. (2018). *Global Powers of Retailing 2018 Transformative change, reinvigorated commerce* (pp. 1–24).

Elhorst, J. P. (2014). *Spatial Econometrics*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40340-8

ESRI. (2016). *ArcGIS REST API*. Environmental Systems Research Institute.

Eurostat. (2018). *E-commerce statistics for individuals—Statistics Explained* (Issue December, pp. 1–12). http://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics_for_individuals

Graham, J. W. (2009). Missing Data Analysis: Making It Work in the Real World. *Annual Review of Psychology*, *60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Grewal, D. (1995). Product quality expectations: Towards an understanding of their antecedents and consequences. *Journal of Business and Psychology*, *9*(3), 225–240. https://doi.org/10.1007/BF02230967

Gu, Z., & Tayi, G. K. (2015). Consumer mending and online retailer fit-uncertainty mitigating strategies. *Quantitative Marketing and Economics*, *13*(3), 251–282. https://doi.org/10.1007/s11129-015-9161-2

Halleck Vega, S., & Elhorst, J. P. (2015). The SLX Model. *Journal of Regional Science*, *55*(3), 339–363. https://doi.org/10.1111/jors.12188

Inouye, D. I., Yang, E., Allen, G. I., & Ravikumar, P. (2017). A review of multivariate distributions for count data derived from the Poisson distribution: A review of multivariate distributions for count data. *Wiley Interdisciplinary Reviews: Computational Statistics*, *9*(3), e1398. https://doi.org/10.1002/wics.1398

Janakiraman, N., & Ordóñez, L. (2012). Effect of effort and deadlines on consumer product returns. *Journal of Consumer Psychology*, *22*(2), 260–271. https://doi.org/10.1016/j.jcps.2011.05.002

Kim, J., & Wansink, B. (2012). How Retailers' Recommendation and Return Policies Alter Product Evaluations. *Journal of Retailing*, *88*(4), 528–541. https://doi.org/10.1016/j.jretai.2012.04.004

King, J. (2018, April 2). *Why Shop Online? It's Easy*. EMarketer. https://www.emarketer.com/content/why-shop-online-it-s-easy

Kushwaha, T., & Shankar, V. (2013). Are Multichannel Customers Really More Valuable? The Moderating Role of Product Category Characteristics. *Journal of Marketing*, *77*(4), 67–85. https://doi.org/10.1509/jm.11.0297

Lacombe, D. J., & McIntyre, S. G. (2016). Local and global spatial effects in hierarchical models. *Applied Economics Letters*, *23*(16), 1168–1172. https://doi.org/10.1080/13504851.2016.1142645

Lemon, K. N., & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, *80*(6), 69–96. https://doi.org/10.1509/jm.15.0420

LeSage, J. P. (2014). What Regional Scientists Need to Know About Spatial Econometrics. *SSRN Electronic Journal*, 13–32. https://doi.org/10.2139/ssrn.2420725

LeSage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics*.

Li, H. (2021). Are e-books a different channel? Multichannel management of digital products. *Quantitative Marketing and Economics*. https://doi.org/10.1007/s11129-021-09235-0

Minnema, A., Bijmolt, T. H. A., Gensler, S., & Wiesel, T. (2016). To Keep or Not to Keep: Effects of Online Customer Reviews on Product Returns. *Journal of Retailing*, *92*(3), 253–267. https://doi.org/10.1016/j.jretai.2016.03.001

Minnema, A., Bijmolt, T. H. A., Petersen, J. A., & Shulman, J. D. (2018). Managing Product Returns Within the Customer Value Framework. In R. W. Palmatier, V. Kumar, & C. M. Harmeling (Eds.), *Customer Engagement Marketing* (pp. 95–118). Springer International Publishing. https://doi.org/10.1007/978-3-319-61985-9_5

Mitchell, V. W. (1992). Understanding Consumers' Behaviour: Can Perceived Risk Theory Help? *Management Decision*, *30*(3), 26–31. https://doi.org/10.1108/00251749210013050

Mitra, D., & Fay, S. (2010). Managing Service Expectations in Online Markets: A Signaling Theory of E-tailer Pricing and Empirical Tests. *Journal of Retailing*, *86*(2), 184–199. https://doi.org/10.1016/j.jretai.2010.02.003

