The Influence of Competitor Prices on Price Sensitivity: A Bayesian Pricing Strategy in E-Commerce

Master Thesis
Econometrics and Management Science

Aron Hartveld
Erasmus School of Economics, Erasmus University Rotterdam
3062 PA Rotterdam, The Netherlands, 344544ah@eur.nl

Abstract. Pricing is a cornerstone in almost every business. Increased price pressure, due to better informed customers online, enhances the need for a sophisticated understanding of price effects. In this paper, we add to existing pricing literature by thoroughly studying effects of competitor prices on price sensitivity of customers in e-commerce. We test our hypothesis that price sensitivity increases as the price approaches the competitor median price. We employ a hierarchical Bayes model, which allows non-linear effects, to test our hypothesis. Furthermore, the model is used to forecast sales and we propose a pricing strategy, based on prices that are optimized by our model. We test the performance of our pricing strategy in a field experiment, implementing our prices in an e-commerce store. We find a high increase in price sensitivity if a product is the cheapest in the market and a linear increase in price sensitivity if the price relative to the market increases. Furthermore, we find that our model is not useful for forecasting and our field experiment results in a 10% sales growth for relatively cheap products, but a 14% decrease in sales for relatively expensive products.

December 2016

Keywords: Pricing Strategy; Non-Linear Price Effects; Competitive Markets; Hierarchical Bayes; Field Experiment

Supervisors:	Prof. dr. D. Fok
	Mr S. Roggeveen, M.Sc.
Co-reader:	Prof. dr. R. Paap

The data used to conduct this research is provided strictly confidential, therefore the author is available to discuss any comment on the content of the paper, but the data can not be shared.

1 Introduction

Pricing is an essential tool in good marketing. Superior pricing analyses and techniques lead to competitive advantages for companies and therefore research on price sensitivity is essential to these companies. The opportunities of digitalization and advanced analytics are exploited to improve pricing, see for example den Boer (2013) and Cope (2007). In contrast, the rapid digitalization of business and the Internet also reduced the search costs for customers and price comparison search engines lead to better informed customers (Bakos, 2001; Ratchford, 2009). Customers who are better aware of prices increase the necessity of competitive prices. As this information is mostly found online, the increase in price transparency is even more important to e-commerce companies than to traditional offline retailers. In addition, the possibilities of product differentiation are limited and customer loyalty is harder to achieve in e-commerce compared to non-Internet markets, this is caused by the lack of the provision of human service (Gefen, 2002).

The need for good pricing policies has been known for a long time and therefore pricing is a well-developed field of research. However, most of the studies assessing price sensitivity limit their research to estimating a constant value for price elasticity, neglecting the complex behavior of price sensitivity in competitive markets (Phillips, 2005). Bijmolt et al. (2005) and Tellis (1988) have written reviews on price elasticity research and all reviewed papers assume constant price elasticity. By assuming a constant price elasticity for all prices, the model is forced to let sales increase as a constant percentage of a price decrease, independent of the absolute value of the price or competitor prices. While these models sometimes incorporate competitor effects on sales, these are independent of price effects. We add to the existing literature by studying the effect of competitor prices on price sensitivity of customers. Our proposed model can reveal these complex effects, as we specify a non-linear model that allows price sensitivity to depend on price and competitor prices.

In an effort to create a model which allows price sensitivity to depend on price, Fok et al. (2007) and Van Heerde et al. (2002) propose models that use regime switching and non-parametric approaches to model price effects respectively. However, these models focus on changes in price sensitivity for different sizes of price changes and do not consider competitor prices in their price sensitivity analysis. The scope of this paper is analyzing changes in price sensitivity as price changes relative to competitor prices. In order to measure the price relative to the competitor prices, we define the price position of a product. The price position of a product is the price of a product divided by the median price of both competitor prices and its own price. A price position of 0.5 means that the price of the product is half the price of the median price and a price position of 2 denotes a price twice as high as the median price.

Modern pricing experts support the theory that demand curves are sigmoidal in shape when plotted against price position (Phillips, 2005). The theory states that price elasticity changes as the price changes relative to the competitors on the market. The intuition behind the theory is that price sensitivity is higher

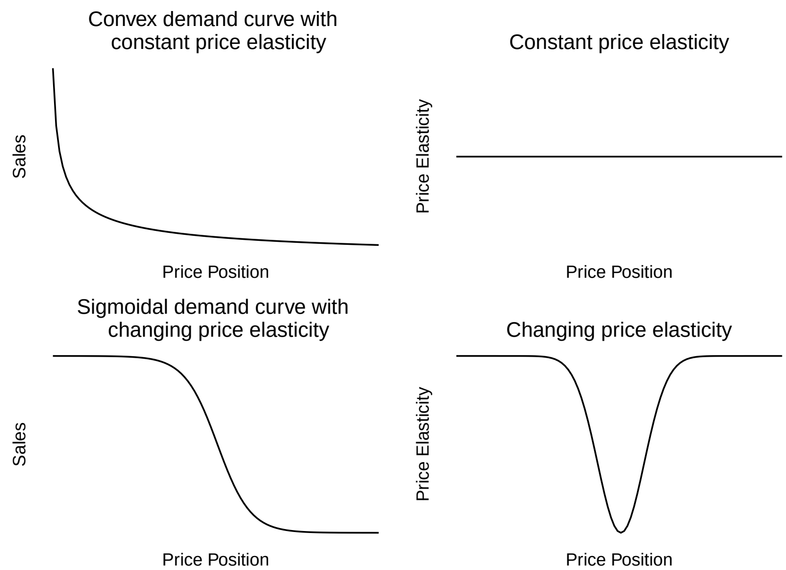

Fig. 1: Demand and price elasticity curves for constant price elasticity in the first row and non-constant price elasticity in the second row.

around the competitor prices, since a small price change can be the difference between being the highest or lowest price in the market. Once price deviates from the competitor prices, we expect price sensitivity to decrease. When a product price is half the price of competitors, increasing or decreasing the price by a percentage should not influence sales a lot, hence price sensitivity is low.

Figure 1 shows the difference in demand and price elasticity curves between a convex and sigmoidal demand model. We can clearly see that ignoring the change in price elasticity leads to biases in expected sales. Pricing strategists therefore over- or underestimate the effects of price changes, missing opportunities to increase profits.

Our model is created in a way that it can estimate the sigmoid demand curve. The model uses state-of-the-art Bayesian techniques to estimate the position, slope and height of the sigmoid curve for products of e-commerce companies. Especially in e-commerce, we expect the sigmoid shaped curves to be present, since price transparency is relatively high (Kocas, 2002; Ratchford, 2009). In addition, we specify the model in a way that it can reject our hypothesis of the sigmoid demand curve, so we can test our hypothesis.

The research is based on data provided by an e-commerce company in the Netherlands. We use daily sales, profit, company and competitor price data of a period consisting of 24 weeks on the level of 65,900 individual products. The sigmoid curve is modeled with respect to the price position of the product, rather than the absolute price. The price position is calculated as a percentage of the median competitor price. The competitors prices are used to determine these

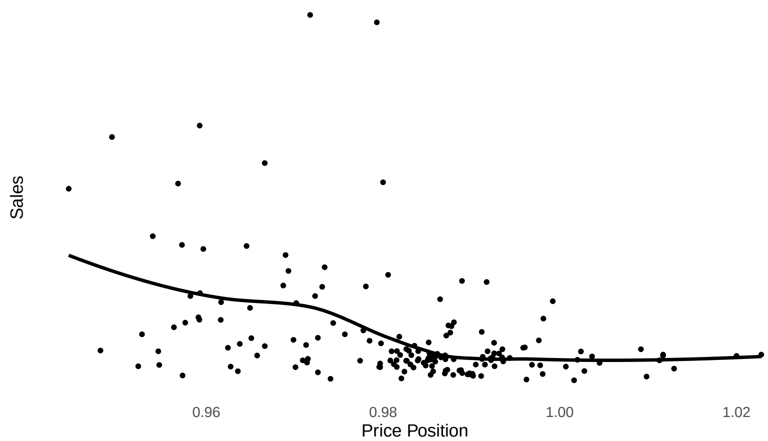

Fig. 2: Aggregated sales plotted against the price position of the product category Jewelry, Watches & Accessories, the resulting shape is an example of a sigmoid demand function.

price positions. Figure 2 shows the aggregated sales of the product category Jewelry, Watches & Accessories plotted against the price position. These sales are an example of a sigmoid demand function, which strengthens our hypothesis.

Additionally to the price position, we measure the additional impact when a company offers a product as cheapest in the market. In the remainder of this paper, we refer to this as the minimal price effect. Lastly, we analyze the effect of the price ticket on price sensitivity of the product. The price ticket is a measure of the absolute value of the price compared to other products, where a product with a low price ticket denotes a cheap product and a product with a high price ticket denotes an expensive product.

Our proposed model uses the well-known Multiplicative Sales Model (MSM) or loglog model as a basis to explain product sales and estimate the price effects. In order to incorporate the changing price elasticities in the model, we employ a hierarchical Bayesian specification. We add a second level regression to the price effect parameter and allow this parameter to depend on the price position. We call this model the Changing Price Response Model (CPRM).

Research by Tellis (1988) and Bijmolt et al. (2005) show a high dispersion in price elasticities for different products and categories, so the model is used to analyze price sensitivity for every product uniquely. This results in a high number of parameters, proportional to the number of products in the model. The hierarchical Bayesian specification allows us to shrink the high number of parameters to reasonable values, by assuming that the parameters for individual products are draws from a common distribution. We call this the Random Effects Specification (RES). This method limits the unwanted variation in the estimation that separate estimations would have. In addition, it allows us to use information of other products, for products that do not have a lot of price changes. Previous research has shown that the hierarchical framework reduces parameter instability and improves forecasting performance, examples of research using the hierarchical Bayesian specification are, among others, Van Nierop et al. (2008) and Boatwright et al. (1999). In addition, this state-of-the-art Bayesian technique is suitable to consider uncertainties in price effects and is therefore capable to find optimal prices even under uncertainty. This makes it an excellent tool to manage risks. We use Markov Chain Monte Carlo (MCMC) simulation techniques to estimate the parameters of the model.

We use the results of the model to assess the opportunities of competitor prices to improve three themes of marketing research. First, we use estimation results to create new insights in price sensitivity. Second, we evaluate forecast possibilities of our new model and lastly, we test the value of our model as a pricing strategy in practice, conducting a field experiment of five weeks.

By studying the estimation results of the CPRM, we find that average elasticities in e-commerce at a price position of the market median is -4.27. Much higher than the average price elasticity found by Bijmolt et al. (2005), -2.62. We find a number of explanations for this difference. First, price sensitivity in e-commerce is higher than price sensitivity in traditional retail. Secondly, we study price elasticity at a certain price position, therefore, other price positions might yield lower price elasticities. Furthermore, we find that high price ticket products are significantly more price sensitive than low price ticket products.

For the effects of price position on price sensitivity, we find that in general price sensitivity increases when price position increases. As a result, the most sensitive price position is also the maximum price position. We do not find evidence for the hypothesized sigmoid shape, however, our model is flexible enough to model the more concave shape suggested by our estimations. In addition, price sensitivity increases when a product has the lowest price position in the market. Lastly, we find very different price sensitivity characteristics across categories and individual products, emphasizing the opportunities for companies of studying products on an individual level.

Besides the estimation of our model, we measure the forecasting performance of the model. We use an out of sample validation technique for one day ahead forecasts. We compare predictive densities of sales forecast of the CPRM to a regular MSM model to study the forecasting power of the changing price elasticity specification. However, our model proves not to be suitable for forecasting sales, as it does not outperform the more simple MSM, without the extensive specification for price position. The complex model is likely to have over fitted the data, finding patterns that only exist in the data sample used for estimation, resulting in an inferior forecasting performance. However, the main scope of our research is analyzing competitor prices and creating a pricing strategy based on those competitor prices. Although lacking forecasting power, the model is still very suitable for these purposes.

The framework is used to create a pricing strategy, applicable for an e-commerce company. The pricing strategy uses optimal prices for multiple products generated by the CPRM. We use draws from the posterior distributions to calculate expected profits and sales of a range of price positions, based on the lagged market price median, considering both the change in price elasticity and the distribution of the price elasticity. The price position that maximizes sales and profitability is determined to be the optimal price. Profits are computed as expected sales multiplied by profit margin of the product. Since we use the proposed framework to optimize prices of numerous different products, the framework is constructed in a way that it is applicable for all products when the appropriate data is available.

