

Loy, Jens-Peter; Ren, Yanjun

Conference Paper — Published Version

Web scraping of food retail prices: An analysis of internet food retail sales prices

Suggested Citation: Loy, Jens-Peter; Ren, Yanjun (2022) : Web scraping of food retail prices: An analysis of internet food retail sales prices, In: Gandorfer, Markus Hoffmann, Christa El Benni, Nadja Cockburn, Marianne Anken, Thomas Floto, Helga (Ed.): Künstliche Intelligenz in der Agrar- und Ernährungswirtschaft, 42. GIL-Jahrestagung, Tänikon, Ettenhausen, 21.-22.02.2022, ISBN 978-3-88579-711-1, Gesellschaft für Informatik e.V., Bonn, pp. 177-182, <https://hdl.handle.net/20.500.12116/38392>

This Version is available at:

<https://hdl.handle.net/10419/251552>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by-sa/4.0/>

Web scraping of food retail prices

An analysis of internet food retail sales prices

Jens-Peter Loy¹ and Yanjun Ren²

Abstract: In this paper, we develop a theory of food retail promotional strategy. We test the theory using online food retail prices. A python code is applied to retrieve information from the web page mytime.de. Mytime.de is an online grocery outlet that belongs to the Bunting Group, a food retailer in North-West Germany. The promotional sales on mytime.de show a complementary relationship between breadth and depth of sales, indicating that in order to attract consumers, stores raise both the number (breadth) and the depth of price promotions.

Keywords: e-food retailing, promotional sales, Germany

1 Introduction

Price promotions (sales) dominate the marketing of many (food) retailers [Si92]. Twenty to fifty percent of stores' food product price variations are due to sales' prices (see [HR04] and [Be08]). Several models have been proposed to explain the sale's phenomenon, see [Va80; Bl81; La86], amongst others. These models also reveal some information on the level of price discounts (depth). However, most of these models focus on a single product or on a small subset of products and fail to address the problem of the optimal number of sales. Actual grocery stores, however, carry between a few thousand up to 35,000 products and offer various products on sale [KP05].

In this paper, we focus on multi-product retailers that carry n products and specifically address the question of the optimal number (breadth) and the optimal depth of promotional sales. We develop a theoretical model of a multiproduct retailer in the fashion of Varian [Va80] to derive the relationship between breadth and depth of promotional sales. Following, we collect data on price promotions by web scraping from the online grocery store mytime.de to test the theory.

¹Christian-Albrechts-Universität zu Kiel, Marktlehre, Olshausenstraße 40, 24118 Kiel, jploy@ae.uni-kiel.de

²Leibniz-Institut für Agrarentwicklung in Transformationsökonomien (IAMO), Theodor-Lieser-Str. 2, 06120 Halle, ren@iamo.de

2 Depth and breadth of sales

We assume that consumers purchase one unit of each (perishable or non-storable) product of a fixed food basket.³ The number of products in the shopping basket (n) is fixed and identical for all consumers. Each retailer k of the m retailers in the market sells all n products. Some products (n_1) are put on sale at a discount price p and all other goods ($n_2 = n - n_1$) are sold at consumer reservation price v . There are two types of consumers: shoppers (non-store loyal) and store-loyal customers. The share of shoppers is $(1 - \alpha)$, the share of store-loyal customers is α/m .⁴ We assume shoppers to be fully informed about store prices at all times and to purchase at the store that offers the lowest costs (price index) for the fixed food basket.⁵ In general, consumers perform one-stop-shopping; however, when price discounts for individual products rise, shoppers start cherry picking and collect items on sale across different stores. Loyal consumers always purchase at the same store. Retailer k chooses depth (the discount price $p_{i,k}$ for product $i \in n_{1,k}$) and breadth of sales (the number of products on sale $n_{1,k}$) simultaneously to maximize profits. For non-sale items, the customer's reservation prices are set. To simplify, the same reservation price is set for all consumers and for all n products. We also assume that buying-in prices or product costs are zero. Following [Va80], a store can choose to put products on sale or not; in either case the expected profit in equilibrium needs to be the same. In case no product is put on sale, only store-loyal customers buy at the store. C indicates the fixed costs of retailing. The profit of an individual store k that offers no sales (ns) is shown in (1). If we assume symmetry over all stores and products, (1) is simplified as shown in (2):⁶