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. https://doi.org/10.1016/j.intmar.2008.10.005

Ofek, E., Katona, Z., & Sarvary, M. (2011). "Bricks and Clicks": The Impact of Product Returns on the Strategies of Multichannel Retailers. *Marketing Science*, *30*(1), 42–60. https://doi.org/10.1287/mksc.1100.0588

Pauwels, K., & Neslin, S. A. (2015). Building With Bricks and Mortar: The Revenue Impact of Opening Physical Stores in a Multichannel Environment. *Journal of Retailing*, *91*(2), 182–197. https://doi.org/10.1016/j.jretai.2015.02.001

Petersen, J. A., & Kumar, V. (2015). Perceived Risk, Product Returns, and Optimal Resource Allocation: Evidence from a Field Experiment. *Journal of Marketing Research*, *52*(2), 268–285. https://doi.org/10.1509/jmr.14.0174

Rekuc, D. (2018, April 28). *Study: Why 92% of Retail Purchases Still Happen Offline*. Ripen. https://ripen.com/blog/ecommerce_survey

Rohm, A. J., & Swaminathan, V. (2004). A typology of online shoppers based on shopping motivations. *Journal of Business Research*, *57*(7), 748–757. https://doi.org/10.1016/S0148-2963(02)00351-X

Satell, G. (2017, September). Toys ‘R’ Us Might Be Dying, but Physical Retail Isn’t. *Harvard Business Review*. https://hbr.org/2017/09/toys-r-us-is-dead-but-physical-retail-isnt

Soper, S. (2018, September 18). *Amazon Will Consider Opening Up to 3,000 Cashierless Stores by 2021*. Bloomberg. https://www.bloomberg.com/news/articles/2018-09-19/amazon-is-said-to-plan-up-to-3-000-cashierless-stores-by-2021

Stan Development Team. (2017). *Stan Modeling Language User’s Guide and Reference Manual*. 397.

Taylor, K. (2019). *The retail apocalypse is far from over as analysts predict 75,000 more store closures*. Business Insider. https://www.businessinsider.com/retail-apocalypse-thousands-store-closures-predicted-2019-4

Thompson, D. (2017). *The 4 Reasons Why 2017 Is a Tipping Point for Retail*. The Atlantic. https://www.theatlantic.com/business/archive/2017/11/future-retail/546119/

van Nierop, J. E. M., Leeflang, P. S. H., Teerling, M. L., & Huizingh, K. R. E. (2011). The impact of the introduction and use of an informational website on offline customer buying behavior. *International Journal of Research in Marketing*, *28*(2), 155–165. https://doi.org/10.1016/j.ijresmar.2011.02.002

Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From Multi-Channel Retailing to Omni-Channel Retailing. Introduction to the Special Issue on Multi-Channel Retailing. *Journal of Retailing*, *91*(2), 174–181. https://doi.org/10.1016/j.jretai.2015.02.005

Wang, K., & Goldfarb, A. (2017). Can Offline Stores Drive Online Sales? *Journal of Marketing Research*, *54*(5), 706–719. https://doi.org/10.1509/jmr.14.0518

Weltevreden, J. W. J. (2007). Substitution or complementarity? How the Internet changes city centre shopping. *Journal of Retailing and Consumer Services*, *14*(3), 192–207. https://doi.org/10.1016/j.jretconser.2006.09.001

Zaczkiewicz, A. (2018). *H&M, Asos and Zara Top List of Most Visited E-commerce Sites*. WWD. https://wwd.com/business-news/business-features/semrush-ecommerce-list-1202566734/

8. Appendix

8.1. Approximation of the Spatial Durbin Error Term

The computation of the error term u relies on the inverse of a function of the fixed weighting matrix and the variable spatial autocorrelation parameter ρ:

(1) $u = f(\rho)\varepsilon = (I - \rho W)^{-1}\varepsilon,$

An approximation using the Neumann series allows us to avoid computing the inverse, which would be computationally expensive, and still gives decent results. The Neumann series for a matrix M is defined as

(2) $(I - M)^{-1} = \sum_{k=0}^{\infty} M^k$

with I being the identity matrix. If we substitute $M = \rho W$, we get

(3) $(I - \rho W)^{-1} = \sum_{k=0}^{\infty} p^k W^k = f(\rho).$

In practice, we use the Neumann series of sixth order, so that we have

(4) $f(\rho) \approx I + p^1W^1 + p^2W^2 + p^3W^3 + p^4W^4 + p^5W^5$

which we use to calculate u in (1).