A field experiment with 9,800 products over 5 weeks is performed to assess the effectiveness of both the CPRM prices and the prices of a simplified Point Estimate-CPRM (PE-CPRM) method. The test is performed in a web shop, using prices corresponding to the strategy proposed in the paper. The results of the two strategies show the trade-off between sales and profit optimization and data storage capacity. In addition, we compare our results with results of a follow-stop pricing strategy to find the potential of this pricing strategy. The follow-stop pricing strategy follows the price of the cheapest competitor until the margins become negative.

The field experiment shows that the PE-CPRM pricing strategy increases sales and profit growth for low ticket items, compared to the follow-stop strategy, with only a small price increase relative to the follow-stop strategy. This shows the capacity of the PE-CPRM to find prices that can be raised effectively. Also, it outperforms the CPRM pricing strategy for low ticket items. The complex CPRM distributions might reduce effectiveness of calculating optimal prices compared to the more straightforward option to use point estimates, as the high variance in the posterior distributions of the parameters are more sensitive to extreme values. Therefore, the optimized prices can be of lower quality.

For high ticket items, sales and profit growth are reduced relative to the follow-stop price strategy. This can be caused by the fact that the follow-stop price strategy is superior for high ticket products, as we find high price sensitivities for these products and following the cheapest competitor in the market can be the best decision. Another explanation can be that our field experiment for high ticket products is polluted, since category managers changed prices of products during the field experiment.

The remainder of this paper is organized as follows. In the next section, relevant research on non-constant price effects is reviewed. The third section discusses data used in this research and in the fourth section we present the technicalities of the CPRM. In the fifth section we illustrate the price optimization approach and the set up of the field experiment. We present the results of our research in the sixth section and we conclude in the seventh section of this paper.

2 Previous Research on Non-Constant Price Effects

The price variable is often considered the most important variable in the marketing mix, since it is the only variable that generates direct revenue (Rao, 1984). Multiple studies have addressed the importance and complexity of good pricing strategies (Gijsbrechts, 1993; Rao, 1984). The complexity of the subject often forces academics to simplify the real world and in a lot of research, price elasticity is assumed to be a constant value. For extensive reviews of papers considering constant price elasticities, we refer to Tellis (1988) and Bijmolt et al. (2005). In this study, we focus on relaxing the restriction of constant price elasticity and modeling the sigmoidal demand curve (Phillips, 2005). For this reason, we provide a review of research conducted on non-constant price effects in this section.

Models such as the multinomial logit model (Nevo, 2000) and the Multiplicative Competitive Interaction Model (Bell et al., 1975; Fok et al., 2002) are used to estimate the sigmoidal curve in market shares. The sigmoidal curve is a natural shape when considering market shares, since market shares approach a maximum of 100% as underlying modeled sales increase. The downside is that these models require market shares or sales data of all competitors on the market. This data is often unavailable to companies and these models are therefore not suitable for pricing strategies.

Literature on sales models, which are truly flexible in price elasticity, is relatively scarce. One paper that does address the concept of changing price elasticities is Fok et al. (2007). Their focus is on customer responses to different absolute sizes of price changes. The paper proposes a two stage Bayesian regression model, where the price effect is a non-linear function. The function allows three regimes that model the effect of a relative large price decrease, a relative small price change and a relative large price increase separately. The results show clear non-linearities in price effects, which emphasizes the importance of the subject.

Another method is proposed by Van Heerde et al. (2002). The research concerns price promotion effects with store-level scanner data. Van Heerde et al. (2002) allow a more flexible decomposition of price effects, more specifically, they use a non-parametric approach that models the effect of the magnitude of the discount. To estimate the parameters of this non-parametric function, they use a local polynomial regression (Fan, 1992). Again, their results showed significant improvement of the model fit, due to the non-linear decomposition of price effects. Considering these two research results, there is enough reason to believe price effects are non-linear and major improvement on understanding price effects is possible when further analyzing these non-linearities.

Our research is an addition to the literature on reference prices (Mazumdar et al., 2005). Research conducted on reference prices analyzes prices relative to competitor and previous prices in a market. In this research we use the competitor prices as reference prices by calculating the price position of our products. This paper extends the line of research in non-linear price effects and reference prices by thoroughly analyzing the changes of price elasticity when the price position of a product changes.

Table 1: Summary of all product categories, their corresponding products types, unique products and the average frequency of price changes per product in the category per month.

Product category	# Product types	# Products	Freq. Δ price
Home Improvement & Gardening	478	6,000	1.76
Sport & Leisure	266	4,000	0.98
Cooking, Dining & Houseware	305	6,800	1.64
Sound & Vision	231	5,800	2.06
Domestic Appliances	241	5,700	2.03
Computer & Games	177	8,300	1.60
Beauty & Care	237	9,900	1.24
Home Furnishing	144	5,300	1.57
Pet	148	3,500	0.78
Baby	140	6,300	0.82
Health & Intimacy	128	3,200	1.75
Jewelry, Watches & Accessories	53	2,600	0.44
Mobile & Tablet	53	4,200	1.33

3 Data

We use data provided by one of the biggest e-commerce companies in the Netherlands. The data concerns daily sales data at the product level over a period of 24 weeks. Additionally, data provided by Google Trends is used to indicate patterns in sales, which are not caused by the retailer. Competitor prices are provided daily by five comparison shopping engines. In this section we elaborate on the content of the data and how we employ the data in our research.

We divide the data at four aggregation levels. The first level is the *product category level*, the second level is the *product group level*, the third level is the *product type level* and the fourth level is the *product level*. At the product category level, the data is split into thirteen categories. Products groups are similar products and the product type level contains products which have the same functionality, but only differ in design or brand, e.g. 'toaster' or 'vacuum cleaner'. The product level consists of all unique products sold by the company.

Table 1 reports the different product categories and the number of products and product types contained by each category. The last column shows the average number of price changes each month per product contained by the category. We can clearly see a difference in the frequency of price changes per category. Domestic Appliances and Sound & Vision are the categories that change their prices most frequently. The prices of products in these categories are changed up to 5 times as often as prices in the Jewelry, Watches & Accessories category. The category Domestic Appliances contains over 63,000 price changes compared to only 6,300 price changes in the same period by Jewelry, Watches & Acces-

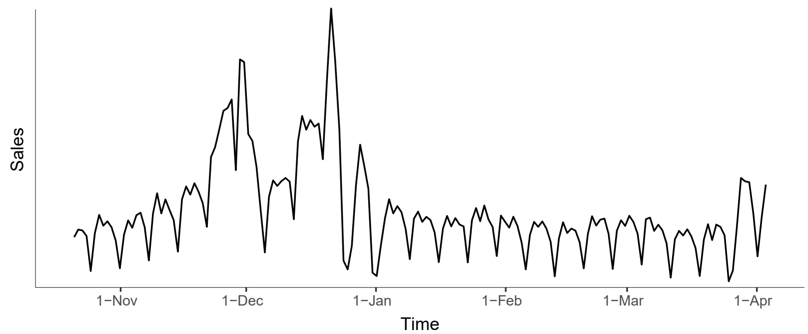

Fig. 3: Daily and holiday patterns in aggregated sales over all categories.

sories. Price elasticities and pricing analysis on the product level is impossible for products which have almost no price changes. The RES enables us to still analyze these price elasticities on the product level, while some categories do not have an average higher than 1 price change per product per month. We do not need a lot of price changes to evaluate price elasticities, since the RES allows us to use the information of other products.

Our research analyses 168 days of sales data on the product level. For every product and every day, sales, number of products sold, profits, variable costs and price are provided. The calculation of profits is based on variable costs and ignore fixed costs, so we use the difference between profits and sales divided by the number of products sold to determine these variable costs. Higher aggregation levels use the sum of the sales and profits, and the unweighted mean of the price positions. Weighted means do not take products without sales into consideration, which do contain valuable information.

Promotional data is not available. While Paap and Franses (2000) point out that effects of promotions are an important factor addressing price effects and Tellis (1988) warns for biases in estimation when omitting promotional variables, we do not expect problems in our setting. Most promotions in the data set are price discounts, which are analyzed by the price variable. These price discounts are almost never combined with promotions, since promotion strategy of the concerning e-commerce company is brand wide and SEA strategy is not based on data of prices. Furthermore, the promotion pressure of the companies brand is equal during the 24 weeks of our time period except for the holidays, which we correct with a dummy variable.

Figure 3 shows the aggregated sales over all categories. We see two clear patterns in the data. First, there is a decrease in sales every Saturday, since products can't be delivered on Sunday. Secondly, we observe a peak in sales during the Dutch holidays, Sinterklaas and Christmas, which are in the period of November 22th until December 2nd and the 13th of December until the 23th of December. We correct both of these patterns in our model by adding a corresponding dummy variable.

In addition, the data contains patterns that can not be explained by the explanatory variables in our model, such as popularity of a product or an advertisement campaign by the producer of the product. These patterns are different for each product, but can be found by using Google Trends. The Google Trends data is an index of the search queries to a specific term relative to other days. The day that had most search queries to the term has a Google Trends score of a 100. The Google Trends data corresponds to these popularity patterns in the data. By adding the Google Trends as an explanatory variable in the model, we correct for these patterns.

We use Dutch product type names to determine Google Trend scores in the Netherlands for each product type. Category names are too general to find a pattern, e.g. 'Domestic Appliances' as a search term does not represent patterns of the sales of the corresponding products. Individual products often do not have enough queries to find a pattern. Google Trends returns 1,146 trends for the 2,601 product type names, when the frequency of searches to a specific term is not sufficient, Google Trends does not return a value. To represent the missing trends, we use the mean values at the product group aggregation level. The average trend within a product category serves as a trend for product groups that do not have any products with a trend.

The competitor data consists of daily prices of all competing products of 2,500 competing online retailers. The resulting price data is enormous and very volatile. For this reason, we consider the median of the price of the product and its competitors per day and refer to this as the products market price. The volatility of the price is the reason we use the median of the prices rather than the mean. The volatility of the data is a result of a big number of outliers within the data. These outliers are caused by mistakes in the dataset of competitors or discounts of competitors. We do not want to take either of these cases into consideration for determining price position. The mistakes in the data do not describe the real world situation and considering big discounts of competitors would have a big effect on the price position, while the market position relative to most competitors does not change. By using a linear loss function and the corresponding metric, the median, we reduce the effect of these outliers compared to using the outlier sensitive quadratic loss function and the corresponding metric, the mean. In addition to the market price, we use provided competitor data to determine the minimal price in the market.

4 The Changing Price Response Model

In this paper, we propose a model that allows price effects to depend on price position. This results in a demand curve that has a sigmoid shape. The sigmoid demand function is an alternative to constant price elasticities, which are often used in price elasticity research (Bijmolt et al., 2005; Tellis, 1988). We use a hierarchical Bayesian approach to allow price effects to depend on price position in a second stage and to shrink the resulting high number of parameters within the model. In addition, the Bayesian framework allows us to assess uncertainty in the estimations. In the remainder of the paper, we refer to this model as the Changing Price Response Model (CPRM), because the changing price effect measures the change in customers response as price position changes.

4.1 Capturing Changes in Price Sensitivity

To model the changing price elasticity, we propose a model with a Multiplicative Sales Model (MSM) as a base. This model describes the sales of a product as a function of the price of the product, lagged sales, indicators for Saturday and holidays, a dummy to indicate whether a product is in stock or not and the Google Trends. The MSM allows us to estimate price elasticity as a parameter corresponding to the price variable. This model is often used to analyze price elasticities in marketing research (Zenor et al., 1998; Desmet and Renaudin, 1998).

The disadvantage of the MSM is that it estimates a constant price elasticity. In order to allow the price elasticities to depend on price position and create a sigmoid demand function, we add a second stage to the model. This second stage models the price elasticity as a function of the price deviation from a point M, where M is the market price or a fraction of this market price. To find the best value for M for every product, we include it as a parameter of the second stage regression.