$$\pi_{ns,k} = \sum_{i=1}^n (\alpha/m)v_{ik} - C_k \quad (1)$$

$$\pi_{ns} = n(\alpha/m)v - C \quad (2)$$

By offering products on sale, a store can attract shoppers and thereby increase the sale volume and profit. This happens if and only if the store wins the sale, meaning that the store offers the product basket at the lowest price index (costs for the fixed food basket). If the store wins the sale and attracts all shoppers, the store's profit consists of regular (n_2) and promotional product (n_1) sales according to Equation (3).

³ To some extent we follow the approach by [Ri06], who also develops a theoretical model to prove complementarity between breadth and depth of sales. [LW09] show that the main result of this model is not valid.

⁴ Without a lack of generality, the number of total consumers is set to one.

⁵ The model does not consider any dynamic interactions but assumes the repetition of actions.

⁶ We assume monopolistic competition; thus the store profit is zero and therefore $C=n(a/m)v$.

$$\pi_s = n_1(\alpha/m + 1 - \alpha)p + n_2(\alpha/m + 1 - \alpha)v - C \quad (3)$$

To see whether there is a certain relationship between depth and breadth of sales, we can take the total derivative of Equation 3:

$$d\pi = (\alpha/m + 1 - \alpha)(p - v)dn_1 + n_1(\alpha/m + 1 - \alpha)dp \quad (3a)$$

For profit maximization ($d\pi=0$), we get $\frac{dn_1}{dp} = -\frac{n_1(\alpha/m + 1 - \alpha)}{(\alpha/m + 1 - \alpha)(p - v)} = \frac{n_1}{v - p} > 0$. This

implies that depth and breadth of sales are substitutes. If the retailer increases the sale's price p (reduces sale's depth), the retailer also increases the number of sales. If the store is not winning the shoppers, the profit is $\pi_f = n_1(\alpha/m)p + n_2(\alpha/m)v - C$. The total derivative in this case is $d\pi_f = (\alpha/m)p dn_1 + n_1(\alpha/m)dp$. For $d\pi_f = 0$ we also

result $\frac{dn_1}{dp} = \frac{n_1}{v - p} > 0$ for the relationship between breadth and depth of sales. We now

introduce a random variable with two values (winning and not winning the shoppers) for the profit and the corresponding probabilities ($F(\cdot), 1 - F(\cdot)$). Because the probabilities winning and not winning the shoppers add up to one, the same result holds for the derivative of the expected profits: $E[\Pi] = F(\cdot)\pi_s + (1 - F(\cdot))\pi_f$. Following, the number of sales and the level of discounts are perfect substitutes at a given level of costs for the food basket. From the consumers' perspective, an offer with many sales and low discounts is equivalent to one with few sales and high discounts. When a retailer varies its level of expenditures for the food basket by putting products on sale to increase the probability of winning the sale, he can freely choose between these options. Thus, we would not expect to find a unique (strictly increasing or decreasing) relationship between the number of items on sale (breadth) and the level of average discounts (depth).⁷ As deciding about prices and publishing of promotional sales information are costly (e.g. store flyers, newspaper advertisements, costs of deciding prices), we introduce non-linear menu costs. Menu costs are assumed to increase at a quadratic rate with the number of sale's items offered by the store (ϕn_1^2). The non-linear nature of menu costs reflects the increasing complexity of determining multiple optimal sales' prices. Every additional promotion needs to consider all interactions between promoted products. For a review of the literature and the empirical significance of menu costs in the field of food and general retailing see [Le97] and [Du99]. Menu costs set incentives for retailers to reduce the number of sales and increase individual discounts.

⁷ Derivation of optimal depth and breadth of sales under these assumptions has no unique solution. Any combination of the number of sales and the level of sales' discounts that leads to the same food basket expenditure is a valid solution. The authors provide a formal proof upon request.