Using this method to approximate u, we found a correlation of more than 0.99 when comparing exact and approximately calculated u for different values of ρ for our concrete weight matrix W. The benefit is that we can pre-compute the powers of the large matrix W, as they do not depend on the parameter ρ, before the estimation. During the estimation, we just need to multiply each pre-computed W^k by a scalar, and the overall sum by a vector. The resulting gain in efficiency is high: Calculating the exact value in R takes about half a minute whereas the approximation – when relying on precomputed powers of W – takes around one second. LeSage and Pace (2009) propose a conceptually related method.

8.2. Model Convergence

In general, there needs to be convergence within and between the so-called chains, to be able to use and interpret the model estimates. A chain is a stochastic process, whose equilibrium distribution is the (posterior density) estimate of the model parameters. There should hence be no continuous movement of a chain in one direction, indicating that the equilibrium distribution has not been reached. There needs to be convergence between chains, i.e., all chains need to converge to a similar value, irrespective of their (random) starting value (Brooks et al., 2011). The so-called potential scale reduction factor (PSRF or $\hat{R}$) is usually used to check convergence. It compares "the variance of the simulations from each chain […], average[s] these within-chain variances, and compare[s] this to the variances of all the chains mixed together" (Brooks et al., 2011, p. 170). The square root of the latter (variance of all chains taken together) divided by the former is $\hat{R}$. The estimates of the burn-in phase are not included in the calculation. A value of one indicates ideal convergence, while a $\hat{R}$ value of less than 1.1 is usually considered as acceptable (Brooks et al., 2011). All relevant parameters of all aforementioned models are estimated with an $\hat{R}$, which lies in the acceptable range of < 1.1.

8.3. Simulation

To verify the estimation process, we simulated data according to the model specification. Our simulated dataset is approximately of the size of the real dataset at each hierarchical level (J=4,000 regions; M=10,000 customers; N=15,000 orders). The assignment of lower levels to higher levels is made so that it mirrors that real dataset. Practically, we base the assignment on an appropriate gamma distribution and then transform the result to assure that all lower level observations are assigned to a higher level. Concerning the spatial weight matrix, we assign up to 18 neighbors to each region, with the distance based on the inverse of the absolute difference in row/column numbers in the matrix. On the order and client level, we use five independent variables, on the regional level two and on the cross-regional level three. For each, we draw random data from U(0,1). Estimation shows that for the both the zero-truncated as well as the regular Poisson case, true parameters are recovered generally well: 19 parameters out of 19 lie in the 95% credibility interval (i.e. 100%) and 17 out of 19 in the 90% credibility interval (i.e. 89%) for the zero-truncated Poisson case. 17 out of 19 lie in the 95% credibility interval and in the 90% credibility interval for the regular Poisson case. The cross-regional parameters are always recovered in the 90% (and thus 95%) credibility interval.

List of research reports

16001-GEM: Hoorn, A. van, How Are Migrant Employees Manages? An Integrated Analysis

16002-EEF: Soetevent, A.R., Te Bao, A.L. Schippers, A Commercial Gift for Charity

16003-GEM: Bouwmeerster, M.C., and J. Oosterhaven, Economic Impacts of Natural Gas Flow Disruptions

16004-MARK: Holtrop, N., J.E. Wieringa, M.J. Gijsenberg, and P. Stern, Competitive Reactions to Personal Selling: The Difference between Strategic and Tactical Actions

16005-EEF: Plantinga, A. and B. Scholtens, The Financial Impact of Divestment from Fossil Fuels

16006-GEM: Hoorn, A. van, Trust and Signals in Workplace Organization: Evidence from Job Autonomy Differentials between Immigrant Groups

16007-EEF: Willems, B. and G. Zwart, Regulatory Holidays and Optimal Network Expansion

16008-GEF: Hoorn, A. van, Reliability and Validity of the Happiness Approach to Measuring Preferences