A problem of the MSM is a result of the logarithmic transformation of the sales. The high number of products offered in the store results in products which are not sold every day. Zero values of a logarithm do not exist, so the MSM does not work for those data points. We do not offer a solution for this problem in this paper, since it is outside the scope of this research. All products that did not sell for 50% of the time period or more are aggregated to the product type level or product group level, if necessary. The remaining zero values are replaced by a value close to zero. Replacing the zero values with values close to zero should not change the model or its economical interpretation a lot and therefore we do not expect a big bias in the results.

We estimate the CPRM for a couple of different replacement values and determine the best replacement value by comparing the forecasting power of the estimated models. The test is conducted on the Domestic Appliances product category. The replacement value of the model with the highest forecast power is used for estimation of the models for the other categories. By separating the data used to determine the substitute of zero values and the data used to evaluate forecasts in later sections, we avoid over fitting the data.

We compute the price position of a product by dividing its price with the market price. This allows us to analyze effects of prices relative to the competitors, rather than the effect of solely the price. This procedure has two other advantages, first, we are able to aggregate the prices of products which are of the same product type. This aggregation can be performed because the price variable is transformed into a price position index, which is comparable for different products. We aggregate the prices by taking the average price position.

Secondly, we neutralize price patterns within the market. The resulting price position variable is defined by function

$$P_t = \frac{Pr_t}{MP_t},$$

where Pr_t is the absolute price at time t and MP_t is the market price at time t. The market price is the median of the price of the product and all competitor prices.

The sales of the products are modeled in the MSM. We regress the sales of product i, $S_{i,t}$, on the lagged sales, average price position $P_{i,t}$ and add dummies $D_{1,t}$ and $D_{2,t}$, that indicate the Saturday and the holidays, since these represent sales bottoms and peaks, respectively. Furthermore, we add a dummy $D_{3,i,t}$ to the model that indicates whether product i is in stock, because delivery time of products that are not in stock increases and these products are therefore less likely to be bought. In case products are aggregated, we use the average of $D_{3,i,t}$. We include the lagged Google Trends of product type p, $gt_{p,t-1}$, where p is the product type of product i. We use lagged values of Google Trends, since the present value can not be used for forecasting, as it is unknown. Lastly, Day (1981) argues that the product live cycle influences sales. This effect is not described by the Google Trends, because Google Trends are on a product type level. Research has shown that ignoring this effect of time causes a bias in the estimation of price elasticity (Tellis, 1988). We add a time variable to the regression at the sales level to avoid this bias. This results into a multiplicative regression formula described by

$$S_{i,t} = \exp\left(\alpha_{1,i} + \varepsilon_{i,t}\right) S_{i,t-1}^{\rho_i} P_{i,t}^{\phi_{i,t}} \beta_{1,i}^{D_{1,t}} \beta_{2,i}^{D_{2,t}} \beta_{3,i}^{D_{3,i,t}} gt_{p,t-1}^{\beta_{4,i}} t^{\beta_{5,i}}, \tag{1}$$

with $\varepsilon_{i,t} \sim \mathcal{N}\left(0, \sigma_{\varepsilon_i}^2\right)$.

The price elasticity is allowed to change over time by adding a second stage to the model. We propose a second stage that corresponds to the sigmoid demand function. A non-linear function of the price position, $\Omega_{i,t}\left(P_{i,t}\right)$, is employed to model this shape. Finally, we expect price elasticity to change when the product is the cheapest in the market. We formalize these effect into the second stage of the regression,

$$\phi_{i,t} = \alpha_{2,i} + \Omega_{i,t}\left(P_{i,t}\right) + \beta_{6,i} D_{4,i,t}, \tag{2}$$

where $\alpha_{2,i}$ is the baseline price elasticity of product i, $D_{4,i,t}$ is a dummy that indicates if the product is sold for the minimal price in the market or the mean of this dummy in the aggregated case.

In order to model the sigmoid demand function, we use $\Omega_{i,t}\left(P_{i,t}\right)$ to create a price elasticity function as a special exponential function, which is shown in Figure 4. The minimal value of the function is at the point M_i. This is the point

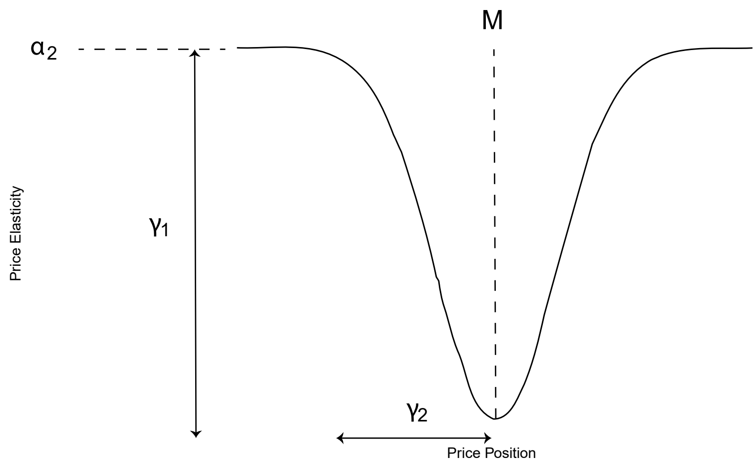

Fig. 4: Graphical interpretation of the effect of the parameters in Ω and α_2 on price elasticity.

where a product is most sensitive to price. In addition, we are interested in the shape of the sigmoid curve. The exponential function is therefore allowed to change in slope by parameter $\gamma_{2,i}$. We also allow the height of the function to change with $\gamma_{1,i}$. The function is formally described by

$$\Omega_{i,t}\left(P_{i,t}\right) = \gamma_{1,i}\exp\left(\gamma_{2,i}\left(M_i - P_{i,t}\right)^2\right), \tag{3}$$

where we expect $\gamma_{2,i}$ to be negative, since we expect price elasticity to increase as price position approaches M_i.

The resulting second stage consists of a big number of parameters, while data in price changes is limited. Additionally, not all products went out of stock in our time period and not all Google Trends varied over time. We use the RES to solve these three problems. Lastly, in previous research by Bekker et al. (2016), it proves hard to estimate changing price effects, because of the high degree of freedom given. The Bayesian framework allows us to add prior information to the model, reducing the degree of freedom in the model. In the next section we discuss how we use these Bayesian techniques to improve our model performance and solve those data problems.

4.2 Bayesian Approach

We use a hierarchical Bayesian specification to be able to shrink the estimations of the big number of parameters to more reasonable and less volatile estimations. In order to do this, we assume that the population location parameters of parameters on the product level are commonly distributed. The resulting common distributions are called the Random Effects Specifications (RES) of the

Table 2: Priors of all variables, an indication whether we added a RES and the prior of the location parameter of the RES, if applicable.

Parameter	Prior	RES	RES prior	Parameter	Prior	RES level	RES prior
σ_ϵ^2	flat	no		γ_2	$\mathcal{N}$	yes	$\mathcal{N}$
α_1	flat	no		M	$\mathcal{N}$	yes	U
ρ	flat	no		β_3	$\mathcal{N}$	yes	flat
β_1	flat	no		β_4	$\mathcal{N}$	yes	flat
β_2	flat	no		β_5	$\mathcal{N}$	yes	flat
α_2	$-\log\mathcal{N}$	yes	$-\log\mathcal{N}$	β_6	$-\log\mathcal{N}$	yes	$-\log\mathcal{N}$
γ_1	$-\log\mathcal{N}$	yes	$-\log\mathcal{N}$				

parameters (Cameron and Trivedi, 2005). Another advantage of the RES is that it allows us to use data points of other products, which is especially useful for products that have low or no volatility in the explanatory variables.

There are three more advantages to the Bayesian approach. First, it allows risks of pricing policies to be managed by analyzing the distributions of the sales forecasts. Secondly, it enables commercial experts to add prior beliefs to the model and lastly, it is easier to estimate, compared to the frequentist approach, that requires maximization of the maximum likelihood of the complex model.

Table 2 shows the prior distribution and an indication if we added a RES to the parameters, for all parameters in the model. Additionally, it shows the priors on the population location parameter of the RES, if applicable. All prior distributions of parameters, for which we did not provide a RES, have a flat prior specification. We don't have any prior beliefs on those parameters. That is,

$$p\left(\sigma_{\varepsilon_i}^2, \alpha_{1,i}, \rho_i, \beta_{1,i}, \beta_{2,i},\right) \propto \sigma_{\varepsilon_i}^{-2}.$$

We shrink the remaining parameters, by adding a RES to the parameter specification. This RES relates the parameters corresponding to the same explanatory variable, reducing variance in the estimation. See Fok et al. (2007) for a similar approach. Let $K_i = \{\kappa_{i,s}\}_{s=1}^{p}$ be the set of all parameters with a RES of product i, where the exact elements of K_i are found in Table 2 and p is the total number of parameters with a RES. Now we define the distribution of the parameters

$$\kappa_{i,s} \sim G_\kappa\left(\mu_\kappa, \sigma_\kappa^2\right), \qquad \forall \kappa_{i,s} \in K_i,$$

where G_κ is the distribution corresponding to κ_i, reported in the second column of Table 2. For each parameter in K_i, we define new parameters μ_κ and σ_κ^2, denoting the population location parameter and population scale parameter of the corresponding RES distribution.

We impose a negative log-normal distribution on α_2, γ_1 and β_6. In Equation 2 and Equation 3 we can see that this results in a strictly negative price effect. Positive price elasticities are inconvenient in pricing strategies, as they always put prices to infinity to maximize profits and sales. Intuitively, this is incorrect as sales always decrease if prices increase, ceteris paribus. Findings of Bijmolt et al. (2005) make us feel comfortable imposing this restriction, as they find that only 2.2% of price elasticities are positive.

However, a positive value for γ_1 can still result in a negative price effect. We choose to use a negative log-normal specification, as we want to analyze the sigmoid shape hypothesized by Phillips (2005). Other shapes are not in the scope of our research, and the model can reject the hypothesis of the sigmoid shape by either estimating a small value for γ_1 or estimating the values of M as a very high or low value. In addition, we limit the degree of freedom of the model by imposing this restriction. Since the model has a very high degree of freedom, we expect this to improve model performance.

Finally, the negative log-normal distribution for β_6 is chosen to reduce freedom in the model as well. As a result, our model can not deal with positive effects in price sensitivity when price becomes the cheapest in the market. We do not expect this to happen often, as price experts of our researched company stated that they always experience a higher sales increase when price becomes the lowest in the market. Furthermore, our model can still estimate a value close to zero for β_6, in the case that this hypothesis is false.

We impose informative priors on some of the newly identified population parameters μ_κ and σ^2_κ. For all population scale parameters σ^2_κ with $\kappa \in K$, we impose relatively uninformative priors. Hobert and Casella (1996) show that imposing a relative uninformative prior on scale parameters in second levels improves convergence and do not have a substantial effect on the results. An inverted Gamma distribution with shape parameter λ_1 and scale parameter λ_2 is a good candidate with limited influence, if parameters are chosen correctly. If we set both λ_1 and λ_2 as 0.1, the shape of the inverted Gamma distribution is nearly flat, resulting in a close to uniform distribution. This means that the prior is relatively uninformative and is therefore a good candidate. Formally, this results in the following prior,

$$p\left(\sigma^2_{\eta_\kappa}\right) \sim \text{IG-2}(0.1, 0.1), \qquad \forall \kappa \in K.$$

We impose proper priors on population location parameters on all parameters in the non-linear function $\mathit{\Omega}_{i,t}\left(P_{i,t}\right)$ and β_6, because the RES is the third level in the model. While parameters in a second level model have already been proven to be difficult to estimate (Bekker et al., 2016), we think adding prior information to reduce freedom in the model can highly improve estimation accuracy. All parameters that have negative log-normal distributions in the second level are assumed to have a negative log-normal distributed population location parameter as well. This results in a maximal value of -1 for the population mean of the parameters. We impose this prior to improve convergence and a value of -1 is reasonable for these parameters. The prior of the population location parameter of γ_2 is a normal distribution. To reduce the influence of the priors, we assume relatively uninformative priors on those population location parameters,

$$\begin{aligned} \mu_\kappa &\sim -\log\mathcal{N}\left(0,100\right) \qquad \kappa \in \left\{\alpha_2, \gamma_1, \beta_6\right\}, \\ \mu_{\gamma_2} &\sim \qquad \mathcal{N}\left(0,100\right) \end{aligned}$$

For μ_M, the population location parameter of the most sensitive price position, we impose a uniform prior on the price position interval [0.5, 1.5],

$$\mu_M \sim U\left(0.5, 1.5\right).$$

Again, we want to limit the freedom of our model. In our dataset, 96% of the 12 million price positions we study are inside this interval, which makes this interval interesting to study. Note that the prior is only imposed on the population location parameter, this allows parameter draws of the normal distribution on the product level to be outside the interval.