When stores offer large discounts on individual products, informed consumers (shoppers) may start buying only products on sale from different stores (cherry picking) instead of buying the entire food basket in one store (one-stop shopping) depending on their transaction costs. If we assume that stores randomize the choice of products they put on sale, a good amount of the products in the food basket will be on sale at the same time over all stores. Cherry picking reduces the profits for the store that wins the shoppers (see also [Mc09: 425]). There are different approaches to introduce cherry picking in the model. We assume that shoppers are not identical with respect to their transaction costs. Depending on the level of discounts offered, shoppers move from one-stop-shopping to cherry picking. As we assume transaction costs to differ between shoppers in a continuous fashion, the share of shoppers that switch to cherry picking strictly increases with the level of discounts offered. Thus, for a low level of discounts we observe few, for a high level of discounts we observe many shoppers starting cherry picking. Cherry pickers buy all the sales' products in the store that is winning the sale.⁸ However, if sales' discounts are too high, these shoppers start to buy the rest of the food basket in competing stores that offer promotional sales for these items. From a store's perspective, cherry picking reduces the share of shoppers that also buy the products at regular prices to fill their food basket. We model this feature by introducing a function ω that strictly decreases with the level of price discounts offered by the store, respectively strictly increases with the sale's price p . ω is the share of shoppers that buy regular products in the store with the lowest price. A first candidate for ω is p/v . According to [Va80], stores choose the level of discount from a U-shaped random distribution function. As informed consumers in the multi-product case make the store choice based on the expenditures for the entire (fixed) food basket, retailers compete by their aggregate prices (food basket expenses). The products in our model are assumed to be identical (symmetric); thus, the decision which products to promote is arbitrary. The process of selecting products for promotion follows a uniform random process. Further, the random distributions of prices are *i.i.d.* across stores. Thus, consumers get a chance to cherry pick promotional sales across retail stores for the entire food basket. For this model, we can derive that breadth and depth are complements.

3 Data

The Bunting Group started its online grocery store www.mytime.de in April 2019. Mytime.de carries around 28,000 items. The Bunting Group is no. 14 in the list of the top 30 food retailers in Germany. We gather the information by employing tools provided with the programming language python. We collect daily prices of sales offers on www.mytime.de. The website promotes between 200 and 600 products per day. The data

⁸ The intuition is as following: when depth and breadth are complements, the store that wins the shoppers also has the highest level of discounts and the most products on sale compared with all other stores. In case we receive a substitutional relationship, we still assume that cherry pickers buy at the store with the most products on sale to reduce transactions costs of shopping caused by store switching.

are gathered daily from February to May 2020. The difference between sales prices and regular prices relative to the regular price in percent represents the level of discounts. Breadth is the number of products on sale. Due to the Corona pandemic, we only use data before the first lock down in Germany starting at March 13th.

4 Results

For each day we calculate the number of promotional sales items and the respective relative price discount for each item. We calculate the unweighted means for the average price discount per day and correlate it with the number of discounts. We further calculate a price-weighted discount by dividing the average sales' price by the average regular prices. Thereby, price cuts from high price get a stronger weight compared to price cuts of low prices. Figure 1 and 2 show the relationships between breadth and depth of sales.

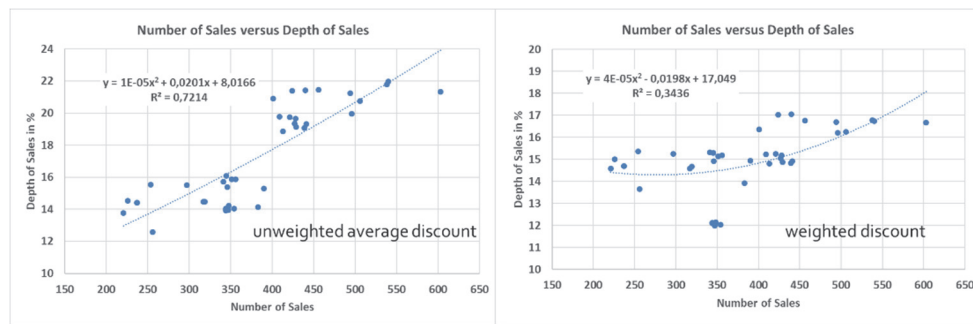


Fig. 1: Breadth and depth of promotional sales (all promotions)