16009-EEF: Hinloopen, J., and A.R. Soetevent, (Non-)Insurance Markets, Loss Size Manipulation and Competition: Experimental Evidence

16010-EEF: Bekker, P.A., A Generalized Dynamic Arbitrage Free Yield Model

16011-EEF: Mierau, J.A., and M. Mink, A Descriptive Model of Banking and Aggregate Demand

16012-EEF: Mulder, M. and B. Willems, Competition in Retail Electricity Markets: An Assessment of Ten Year Dutch Experience

16013-GEM: Rozite, K., D.J. Bezemer, and J.P.A.M. Jacobs, Towards a Financial Cycle for the US, 1873-2014

16014-EEF: Neuteleers, S., M. Mulder, and F. Hindriks, Assessing Fairness of Dynamic Grid Tariffs

16015-EEF: Soetevent, A.R., and T. Bružikas, Risk and Loss Aversion, Price Uncertainty and the Implications for Consumer Search

16016-HRM&OB: Meer, P.H. van der, and R. Wielers, Happiness, Unemployment and Self-esteem

16017-EEF: Mulder, M., and M. Pangan, Influence of Environmental Policy and Market Forces on Coal-fired Power Plants: Evidence on the Dutch Market over 2006-2014

16018-EEF: Zeng,Y., and M. Mulder, Exploring Interaction Effects of Climate Policies: A Model Analysis of the Power Market

16019-EEF: Ma, Yiqun, Demand Response Potential of Electricity End-users Facing Real Time Pricing

16020-GEM: Bezemer, D., and A. Samarina, Debt Shift, Financial Development and Income Inequality in Europe

16021-EEF: Elkhuizen, L, N. Hermes, and J. Jacobs, Financial Development, Financial Liberalization and Social Capital

16022-GEM: Gerritse, M., Does Trade Cause Institutional Change? Evidence from Countries South of the Suez Canal

16023-EEF: Rook, M., and M. Mulder, Implicit Premiums in Renewable-Energy Support Schemes

17001-EEF: Trinks, A., B. Scholtens, M. Mulder, and L. Dam, Divesting Fossil Fuels: The Implications for Investment Portfolios

17002-EEF: Angelini, V., and J.O. Mierau, Late-life Health Effects of Teenage Motherhood

17003-EEF: Jong-A-Pin, R., M. Laméris, and H. Garretsen, Political Preferences of (Un)happy Voters: Evidence Based on New Ideological Measures

17004-EEF: Jiang, X., N. Hermes, and A. Meesters, Financial Liberalization, the Institutional Environment and Bank Efficiency

17005-EEF: Kwaak, C. van der, Financial Fragility and Unconventional Central Bank Lending Operations

17006-EEF: Postelnicu, L. and N. Hermes, The Economic Value of Social Capital

17007-EEF: Ommeren, B.J.F. van, M.A. Allers, and M.H. Vellekoop, Choosing the Optimal Moment to Arrange a Loan

17008-EEF: Bekker, P.A., and K.E. Bouwman, A Unified Approach to Dynamic Mean-Variance Analysis in Discrete and Continuous Time

17009-EEF: Bekker, P.A., Interpretable Parsimonious Arbitrage-free Modeling of the Yield Curve

17010-GEM: Schasfoort, J., A. Godin, D. Bezemer, A. Caiani, and S. Kinsella, Monetary Policy Transmission in a Macroeconomic Agent-Based Model

17011-I&O: Bogt, H. ter, Accountability, Transparency and Control of Outsourced Public Sector Activities

17012-GEM: Bezemer, D., A. Samarina, and L. Zhang, The Shift in Bank Credit Allocation: New Data and New Findings

17013-EEF: Boer, W.I.J. de, R.H. Koning, and J.O. Mierau, Ex-ante and Ex-post Willingness-to-pay for Hosting a Major Cycling Event

17014-OPERA: Laan, N. van der, W. Romeijnders, and M.H. van der Vlerk, Higher-order Total Variation Bounds for Expectations of Periodic Functions and Simple Integer Recourse Approximations

17015-GEM: Oosterhaven, J., Key Sector Analysis: A Note on the Other Side of the Coin

17016-EEF: Romensen, G.J., A.R. Soetevent: Tailored Feedback and Worker Green Behavior: Field Evidence from Bus Drivers