The remaining population location parameters do not have information added through priors, as the RES of these parameters are only in the second level of the model and we do not expect estimation difficulties. Formally, that is,

$$p\left(\beta_3, \beta_4, \beta_5\right) \propto 1.$$

To estimate the posterior distributions of the parameters, we use Markov Chain Monte Carlo (MCMC) techniques (Tierney, 1994). More specifically, we use the Gibbs sampling technique proposed by Geman and Geman (1984)[1] and the Metropolis-Hasting sampler which was developed by Metropolis et al. (1953) and subsequently generalized by Hastings (1970).

The Gibbs sampler requires the full conditional probability density function of all the parameter distributions. These distributions are of known form for all parameters, except for $\gamma_{2,i}$ and M_i. The distributions can be found in Greenberg (2012). The remaining parameters $\gamma_{2,i}$ and M_i have unknown full conditional posterior distributions. We estimate the posterior distribution with the Metropolis-Hasting technique. A detailed explanation of the Metropolis-Hasting technique is found in Chib and Greenberg (1995). We use a normally distributed candidate function and use the first 4,000 draws from the sampler to calibrate parameters of the candidate. The candidate is calibrated in such a fashion that it attempts to realize an acceptance rate between 20% and 40%.

4.3 Forecast Evaluation

The draws of the posterior distributions are used to forecast sales. These forecasts are compared to forecasts of a Bayesian MSM with constant elasticities. The

[1] For a thorough elaboration of the Gibbs sampler we refer to Casella and George (1992).

only difference between the CPRM and the MSM is the second level in the CPRM, where we allow price sensitivity to depend on price position. Insights in forecasting performance allows us to assess the added value of the changing price sensitivity specification. In addition, we use forecasting evaluation to determine the best substitute value for the remaining data points with sales values of zero, as described in Subsection 4.1.

We evaluate the forecast performance of the model by analyzing the predictive Bayes factor. This is a well established concept for model comparison (Geweke, 1994). Geweke and Whiteman (2006) present a thorough elaboration on the predictive Bayes factor, in this section, we recapitulate the technique for a one step ahead forecast, as we only use these in our paper. The purpose of our model is optimizing prices, and as the parameters of the model can be estimated every day, there is no need to forecast further in the future. In addition, we evaluate over 65.000 products, so we do not need more forecasts to increase reliability of our results. We use 167 days to estimate the parameters of the model and 1 day to evaluate forecasts. Although one could argue that a cross-validation method generates more reliable results and Kohavi (1995) argues that ten-fold stratified cross validation performs best in real-world datasets, we do not use those advanced techniques. Our dataset is extensive enough to return reliable results and the computational expensive method is not suitable for these high number of validation runs.

The predictive Bayes factor is a technique to compare the posterior predictive densities of two models. These posterior predictive densities represent the likelihood of future values, conditional on all available data and the model. Let $y_{i,T+1}$ be the sales of product i in future period $T+1$, Y_T the set of all sales data of all products from time 1 to time T, $\{y_{i,t}\}_{i=1}^{N}\ \forall t=1,\ldots,T$, where N is the number of products. We define X_T as the set of all explanatory variables of all products at time 1 to time T and $\theta_{i,A}$ as the set of all the parameters of product i and model A. To compare model A and B, we calculate the predictive Bayes factor,

$$PBF_{A|B} = \frac{p\left(Y_{T+1}|Y_T, X_T, A\right)}{p\left(Y_{T+1}|Y_T, X_T, B\right)} = \frac{\prod_{i=1}^{N} p\left(y_{i,T+1}|Y_T, X_T, A\right)}{\prod_{i=1}^{N} p\left(y_{i,T+1}|Y_T, X_T, B\right)}. \tag{4}$$

The predictive likelihood of model A for product i can be rewritten as

$$\begin{aligned} p\left(y_{i,T+1}|Y_T, X_T, A\right) &= \int_{\theta_{i,A}} p\left(y_{i,T+1}, \theta_{i,A}|Y_T, X_T, A\right) d\theta_{i,A} \\ &= \int_{\theta_{i,A}} p\left(y_{i,T+1}|\theta_{i,A}, Y_T, X_T, A\right) p\left(\theta_{i,A}|Y_T, X_T, A\right) d\theta_{i,A}. \end{aligned}$$

The Gibbs sampler provides us with draws from the posterior density of the parameters, $p\left(\theta_{i,A}|Y_T, X_T, A\right)$. We can use these draws to approximate the predictive likelihood,

$$p\left(y_{i,T+1}|Y_T, X_T, A\right) \approx \frac{1}{M}\sum_{j=1}^{M} p\left(y_{i,T+1}|\theta_{i,A}^{(j)}, Y_T, X_T, A\right)$$

where (j) denotes the j^{th} draw of the Gibbs sampler and M is the total number of draws.

The full conditional posterior distribution of $y_{i,T+1}$ is known for our CPRM and the MSM. Equation 1 shows the mathematical representation of $S_{i,T+1}$, that corresponds to $y_{i,T+1}$ in the example. This results in the distribution

$$y_{i,T+1}|\theta_{i,A}^{(j)}, Y_T, X_T, A \sim \log\mathcal{N}\left(\log \hat{y}_{i,T+1}^{(j)}, \sigma_{\varepsilon_i}^{(j)2}\right),$$

with

$$\hat{y}_{i,T+1}^{(j)} = \exp\left(\alpha_{1,i}^{(j)}\right) y_{i,T}^{\rho_i^{(j)}} P_{i,T+1}^{\phi_{i,T+1}^{(j)}} \beta_{1,i}^{(j)D_{1,T+1}} \beta_{2,i}^{(j)D_{2,T+1}} \beta_{3,i}^{(j)D_{3,i,T+1}} gt_{p,T}^{\beta_{4,i}^{(j)}} \left(T+1\right)^{\beta_{5,i}^{(j)}}.$$

We calculate $\phi_{i,T+1}^{(j)}$ using Equation 2, that is,

$$\phi_{i,T+1}^{(j)} = \alpha_{2,i}^{(j)} + \Omega_{i,T+1}^{(j)}\left(P_{i,T+1}\right) + \beta_{6,i}^{(j)} D_{4,i,T+1}, \tag{5}$$

where $\Omega_{i,T+1}^{(j)}\left(P_{i,T+1}\right)$ is the result of Equation 3 for draw j of the parameter distribution,

$$\Omega_{i,T+1}^{(j)}\left(P_{i,T+1}\right) = \gamma_{1,i}^{(j)} \exp\left(\gamma_{2,i}^{(j)}\left(M_i^{(j)} - P_{i,T+1}\right)^2\right).$$

The resulting predictive Bayes factor indicates the predictive performance of model A compared to model B. A predictive Bayes factor higher than 1 represents superior performance of model A and a predictive Bayes factor lower than 1 represents superior performance of model B.

5 Price Optimization

The proposed CPRM enables us to forecast sales in units and profits for different price positions of products and can therefore be used to determine optimal prices corresponding to these forecasts. In this section, we discuss the process of determining those prices and set up an empirical test to assess the quality of the CPRM pricing strategy.

5.1 Determining Optimal Prices

Because of the complex construction of the multiple stages in the model, an analytical solution for optimal prices using derivatives in infeasible. For this reason, we use a simulation method to determine optimal prices. First, we study the distributions of expected profits and sales for a grid of different prices. We use draws from the Gibbs sampler to simulate from these distributions. Secondly, we use the distributions of the expected profits and sales for all different prices to determine the optimal price. We employ an objective function to allow managers to balance profits and sales, as both are important targets.

E-commerce companies can change their prices with almost no time costs. This causes companies to continuously change prices. Since companies can alter their price the next day, we decide to optimize the price for a one day ahead forecasts.

First, we determine expected sales distributions given all data, as profits are easily computed when expected sales are known. For each draw j of the Gibbs sampler, we calculate the expected sales conditional on the model parameters. All draws are combined to create the unconditional posterior distribution of sales.

To decrease computation time of the algorithm and storage space required, we simplify the optimization of the sales to a sales multiplier. By the construction of the sales, $S_{i,t}$, described in Equation 1, we can see that the effect of the price on $S_{i,t}$ is fully captured by $P_{i,t}$ and $\phi_{i,t}$. All other variables and parameters are multiplied constants. These parameters and the error term can be ignored, as we only want to optimize the sales for the price position.

Let X_t be a collection of all explanatory variables used in the CPRM at time t and $\Psi_{i,t}$ a collection of all parameters in the CPRM of product i at time t. The draws from the expected sales multiplier distribution are calculated by

$$ESF_{i,t}^{(j)}\left(P_{i,t}\right) = \mathrm{E}\left[S_{i,t}^{*}\left(P_{i,t}\right)|\Psi_{i,t}^{(j)}, X_{i,t}\right] = P_{i,t}^{\phi_{i,t}^{(j)}},$$

where the $ESF_{i,t}^{(j)}$ is the expected sales multiplier forecast of product i at time t conditional on parameter draw j and all data, we removed the conditional notation to simplify notation in later parts of this section. The asterisk denotes a sales multiplier, rather than absolute sales and $P_{i,t}$ is the price position subject to optimization.

We have no forecasts of the market price, so we assume that it remains constant over time. Although this assumption is unlikely to hold, we do not expect market price to change a lot in one day. For this reason, we neglect the estimation bias caused by the omitted forecasts. $\phi_{i,t}^{(j)}$ is calculated similar to the method shown in Equation 5.

The expected profit $\pi_{i,t}$ given X_t is calculated by multiplying the margin with the expected sales,

$$\mathrm{E}\left[\pi_{i,t}\left(P_{i,t}\right)|\Psi_{i,t}^{(j)}, X_{i,t}\right] = \mathrm{E}\left[S_{i,t}\left(P_{i,t}\right)|\Psi_{i,t}^{(j)}, X_{i,t}\right]\left(Pr_{i,t} - C_i\right),$$

where C_i are the variable costs of product i and $Pr_{i,t}$ is the absolute price of product i at time t.

We can simplify the profits to a profit multiplier, by eliminating all irrelevant parameters, similar to the method we used to simplify sales to a sales multiplier. We determine the expected profit multiplier forecast, $EPF_{i,t}$, similar to the ESF,

$$EPF_{i,t}^{(j)}\left(P_{i,t}\right) = \mathrm{E}\left[\pi_{i,t}^{*}\left(P_{i,t}\right)|\Psi_{i,t}^{(j)}, X_{i,t}^{(j)}\right] = P_{i,t}^{\phi_{i,t}^{(j)}}\left(Pr_{i,t} - C_i\right),$$

After calculating all distributions of expected profit and sales multipliers for a different number of price positions, we determine the optimal price by selecting a price that optimizes an objective function. This objective function can be adjusted to optimize both the average expected profits and average expected sales, because both can be relevant in determining a price. To incorporate this multi component objective, we need to be able to compare both sales and profit multipliers. We do this by calculating maximum forecast sales multiplier and profit multiplier in our price grid and divide forecast sales multiplier and profit multiplier by these numbers.

Maximum profit multiplier is calculated by evaluating the average expected profit multipliers for all price positions on the price grid, that is,

$$EPF_{i,t}^{\max} = \max_{P}\left(\frac{1}{M}\sum_{j=1}^{M} EPF_{i,t}^{(j)}\left(P_{i,t}\right)\right).$$

The maximum sales multiplier is always at the lowest price position, as we impose a restriction for price elasticity to be negative and sales are in units. To determine the maximum sales multiplier, we calculate the ESF for the minimal price,

$$ESF_{i,t}^{\max} = \frac{1}{M}\sum_{j=1}^{M} ESF_{i,t}^{(j)}\left(P_{i,t}^{\min}\right),$$

where $P_{i,t}^{\min}$ is the minimal price position on the price grid.