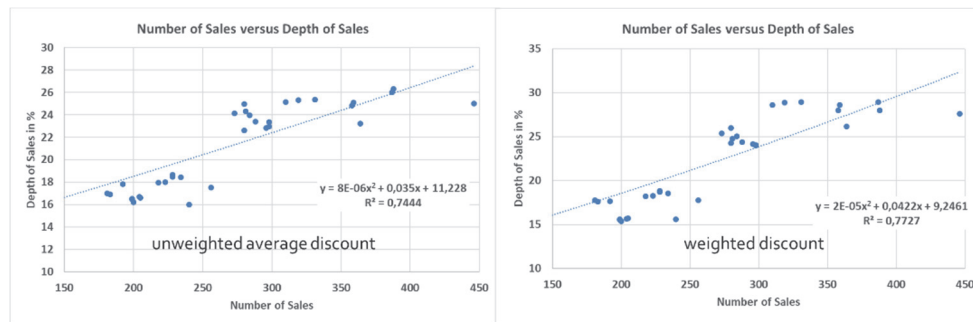


Fig. 2: Breadth and depth of promotional sales (promotions of items below 10 euros)

We also use only sales of less than 10 euros to exclude some sales of big packs which might not be so relevant to most consumers. For all variations in Figure 1 and 2, we find a complementary relationship between breadth and depth of promotional price sales.

5 Summary

Price promotion (sales) are an important tool in food retail marketing. A theory using menu costs and cherry picking predicts a complementary relationship between breadth and depth of sales. We test this theory by web scraping price data from an online food retailer. First evidence shows that breadth and depth of promotional sales are complements. The model may need to be extended to capture other factors influencing the promotional behavior of retailers, such as demand peaks (e. g. holidays), product characteristics (e. g. signpost items), contracts with the processing industry (e. g. promotion commitments) etc. The results allow us to better understand the factors determining the most important marketing tool in food retailing.

Literaturverzeichnis

- [Be08] Berck P., J. Brown, J. M. Perloff and S. Villas-Boas (2008): Sales: tests of theories of causality and timing. *International Journal of Industrial Organization*, 26: S. 1257-1273.
- [Bl81] Blattberg R. C., G. D. Eppen and J. Lieberman (1981): A theoretical and empirical evaluation of price deals for consumer nondurables. *Journal of Marketing*, 45(1): S. 116-129.
- [Du99] Dutta S., M. Bergen, D. Levy and R. Venable (1999): Menu costs, posted prices, and multiproduct retailers. *Journal of Money, Credit and Banking*, 31(4): S. 683-703.
- [HR04] Hosken D. and D. Reiffen (2004): How retailers determine which products should go on sale: evidence from store-level data. *Journal of Consumer Policy*, 27(2): S. 141-177.
- [KP05] KPMG (2005): Consumers markets and retail. *Der deutsche Lebensmitteleinzelhandel aus Verbrauchersicht*. Cologne. (Download at www.kpmg.de)
- [La86] Lazear E.P. (1986): Retail pricing and clearance sales. *The American Economic Review*, 76(1): S. 14-32.
- [Le97] Levy D., M. Bergen, S. Dutta and R. Venable (1997): The magnitude of menu costs: direct evidence from large U.S. supermarket chains. *Quarterly Journal of Economics*, 112: S. 791-825.
- [LW09] Loy J.-P. and C. R. Weiss (2009): Promotional food retail sales: frequency versus depth: a comment. *Managerial and Decision Economics*, 30: S. 513-515.
- [Mc09] McAlister L., E. I. George and Y.-H. Chien (2009): A basket-mix model to identify cherry-picked brands. *Journal of Retailing*, 85 (4): S. 425-436.
- [Ri06] Richards T. J. (2006): Sales by multi-product retailers. *Managerial and Decision Economics*, 27: S. 261-277.
- [Si92] Simon H. (1992): *Preismanagement – Analyse – Strategie – Umsetzung*. 2. Edition Gabler, Würzburg.
- [Va80] Varian H. R. (1980): A model of sales. *The American Economic Review*, 70(4): S. 651-659.