17017-EEF: Trinks, A., G. Ibikunle, M. Mulder, and B. Scholtens, Greenhouse Gas Emissions Intensity and the Cost of Capital

17018-GEM: Qian, X. and A. Steiner, The Reinforcement Effect of International Reserves for Financial Stability

17019-GEM/EEF: Klasing, M.J. and P. Milionis, The International Epidemiological Transition and the Education Gender Gap

2018001-EEF: Keller, J.T., G.H. Kuper, and M. Mulder, Mergers of Gas Markets Areas and Competition amongst Transmission System Operators: Evidence on Booking Behaviour in the German Markets

2018002-EEF: Soetevent, A.R. and S. Adikyan, The Impact of Short-Term Goals on Long-Term Objectives: Evidence from Running Data

2018003-MARK: Gijsenberg, M.J. and P.C. Verhoef, Moving Forward: The Role of Marketing in Fostering Public Transport Usage

2018004-MARK: Gijsenberg, M.J. and V.R. Nijs, Advertising Timing: In-Phase or Out-of-Phase with Competitors?

2018005-EEF: Hulshof, D., C. Jepma, and M. Mulder, Performance of Markets for European Renewable Energy Certificates

2018006-EEF: Fosgaard, T.R., and A.R. Soetevent, Promises Undone: How Committed Pledges Impact Donations to Charity

2018007-EEF: Durán, N. and J.P. Elhorst, A Spatio-temporal-similarity and Common Factor Approach of Individual Housing Prices: The Impact of Many Small Earthquakes in the North of Netherlands

2018008-EEF: Hermes, N., and M. Hudon, Determinants of the Performance of Microfinance Institutions: A Systematic Review

2018009-EEF: Katz, M., and C. van der Kwaak, The Macroeconomic Effectiveness of Bank Bail-ins

2018010-OPERA: Prak, D., R.H. Teunter, M.Z. Babai, A.A. Syntetos, and J.E. Boylan, Forecasting and Inventory Control with Compound Poisson Demand Using Periodic Demand Data

2018011-EEF: Brock, B. de, Converting a Non-trivial Use Case into an SSD: An Exercise

2018012-EEF: Harvey, L.A., J.O. Mierau, and J. Rockey, Inequality in an Equal Society

2018013-OPERA: Romeijnders, W., and N. van der Laan, Inexact cutting planes for two-stage mixed-integer stochastic programs

2018014-EEF: Green, C.P., and S. Homroy, Bringing Connections Onboard: The Value of Political Influence

2018015-OPERA: Laan, N. van der, and W. Romeijnders, Generalized aplha-approximations for two-stage mixed-integer recourse models

2018016-GEM: Rozite, K., Financial and Real Integration between Mexico and the United States

2019001-EEF: Lugalla, I.M., J. Jacobs, and W. Westerman, Drivers of Women Entrepreneurs in Tourism in Tanzania: Capital, Goal Setting and Business Growth

2019002-EEF: Brock, E.O. de, On Incremental and Agile Development of (Information) Systems

2019003-OPERA: Laan, N. van der, R.H. Teunter, W. Romeijnders, and O.A. Kilic, The Data-driven Newsvendor Problem: Achieving On-target Service Levels.

2019004-EEF: Dijk, H., and J. Mierau, Mental Health over the Life Course: Evidence for a U-Shape?

2019005-EEF: Freriks, R.D., and J.O. Mierau, Heterogeneous Effects of School Resources on Child Mental Health Development: Evidence from the Netherlands.

2019006-OPERA: Broek, M.A.J. uit het, R.H. Teunter, B. de Jonge, J. Veldman, Joint Condition-based Maintenance and Condition-based Production Optimization.

2019007-OPERA: Broek, M.A.J. uit het, R.H. Teunter, B. de Jonge, J. Veldman, Joint Condition-based Maintenance and Load-sharing Optimization for Multi-unit Systems with Economic Dependency

2019008-EEF: Keller, J.T. G.H. Kuper, and M. Mulder, Competition under Regulation: Do Regulated Gas Transmission System Operators in Merged Markets Compete on Network Tariffs?