In order to compare both revenue and sales, these two metrics are used to scale the averages of the EPF and ESF distributions and calculate EPF and ESF scores, denoted as $EPFS$ and $ESFS$ respectively, that is,

$$EPFS_{i,t}\left(P_{i,t}\right) = \frac{\frac{1}{M}\sum_{j=1}^{M} EPF_{i,t}^{(j)}\left(P_{i,t}\right)}{EPF_{i,t}^{\max}},$$

$$ESFS_{i,t}\left(P_{i,t}\right) = \frac{\frac{1}{M}\sum_{j=1}^{M} ESF_{i,t}^{(j)}\left(P_{i,t}\right)}{ESF_{i,t}^{\max}}.$$

Note that these metrics of the multipliers are slightly different from the metrics of the absolute sales and profits, caused by the correlation between the draws of the Gibbs sampler for different parameters. We do not expect a large difference, as the sales and profit multipliers are very correlated with the absolute sales and profits. Thus, the benefits of calculating all the expected sales and profits do not outweigh the costs of performing the exorbitant number of computations needed to compute the absolute values.

We combine these two metrics with two weights to create the objective function, subject to optimization,

$$\Gamma\left(\omega_1, \omega_2\right) = \max_{P}\left(\omega_1 EPFS_{i,t}\left(P_{i,t}\right) + \omega_2 ESFS_{i,t}\left(P_{i,t}\right)\right), \tag{6}$$

with $\omega_1 + \omega_2 = 1$. The price position that yields the highest value for $\Gamma\left(\omega_1, \omega_2\right)$ is determined to be the optimal price position. Managers can determine their own weights for ω_1 and ω_2, so price optimization corresponds to their chosen strategy.

One could extend the objective function to enable managers to reduce the risk of their price policy. This can be done by adding a metric of risk to the objective function, such as the variance in the profit and sales multiplier distributions.

Although the algorithm is a theoretical solid method, it requires a lot of storage capacity to store all the draws for these products. To limit the required storage, we propose a simplified method to the previously proposed simulation method. We calculate forecast loss and efficiency gain to assess the quality of the procedure. The simplified procedure uses the mean of the parameter distribution as a point estimation for the parameters, so we do not need to store all draws. Using a point estimate is an accepted method to simplify posterior Bayesian distribution results (Greenberg, 2012). We refer to this pricing strategy as the Point Estimate CPRM strategy (PE-CPRM strategy).

5.2 Field Experiment

Rather than only assessing the theoretical value of our model, we test our pricing strategy in a field experiment. In this field experiment we analyze and compare the effects of the CPRM price strategy and the PE-CPRM price strategy in the store of an e-commerce company. Figure 5 shows the set-up of the field experiment. First, we divide the products into three groups. Three different price strategies are used to price the products in these three groups, the CPRM strategy, PE-CPRM strategy and the follow-stop strategy, in which we always follow the cheapest competitor unless the profit margin is negative. Secondly, we evaluate resulting revenues, sales (in units), profits, objective score and the price indices in the test period. Objective scores denote the weighted percentage increase in sales and profits, corresponding to the weights of the objective function subject to optimization in the field experiment, described in Equation 6. We use percentage increase rather than absolute sales and profits, because these metrics can be compared.

The products are assigned by stratified sampling to each group (Cochran, 1953). We divide all products into four price ticket groups: low prices, medium-low prices, medium-high prices and high prices. In addition, we split the products

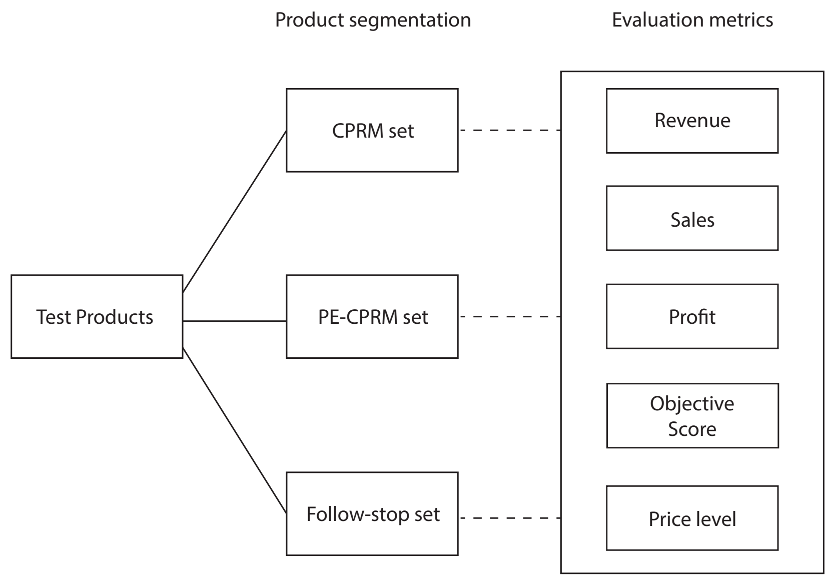

Fig. 5: Set-up of the field experiment.

into four revenue groups: low revenue, medium-low revenue, medium-high revenue and high revenue. The final result is a grid of 16 product groups. Each of the three samples contain an equal share of products from each group. Using this procedure allows us to compare results of the three samples, since our samples contain comparable products.

Neyman (1934) argues that sampling methods are not always perfectly representative. We correct this problem by evaluating our results with a difference-in-difference (DiD) evaluation. The DiD approach is a widely used method to correct for differences in groups, if pre-treatment information is available. We refer to Lechner (2011) for an overview of research conducted with DiD. Rather than comparing absolute values of profit, sales and revenues of the three strategy groups, we compare growth of the three groups relative to a benchmark period before our test. We show a graphical representation of this growth in Figure 6, where Δ represents the growth. To make the benchmark period comparable to the test period, our benchmark period consist of the same number of weeks as the test period. Additionally, our benchmark period is the period right before the experiment starts. This is the most comparable period, due to the high number of changes in the e-commerce business.

Let $S_{g,b}$ be the sales of strategy group g in the benchmark period and $S_{g,e}$ the sales of strategy group g in the test period. We can now calculate the difference Δ_g by

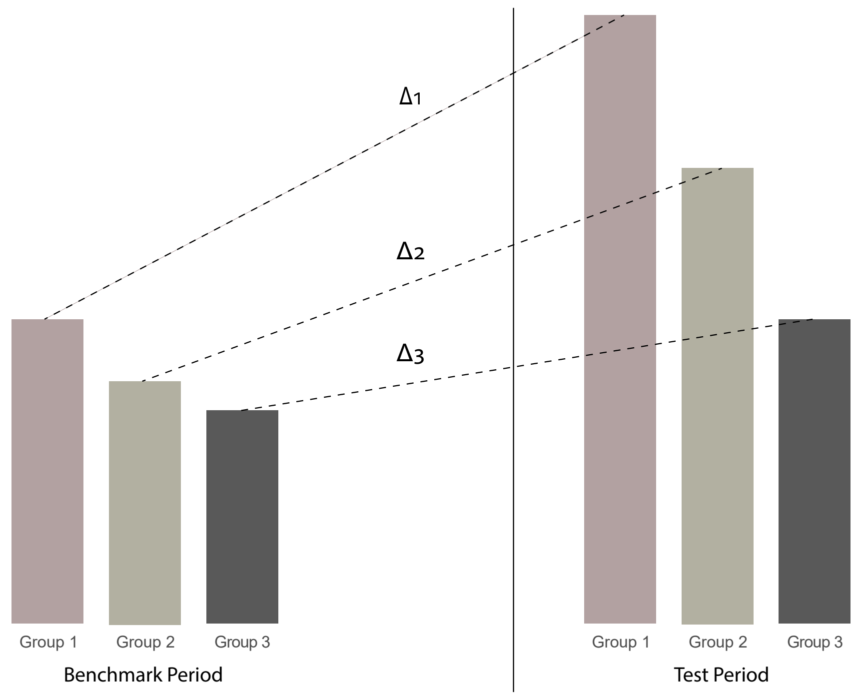

Fig. 6: Example of Difference-in-Difference evaluation of our field experiment.

$$\Delta_g = S_{g,e} - S_{g,b}$$

for all strategy groups. In addition, we compute the percentage change, $\Delta_{\%,g}$, and the average change per product per week, $\Delta_{p.p.p.w,g}$,

$$\Delta_{\%,g} = \frac{S_{g,e} - S_{g,b}}{S_{g,b}},$$

$$\Delta_{p.p.p.w,g} = \frac{1}{T}\frac{1}{N_g}\sum_{t=1}^{T}\sum_{i=1}^{N_g} S_{g,e,i,t} - S_{g,b,i,t},$$

where i denotes a product, t the week in the test or benchmark period, N_g is the total number of products in strategy group g and T is the length of the test and benchmark period in weeks.

Comparing the different changes gives us the DiD results. The results of the CPRM group compared to the follow-stop group show us the potential of the pricing strategy. The results of the PE-CPRM sample assess the trade-off between storage capacity and sales and profit optimization of using the simulation method compared to the point estimate method.

We analyze the effect of price changes during the test period to see what the effect of the strategies is on the prices. We assess these price changes by measuring the price index of the products in the strategy group. The price index is the average indexed change in price relative to the week before the experiment starts. Let $P_{g,i,t}$ be the price of product i in strategy group g at time t. We denote t_0 as the week before the test starts, the last week of the benchmark period. This is our benchmark for price changes. The price index, PI_g of a strategy group is calculated as follows,

$$PI_g = \frac{1}{T}\frac{1}{N_g}\sum_{t=1}^{T}\sum_{i=1}^{N_g}\frac{P_{g,i,t}}{P_{g,i,t_0}}.$$

The resulting price index is used to analyze the results of price changes, but also to confirm hypotheses about the sales and profit results in the next section.

6 Empirical Results

In this section, we discuss the results of the CPRM model. We estimate the parameters of our model for all product categories described in Section 3. The results are used to evaluate three subjects of marketing research. First, we analyze the effects of price position on price sensitivity, using the estimation results. Secondly, we investigate the forecasting possibilities of the model, calculating the predictive Bayes factor, and lastly, we asses the practical use of our model as a price strategy in a field experiment.

6.1 Estimation Results

We estimate the parameters of the CPRM using a combination of Gibbs sampling and Metropolis Hasting sampling, as described in Subsection 4.2. The results are based on 30,000 draws. Unreported plots of parameter draws show that the Markov chain is converged after 3,000 draws, so the first 3,000 are used as burn in. To remove autocorrelation in the Markov chain, we use a thin value of 10, which means that we only consider every tenth draw of the sampler. The analyses below only concern price effects, as all other effect are out of the scope of our research.

Before we estimate the parameters of the model, we solve a problem of the logarithmic transformation we employ. The logarithmic transformation does not allow for zero values in the sales, as described in Subsection 4.1. The CPRM is estimated for different replacement values for the zero values. To determine the replacement value with the best model fit, we assess forecasting power of the different estimated models for one day ahead forecasts of 5,657 products in the Domestic Appliances product category. The predictive Bayes factor is used to evaluate forecasts, as described in Subsection 4.3. We perform a t-test on the average predictive likelihood to test the significance of the difference

Table 3: Heat map of average price elasticities of all products per product category and price ticket group. Darker color represents a higher price elasticity. Elasticities are calculated at a price position of 1, excluding the minimal price effect. The cells that are empty do not contain enough products to produce reliable results. Averages denote the average price elasticity of all products in the category or price ticket group.

Product category \ Price ticket (€)	< 10	10 - 19.99	20 - 49.99	50 - 99.99	> 100	Average
Beauty & Care	-2.40	-2.67	-4.52	-7.96	-4.31	-2,99
Cooking, Dining & Houseware	-2.38	-2.91	-3.79	-3.81	-3.62	-3.36
Sound & Vision	-2.31	-3.04	-4.01	-5.32	-7.42	-5,07
Sport & Leisure	-2.45	-2.97	-4.20	-5.01	-7.96	-4.11
Jewelry, Watches & Accessories	-4.95	-8.45	-11.23			-6,14
Health & Intimacy	-2.42	-2.74	-2.97	-3.17		-2,66
Pet	-3.01	-4.16	-6.87	-5.61	-4.82	-4.92
Home Furnishing	-1.55	-3.17	-3.65	-4.32	-4.46	-3.66
Home Improvement & Gardening	-2.88	-3.44	-3.57	-6.07	-5.42	-4.28
Domestic Appliances	-1.99	-4.00	-5.71	-7.12	-9.39	-6,60
Mobile & Tablets	-1.66	-2.59	-4.02	-5.51	-7.32	-5,25
Computer & Games	-2.92	-3.71	-4.11	-8.73	-6.74	-5.16
Baby	-2.58	-3.10	-3.77	-4.12	-5.32	-3.31
Average	-2.47	-3.19 ***	-4.42 ***	-5.52 ***	-7.30 ***	-4.55

Note: We perform a t-test to test for significant difference in average price elasticity compared to the lowest price ticket group. *, ** and *** represent a significant difference in average price elasticity with a 0.05, 0.01 and 0.001 p-value threshold respectively.

in forecast performance for different replacement values. Because of the high computational costs of estimating the model, we only analyze the difference in forecasting performance for replacement values 0.01 and 0.001. We do not find any significant differences in the results, so we decide to use a replacement value of 0.001.

We start our analyses of price effects by summarizing the average estimated price elasticities of all product categories for a number of price ticket groups. The price elasticities are calculated at price position $M = 1$, because by definition, this is the price position at which products are most likely to be priced. We do not yet consider the increase of sensitivity caused by the minimal price position. With a price position of $M = 1$, it is unlikely that the product is the lowest price in the market, as this only occurs if almost all competitors have the same price. The average elasticities are reported in Table 3. The darker colors in the heat map represent a higher price elasticity and lighter colors a lower price elasticity. The last column and row state the average price elasticities of the corresponding product category or price ticket group. We use a t-test to test for a significant difference in average price elasticities for the different price ticket groups, compared to the lowest price ticket group.

Table 4: Estimation results of the second level parameters. The first half of the table shows the average posterior mean of all products of the estimated second level parameters in the CPRM per product category. In the second part we calculate points of interest, where the maximum price elasticity is the maximum increase added to the baseline price elasticity without considering minimal price effect, the price position multiplier denotes the ratio of the maximum price elasticity to the baseline price elasticity and the minimal price multiplier denotes the ratio of the price elasticity of products with minimal price effects to products without minimal price effects, at the midpoint price elasticity. The midpoint price elasticity is calculated as the average of the baseline and maximum price elasticity.

	Beauty & Care	Cooking, Dining & Houseware	Sound & Vision	Sport & Leisure	Jewelry, Watches & Accessories	Health & Intimacy	Pet	Home Furnishing	Home Improvement & Gardening	Domestic Appliances	Mobile & Tablets	Computer & Games	Baby
Coefficients													
Baseline	-1.69	-2.43	-3.20	-2.95	-4.98	-1.80	-3.34	-2.67	-3.12	-3.44	-3.27	-2.21	-2.48
Maximum increase	-3.79	-2.57	-4.66	-3.46	-3.05	-2.21	-3.07	-2.80	-2.87	-4.57	-4.84	-6.04	-2.48
Slope	-1.40	-0.71	-0.62	-0.50	-0.23	-0.61	-0.43	-0.44	-0.46	-0.68	-0.51	-0.96	-0.56
Most sensitive point	1.60	1.47	1.47	1.13	0.99	1.64	1.49	1.68	1.50	1.48	1.40	1.50	1.52
Minimal price effect	-1.50	-1.84	-2.13	-2.60	-2.77	-1.59	-2.36	-2.23	-2.43	-2.38	-2.63	-1.97	-2.05
Points of interest													
Maximum price elasticity	-5.48	-5.00	-7.86	-6.41	-8.03	-4.01	-6.41	-5.47	-5.99	-8.01	-8.11	-8.25	-4.96
Price position multiplier	3.24	2.06	2.46	2.17	1.61	2.23	1.92	2.05	1.92	2.33	2.48	3.73	2.00
Minimal price multiplier	1.42	1.50	1.39	1.56	1.43	1.55	1.48	1.55	1.53	1.42	1.46	1.38	1.55

There are three remarkable results in the heat map. First, we see a significant overall increase in price elasticity as price ticket increases. The average price elasticity for high ticket products is more than twice as high as the average price elasticity for low ticket products, more specifically the average price elasticity rises from -2.47 to -7.30. We hypothesize that this increase in price elasticity is caused by the fact that customers spent more time and energy in comparing prices for more expensive products and therefore become more price sensitive. Further research can improve our understanding of this subject.

Secondly, we find a clear difference in price sensitivity across product categories. More specifically, we find that the average elasticities differ from -2.66 to -6.60. This is in line with our expectations and emphasizes the importance of estimating models for different product categories separately to reduce estimation errors.

The last notable finding is that our estimated price elasticities are much higher than the reported average price elasticities of the survey studies on price elasticities. Tellis (1988) finds an average price elasticity of -1.76 and Bijmolt et al. (2005) find an average price elasticity of -2.62. This is substantially lower than our average price elasticity of -4.27 at price position $M = 1$. There are two explanations for this finding. First, we consider the products price position, so in our estimated price elasticity only the price elasticity at the market median is considered, whereas the research of Tellis (1988) and Bijmolt et al. (2005) consider all price positions. Other price positions than the median can result in lower price elasticities. Secondly, our data is gathered solely by an e-commerce company. Customers on the Internet are better informed because of easily accessible information (Bakos, 2001) and can therefore be more price sensitive.

To analyze the effect of price position on price elasticity, we summarize the average posterior means of the second level parameters of all products in a product category in Table 4. Most parameters are subject to a log-normal prior distribution and are therefore never zero. For this reason, we do not test for significant difference from zero. We report points of interests in the second part of the table, which shows the average maximum price elasticity, the ratio of the average maximum price elasticity to the average minimum price elasticity, and the ratio of the price elasticity of products with minimal price effects to products without minimal price effects, at the midpoint price elasticity. The midpoint price elasticity is calculated as the average of the baseline and the maximum price elasticity.

We find that price position has a strong influence on the price elasticity of a product. The maximal price elasticities of almost all product categories are twice as high as the baseline price elasticity. This confirms our hypothesis and the theory of Phillips (2005) that price elasticity changes as price position changes. Additionally, we see that both the baseline price elasticity and the increase in price elasticity are different across categories. Products in the Computer & Games product category have low price elasticities when their price position is low, while their price elasticities increase relatively fast, when their price position approaches 1.5. The opposite is true for the product category Jewelry, Watches

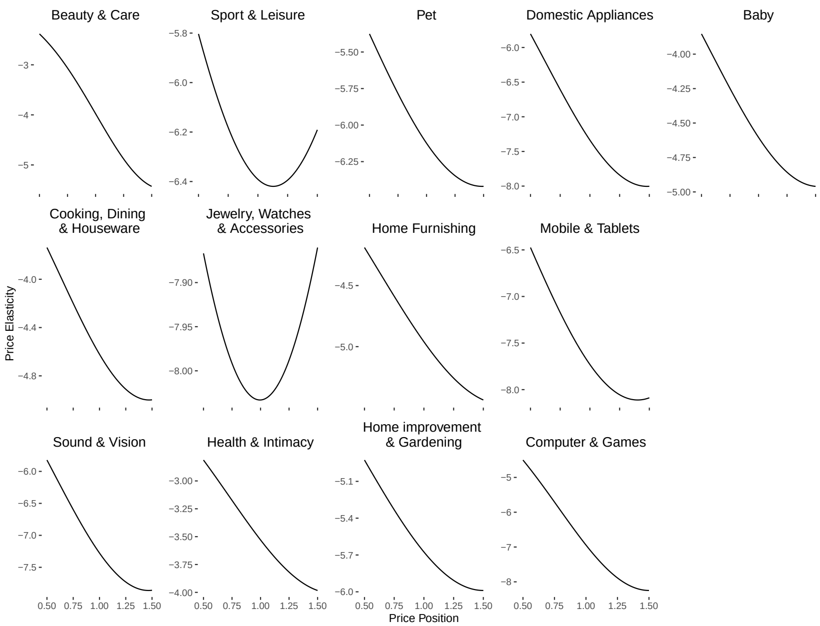

Fig. 7: Overview of average price elasticities of all product categories from a price position of 0.5 to 1.5, without considering minimal price effect.

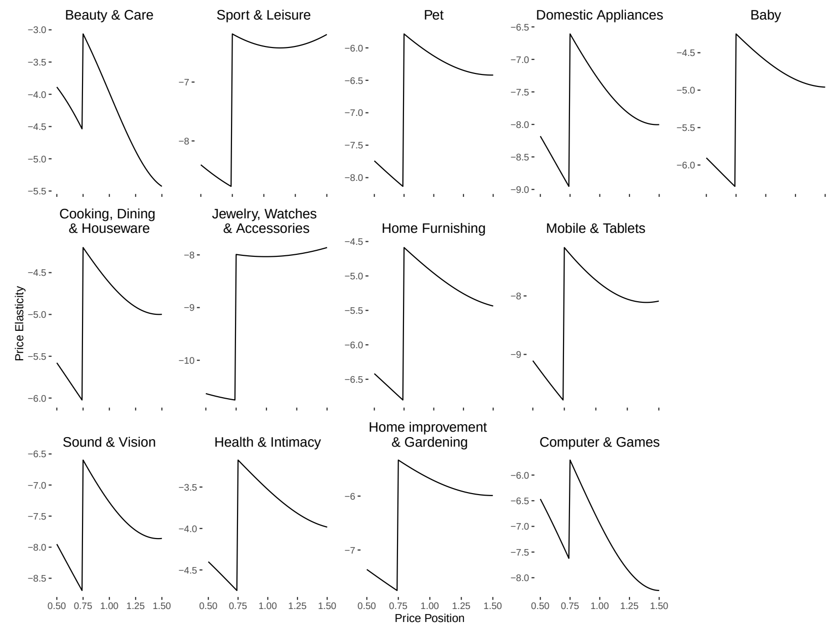

Fig. 8: Overview of average price elasticities of all product categories from a price position of 0.5 to 1.5, including a minimal price effect at a price position of 0.75.

& Accessories. Products in this product category have high price elasticities and price elasticities change relatively little when price position changes.

Contrary to our initial thoughts, the most price sensitive point in the market turns out to be at a higher price position than the market median for most categories. This means that in general, price sensitivity increases when price position increases. Figure 7 shows the average price elasticities of all product categories for price positions ranging from 0.5 to 1.5. We can clearly see the substantial change in price elasticity when price position increases, but we also see that, for most categories, this change is solely negative. Note that we did not consider the minimal price effect, which increases price sensitivity again when price decreases.

Our hypothesis that price elasticity of a product is higher when it has the lowest price in the market, the minimal price effect, is confirmed by our results. Average price elasticities increase between 39% and 56% when products become the lowest in the market. This means that customers are very sensitive to price comparisons and tend to buy products which have the lowest price in the market, regardless of the store that offers the product.

Figure 8 shows the average price elasticity of the product categories of price positions from 0.5 to 1.5, including the minimal price effect when price position is lower than 0.75. We can see that the change of price elasticity, caused by the minimal price effect, is considerable for all product categories. Relative to the change in price elasticity, caused by price position, the minimal price effect differs across product categories. Minimal price effect for products in the Sport & Leisure category is high relative to the effect of a change in price position. The opposite is true for products in the product category Computer & Games, which are relatively more effected by changes in price position.

To assess the uncertainty in our estimations, we analyze the standard deviations of the posterior distributions of the parameters in the second level of the model. Figure 9 shows the standard deviation of the posterior distribution of each product, for each parameter in the second level of the model.

An interesting finding is that the standard deviations of the log-normal distributed parameters (α_2, γ_1 and β_6) are high. This means that the data did not contain enough information to estimate the parameters with a high degree of certainty on the product level. This uncertainty is caused by complex specification in the second level of the model combined with the high degree of freedom. Although there is a high degree of uncertainty in the parameters on the product level, the conclusions made on the parameter estimations are valid, because of the high number of evaluated products. The uncertainty might cause problems for the models applications on the product individual level, such as forecasting.

For the most sensitive price position M, we also find relatively high standard deviations, as the population location parameter, corresponding to the population mean, is priori uniformly distributed on the interval [0.5, 1.5] and average standard deviation is 1.66. This means that the data did not contain enough information to find certainty in the distribution of M, possibly caused by the fact that the population mean of M was not in the interval [0.5, 1.5].