2019009-EEF: Hulshof, D. and M. Mulder, Renewable Energy Use as Environmental CSR Behavior and the Impact on Firm Profit

2019010-EEF: Boot, T., Confidence Regions for Averaging Estimators
2020001-OPERA: Foreest, N.D. van, and J. Wijngaard. On Proportionally Fair Solutions for the Divorced-Parents Problem

2020002-EEF: Niccodemi, G., R. Alessie, V. Angelini, J. Mierau, and T. Wansbeek. Refining Clustered Standard Errors with Few Clusters

2020003-I&O: Bogt, H. ter, Performance and other Accounting Information in the Public Sector: A Prominent Role in the Politicians' Control Tasks?

2020004-I&O: Fisch, C., M. Wyrwich, T.L. Nguyen, and J.H. Block, Historical Institutional Differences and Entrepreneurship: The Case of Socialist Legacy in Vietnam

2020005-I&O: Fritsch, M. and M. Wyrwich. Is Innovation (Increasingly) Concentrated in Large Cities? An Internatinal Comparison

2020006-GEM: Oosterhaven, J., Decomposing Economic Growth Decompositions.

2020007-I&O: Fritsch, M., M. Obschonka, F. Wahl, and M. Wyrwich. The Deep Imprint of Roman Sandals: Evidence of Long-lasting Effects of Roman Rule on Personality, Economic Performance, and Well-Being in Germany

2020008-EEF: Heijnen, P., On the Computation of Equilibrium in Discontinuous Economic Games

2020009-EEF: Romensen, G.J. and A.R. Soetevent, Improving Worker Productivity Through Tailored Performance Feedback: Field Experimental Evidence from Bus Drivers

2020010-EEF: Rao, Z., M. Groneck, and R. Alessie, Should I Stay or Should I Go? Intergenerational Transfers and Residential Choice. Evidence from China

2020011-EEF: Kwaak, C. van der, Unintended Consequences of Central Bank Lending in Financial Crises

2020012-EEF: Soetevent, A.R., Determinants choice set variation in demand estimation – with an application to the electric vehicle public charging market

2020013-EEF: Kwaak, C. van der, Old-Keynesianism in the New Keynesian model

2020014-EEF: Plaat, m. van der, Loan Sales and the Tyranny of Distance in U.S. Residential Mortgage Lending

2020015-I&O: Fritsch, M., and M. Wyrwich, Initial Conditions and Regional Performance in the Aftermath of Disruptive Shocks: The Case of East Germany after Socialism

2020016-OPERA: Laan, N. van der, and W. Romeijnders, A Converging Benders' Decomposition Algorithm for Two-stage Mixed-integer Recourse Models

2021001-OPERA: Baardman, L., K.J. Roodbergen, H.J. Carlo, and A.H. Schrotenboer, A Special Case of the Multiple Traveling Salesmen Problem in End-of-aisle Picking Systems

2021002-EEF: Wiese, R., and S. Eriksen, Willingness to Pay for Improved Public Education and Public Health Systems: The Role of Income Mobility Prospects.

2021003-EEF: Keller, J.T., G.H. Kuper, and M. Mulder, Challenging Natural Monopolies: Assessing Market Power of Gas Transmission System Operators for Cross-Border Capacity

2021004-EEF: Li, X., and M. Mulder, Value of Power-to-Gas as a Flexibililty Option in Integrated Electricity and Hydrogen Markets

2021005-GEM: Rozite, K., J.P.A.M. Jacobs, and D.J. Bezemer, Investor Sentiment and Business Investment

2021006-EEF: Spierdijk, L., and T. Wansbeek, Differencing as a Consistency Test for the Within Estimator

2021007-EEF: Katz, M., and C. van der Kwaak, To Bail-in or to Bailout: that's the (Macro) Question

2021008-EEF: Haan, M.A., N.E. Stoffers, and G.T.J. Zwart, Choosing Your Battles: Endogenous Multihoming and Platform Competition

2021009-I&O: Greve, M., M. Fritsch, and M. Wyrwich, Long-Term Decline of Regions and the Rise of Populism: The Case of Germany

2021010-MARK: Hirche, C.F., T.H.A. Bijmolt, and M.J. Gijsenberg, When Offline Stores Reduce Online Returns www.rug.nl/feb