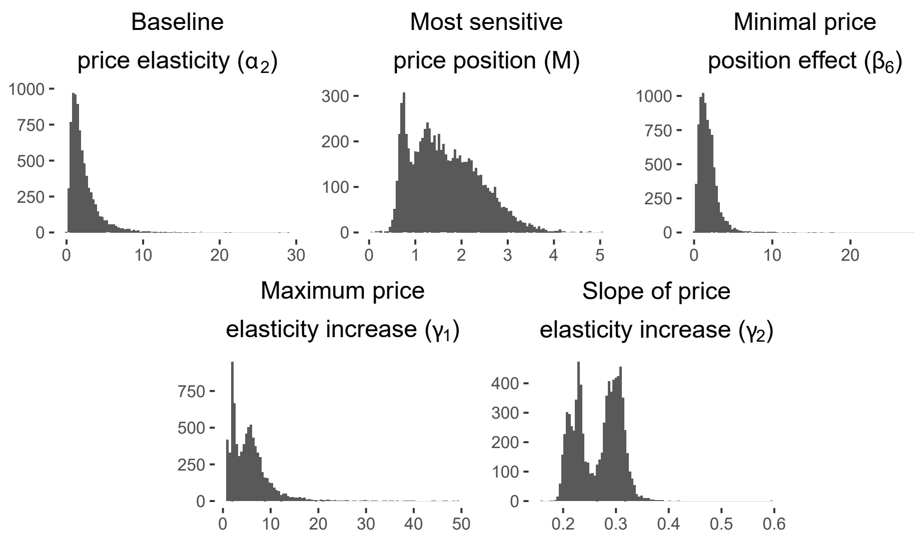

Fig. 9: Histogram of posterior standard deviations of the parameters in the second level of the model across products. Every standard deviation corresponds to one product in the model.

Figure 10 shows the posterior distributions of the population location parameters of the parameters in the second level of the model for all categories. One thing that stands out is the skewness of the posterior distribution of the population location parameter of the most sensitive price position parameter M. The prior distribution limits the possibility for a posterior population location parameter above 1.5, while the results show that the posterior population location parameter is probably higher than 1.5. This explains the high standard deviation in Figure 9. A higher value of M, would result in a more linear decrease in price elasticity as price position increases on the interval [0.5, 1.5]. Again, this is in line with our earlier findings of Figure 7, where we see that most average price elasticities increase as price position increases.

Given the results of Figure 7, Figure 9 and Figure 10, we reject our hypothesis of the U-shaped price elasticity as price position changes proposed earlier in this paper and conclude that, in general, there is a linear decreasing trend in price elasticity if price position changes. Our model is able to incorporate this linear decreasing trend by estimating a high value for the most price sensitive price position. Note that the limits on M do not necessarily cause problems for our model, as the slope parameter γ_2 can still model the linear downward trend. Rather than our hypothesized sigmoid shaped demand function, this creates a more concave demand function, possibly to reduce the convex shape of the first level, that is created by the MSM model shown in Figure 1. Further research could prove this hypothesis by estimating the CPRM with an additive first level, rather than a multiplicative level.

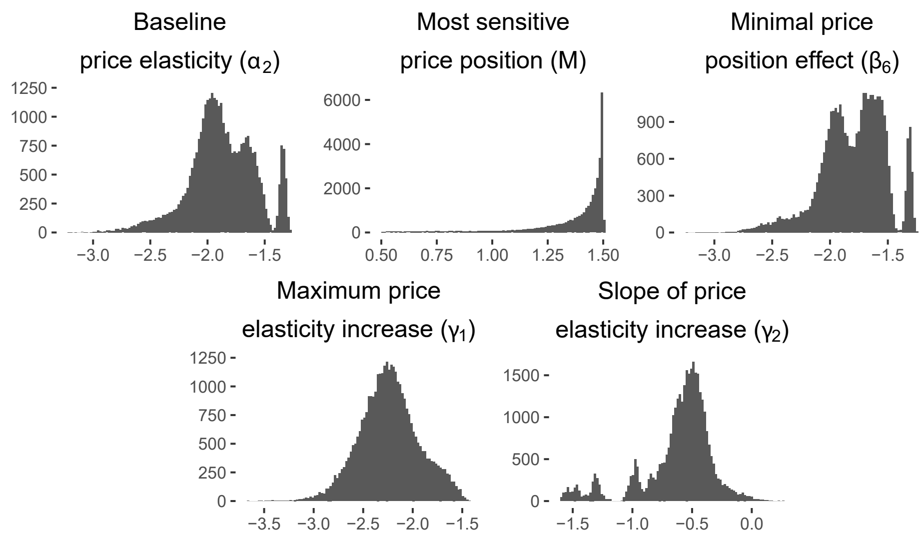

Fig. 10: Posterior distributions of the population location parameters of the parameters in the second level of the model for all categories

The negative linear trend in price elasticity as price position changes, when corrected for the minimal price effect, shows that a price decrease when price position is high, results in a relative bigger customer increase than when price position is low.

For the remaining parameters, we see multiple peaks at different values. This represents the difference across categories. These findings are in line with our earlier results, presented in Table 4.

6.2 Forecast Results

To determine predictive power of the changing price sensitivity specification, we create one day ahead forecasts with the CPRM and the MSM. The only difference between the models is the changing price sensitivity in the second level of the CPRM, so the difference in forecast performance represents the added forecasting power by the second level. We forecast sales of 65,900 products divided over all product categories, except for the Domestic Appliances product category. We use the forecasts of the product category Domestic Appliances to determine the replacement of zero values and therefore these forecasts could be subject to over fitting.

The MSM specification of the model can not forecast zero values, so we replace all zero values in the forecast sample with 0.001. This corresponds to the values used to estimate the parameters of the model.

Table 5 shows the average predictive likelihood of the forecast distributions of all individual products by the CPRM and MSM, the predictive Bayes factor and

Table 5: Forecast comparison of the CPRM and MSM. First two columns show the average predictive likelihood, third column shows the predictive Bayes factor and the final column shows the number of forecasts. The last row denotes the aggregated results for all products.

	CPRM	MSM	PBF		# Forecasts
Home Improvement & Gardening	7.29	7.79	0	***	6,000
Sport & Leisure	7.37	7.97	0	***	4,000
Cooking, Dining & Houseware	10.82	11.40	0	***	6,800
Sound & Vision	7.44	7.79	0	***	5,800
Computer & Games	4.45	4.78	0	***	8,300
Beauty & Care	6.21	6.59	0	***	9,900
Home Furnishing	13.25	13.74	0	**	5,300
Pet	8.64	11.19	0	***	3,500
Baby	7.77	9.11	0	***	6,300
Health & Intimacy	7.14	8.31	0	***	3,200
Jewelry, Watches & Accessories	11.27	10.80	∞		2,600
Mobile & Tablet	11.38	11.27	1,900		4,200
Total	8.17	8.77	0	***	65,900

Note: We perform a t-test to test for significant difference in average predictive likelihood. *, ** and *** represent a significant difference with a 0.05, 0.01 and 0.001 p-value threshold respectively.

the number of unique products per category. The bottom row shows the results for all products aggregated. The difference in the average predictive likelihood of the CPRM and MSM is tested for significance with a t-test, to determine the reliability of the results.

When we compare the predictive performance of the CPRM and the MSM, aggregated over all categories, we find a predictive Bayes factor of 0. This means that in general, the forecasts created by the MSM are better than the ones created by the CPRM. Furthermore, all categories show either a Bayes factor of 0 or an insignificant difference in predictive likelihoods. As a result, we conclude that the second level specification does not improve forecasting performance. This is caused by the fact that the second level has a complex specification for price effects, which is useful for analyzing theoretical effects of price position on price sensitivity, but not for forecasting purposes. One explanation for this problem is that we over fit the data. This occurs when the high degree of freedom in our model finds non general patterns, only existent in our data sample, but not in future observations. We refer to Babyak (2004) for a more elaborate explanation of overfitting. Another explanation is that the high degree of uncertainty in the posterior distributions of the second level parameters reduced the forecast accuracy. Further research can find other applications of competitor prices that do improve forecasting. In the next subsection, we discuss the result of the CPRM in a field experiment.

6.3 Field Experiment Results

The effectiveness of the CPRM price strategy is tested in a field experiment, where the optimized prices described in Section 5 are used as prices in the store of an e-commerce company. The results of this strategy are compared to the results of the PE-CPRM strategy and the follow-stop strategy. The test period consist of 35 days. The benchmark period is equal in size, to make the metrics comparable in our DiD approach. Prices are changed twice a week. We evaluate the metrics revenue, sales (in number units sold), profits and objective score. The objective score is the weighted percentage increase in sales and profits, corresponding to the objective function employed in the field experiment. In addition, we analyze the changes in average price indices as a result of the different price strategies.

The price grid used to find optimal prices consist of 100 price positions, uniformly distributed in the range of the lowest competitor price and the price 5% higher than the highest competitor price. This range has two advantages. First, we never price ourself lower than the market, avoiding a price war with competitors. Secondly, the strategy is able to increase the price higher than all competitors if deemed necessary, but we never increase price much above the market price to protect the companies price perception.

We divide our products in two groups, low price ticket and high price ticket groups. These price ticket groups contain products cheaper than €20,- and products equally or more expensive than €20,- respectively. The products are separated, because Table 3 shows a strong increase in price sensitivity as absolute price increases. It is interesting to see whether this influences the performance of the CPRM or not.

Our test group contains 9,824 products of the product category Domestic appliances. These products are distributed over the three strategy groups, using the method described in Subsection 5.2. In the test period, numerous price changes occurred that did not correspond to the pricing strategy. The prices are changed by the category managers, if they put an item in promotion, the supplier provides a special discount to lower the price or the product gets marked as dead stock. We call these price changes pollution of the test. All polluted products are removed from the test sample and we perform a quantitative analysis on the remaining products.

Fortunately, the unpolluted products are almost distributed similar to the segmentation prior to the test. The CPRM group contains 694 products, the PE-CPRM group consist of the same number of products and the follow-stop group holds 701 products. Evaluated metrics in the benchmark period are not equal among the different groups. Our employed DiD method corrects for this pollution.

The pricing expert of our studied company determined the weights in objective function to optimize prices, as these need to be in line with their own

Table 6: Change of revenue, sales and profit of test products. Results are split in groups for low ticket products (<€20) and high ticket products ($\geq$ €20). The first column shows the absolute change in euro or units, where revenue and profit are in euro and sales are in units, the second column shows the percentage increase and the third column shows average euro or units increase per product per week (p.p.p.w.). The objective score denotes the weighted percentages of sales and profit increase, corresponding to the objective function. The final column shows the number of unique products in each group.

	Revenue			**Sales**			**Profit**			Obj.	# unique
Test group	Increase	% Increase	p.p.p.w.	Increase	% Increase	p.p.p.w.	Increase	% Increase	p.p.p.w.	score	products
Low ticket products											
CPRM	$-2,298$	-1.9%	-1.58	-199	-2.7%	-0.14	842	6.0%	0.59	0.78	287
PE-CPRM	9,949	5.9%	6.17 **	513	4.7%	0.32 **	4,447	27.8%	3.11 **	13.91	303
Follow-stop	$-3,577$	-3.0%	-2.24	-384	-5.5%	-0.24	1,662	13.8%	1.04	2.20	319
High ticket products											
CPRM	$-39,745$	-8.6%	-19.44 ***	$-1,161$	-14.5%	-0.57 **	8,839	16.8%	4.35 **	-1.99	407
PE-CPRM	$-60,558$	-14.1%	-31.09 ***	$-1,222$	-18.4%	-0.63 ***	2,109	4.4%	1.22 ***	-9.25	391
Follow-stop	38,524	10.6%	20.17	-413	-6.3%	-0.22	17,147	34.7%	8.98	10.08	382

Note: Significance is tested for the average increase per product per week. We perform a t-test to test for a difference in average increase compared to the follow-stop strategy. *, ** and *** represent a significant difference in average increase with a 0.1, 0.01 and 0.001 p-value threshold respectively.

pricing strategy. Effects of these weights are left to further research. We use a weight for ω_1 of 0.4 and 0.6 for ω_2.

We begin by summarizing the resulting changes in revenue, sales, profit and objective score, where sales always denote sales in units, revenue and profits are in euro and the objective score denotes the percentage increase of both sales and profits, corresponding to the weights of the objective function used in the price test. We assess the effectiveness of optimizing the objective function with the objective score.

Table 6 shows the absolute increase in euro or units, the percentage increase, the average increase per product per week of our different metrics. In addition, it shows the objective score and the last column shows the number of products in each group. Significance is tested for average increase per product per week, as total increases can be highly influenced by one or two products. To limit these influences further, we evaluate both the average increase of revenue and sales per unit to determine significance. The significance is tested for the CPRM and PE-CPRM groups, compared to the results of the follow-stop group. The null-hypothesis is an equal average increase per product per week.

For the low ticket products, we find that the CPRM strategy does not significantly change the results of all metrics. However, the PE-CPRM strategy does improve all three metrics significantly compared to the follow-stop group. The PE-CPRM increases revenue growth by 9.9%, sales growth by 10.2%, profit growth by 14.0% and the objective score by 11.7. We did not expect the PE-CPRM to outperform the CPRM, as the CPRM uses more information to cal-

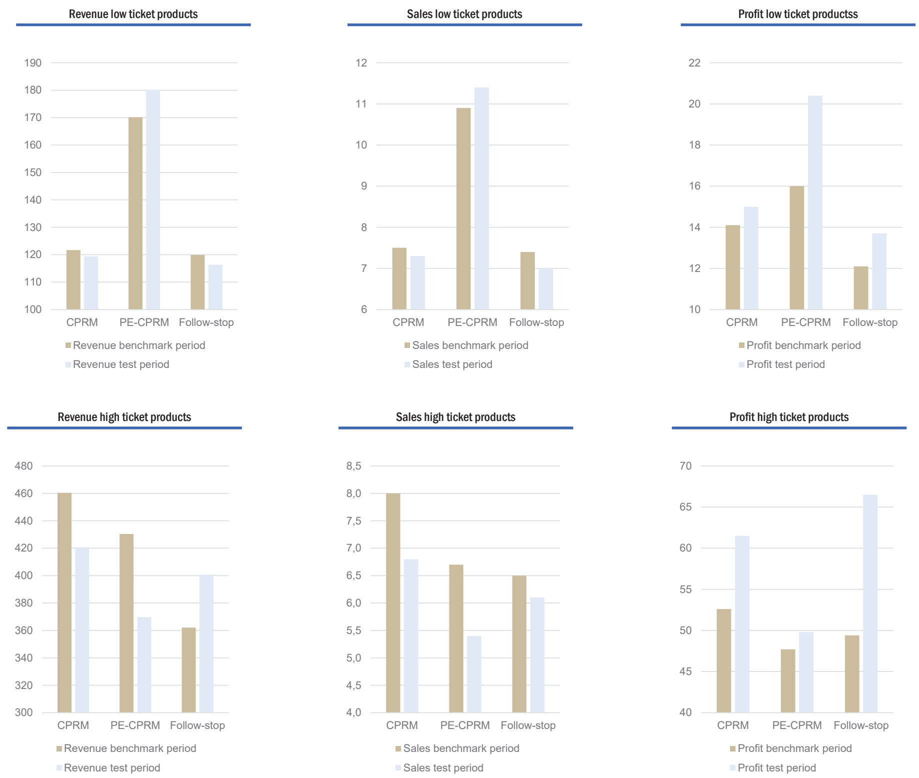

Fig. 11: Revenue, sales and profits of each test group in the benchmark and test period by ticket size. Reported numbers are in thousands.

culate optimal prices. Evidently, the complex information in the full distribution of the price elasticity does not improve the calculation of optimal prices. This is probably caused by the high variance in the posterior distributions of the parameters, which are more sensitive to extreme values. Therefore, the high variance can reduce quality of the optimized prices.

In the high ticket group, both the CPRM and PE-CPRM reduce growth in all metrics. Revenue, sales and profit growth are reduced up to 24.7%, 12.1% and 30.4%. The objective score is reduced up to 19.3 points. Two arguments can explain these results. First, the follow-stop strategy follows the cheapest competitor in the market, as long as the margin is positive. Because of the relative high price elasticity in the high ticket product group, this might always be the best strategy. Note that our strategy is able to incorporate a strong

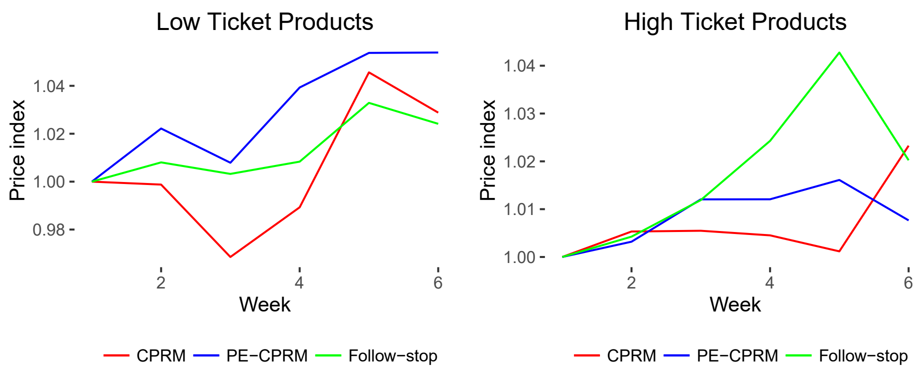

Fig. 12: Price indices of the CPRM, PE-CPRM and follow-stop test groups relative to week 1, the week before the test started. Price indices are shown for each week of the test period and for each price ticket group.

competitive strategy like this, by choosing a high weight on sales, rather than profit.

The second reason is the result of pollution in the test. We hypothesize that the prices of the high ticket products of the benchmark period of the follow-stop group are relatively cheap. This can only be caused by pollution, because the segmentation described in Subsection 5.2 also segmented the products into equally priced groups. The origin of our hypothesis is that we find that revenue and profits increase much more than sales in units. This indicates relatively cheap products in the benchmark period and an increase in price during the test period. Because the strategy does not change between benchmark and test periods for the follow-stop strategy, the cheapest competitor must have raised its prices during the test period or variable costs have risen, which is unlikely as profits have risen as well. This increase in price index results in a good performance of the follow-stop group, not caused by the strategy, which makes the DiD results of the CPRM and PE-CPRM poor. To test this hypothesis we look at our benchmark period.

Figure 11 shows revenues, sales and profits in the benchmark and test period for the CPRM, PE-CPRM and follow-stop strategies for the low and high ticket groups. As expected, we find low revenue for the benchmark period for high ticket products of the follow-stop group, combined with relatively high sales. The average price of sold products in the benchmark period is €59, compared to €68 and €80 for the CPRM and PE-CPRM groups respectively. The average price index of the benchmark period can only increase if the market price increases, which is not caused by the follow-stop strategy.

For the low ticket items, we find that nothing out of the ordinary appears, except for the profit increase of the PE-CPRM. This indicates the ability of the

Table 7: Percent point difference in average price index relative to the week before the test started. Difference is measured as price index of the row group minus price index of the column group.

	Low ticket products			**High ticket products**		
	CPRM	PE-CPRM	Follow-stop	CPRM	PE-CPRM	Follow-stop
CPRM	0	−2.44%. **	−0.86%.	0	−0.24%.	−1.00%. *
PE-CPRM		0	1.58%. *		0	−0.76%. *
Follow-stop			0			0

Note: We perform a t-test to test for a significant difference in average price index. *, ** and *** represent a significant difference with a 0.1, 0.01 and 0.001 p-value threshold respectively.

PE-CPRM to find prices which can be raised while increasing both revenue and sales in units.

To test our hypotheses for both low and high ticket products, we first look at the time line of the price indices during the price test, to find an indication whether the price indices differences are structural. Secondly, we look at the significance of the changes in price, to statistically confirm our hypotheses.

We show the movement in average price index of all products for the CPRM, PE-CPRM and follow-stop groups in Figure 12. As hypothesized, the PE-CPRM can raise prices of low ticket products, while improving the revenue, profits and sales. The latter is impressive and unintuitive. This result can be explained with the reasoning that the PE-CPRM raised products in price to increase profits and lowered some too high prices to also increase sales. We see that the PE-CPRM has a higher price index in all weeks, which indicates that the price increase of the PE-CPRM might be structural.

In addition, we find that the low prices of the follow-stop strategy for high ticket products in the benchmark period are raised over 4%. The only way the follow-stop strategy can raise its price index is if the cheapest competitor raises its prices. Additionally, this price index increase is an increase over time, which indicates that the cheapest prices in the market of the follow-stop strategy group steadily increased. This confirms our hypothesis that the pollution biased the test results.

The increase in price index also appears in the follow-stop group for low ticket products, however, the CPRM and PE-CPRM also show price index increases in the latter weeks, when the follow-stop price index increase starts. This shows that in the low ticket product groups, the pollution did not influence the segmentation prior to the field experiment negatively.

Finally, Table 7 shows the percent point difference in average price index per strategy group and price ticket, that is, the difference in average price index of all products in all weeks of the test period, compared to the other strategy groups. For the low ticket products, we find that the PE-CPRM raised its prices significantly more than the other strategies. For the high ticket products, we see that the price index of the follow-stop strategy increased significantly compared to the other strategies. Both confirm our earlier findings statistically.

7 Conclusion

Our paper provides new insights in the effects of competitor prices on price sensitivity. In addition, we test the hypothesis of a sigmoid demand curve, hypothesized by (Phillips, 2005), using our proposed Changing Price Response Model (CPRM). We assess forecast performance of our model and we propose a price strategy, based on our model, that determines optimal prices based using information of competitor prices. The quality of these optimal prices is tested in a field experiment, where products are priced according to our proposed strategy for five weeks.

The CPRM is a hierarchical Bayes model with a multiplicative sales model as first level and a second level to model the price effects. We propose a non-linear second level, to model the sigmoid demand curve with respect to price position. We estimate the parameters of our model for 13 product categories. The parameter estimates are obtained using MCMC.

Our first findings are general findings on price sensitivity estimation. We find that the average price elasticity is -4.27, much higher than the average price elasticity of -2.62 found by Bijmolt et al. (2005). This is likely caused by the fact that our research only concerns e-commerce, where price elasticity might be higher than traditional offline retailers, because of better information streams. Furthermore, we find that price elasticities of expensive products are higher than cheaper products. The average price elasticities differ between -2.47 and -7.30, for lowest and highest price tickets respectively.

Our most remarkable result is an unexpected one. Rather than the U-shaped price elasticity curve we expected, we find that price elasticity, in general, when corrected for the minimal price effect, decreases almost linearly when price position increases. This means that we do not find evidence of the hypothesized sigmoid demand curve. Furthermore, this shows that returns of price decreases for products with a higher price position are relatively higher than returns of price decreases for products with a lower price position. Our model is able to model this linear decrease, because of its high degree of freedom in the modeling of price effects through the second level. In addition, we find a considerable minimal price effect. Price sensitivity for products with the lowest price in the market can increase up to 56%.

The forecast results of the CPRM are not more accurate than the forecast results of the MSM. Therefore, the added second level, which is the only difference between the two models, does not improve forecasting performance. Although we do not improve sales forecast, we can still use the complex specification of our model to study price effects and determine optimal prices.

The field experiment provided noteworthy results. First, the results of the PE-CPRM strategy for low ticket products show a 10% increase in growth in both sales and revenue compared to the follow-stop strategy. Profit growth increases by 14% compared to the follow-stop strategy. In addition, the average price index of the products following the PE-CPRM strategy increases by 1.58% compared to the follow-stop strategy, without reducing sales. The PE-CPRM is likely able to both lower prices of products that are too highly priced and higher prices that are priced too low.

Surprisingly, the CPRM strategy did not show significantly improved results for the low ticket prices. This is probably caused by the high complexity and variance of the distributions of price elasticities in the CPRM, reducing effectiveness of the strategy.

For the high ticket products, we find a negative performance of both the CPRM and PE-CPRM. This result can be explained by two reasons. First, the follow-stop strategy might be superior for high ticket products, as these are relatively price sensitive and always following the cheapest competitor in the market can be the best strategy. Secondly, we find pollution in our test, possibly influencing the benchmark and therefore the outcome of our field experiment. To eliminate the last option, another field experiment has to be conducted in a non polluted environment.

A limitation of our research is the high number of zero values in the sales data. These are transformed to 0.001, biasing our parameter estimates and forecasts. A new research can transform our proposed CPRM to a model with either a additive first level or a logit level to model zero values. The first option can also test our hypothesis, that the linear decrease in price sensitivity as price position increases is estimated to reduce the convex demand curve which the multiplicative specification creates.

Another limitation are the weights of our objective function to optimize prices, described in Equation 6. We did not study the effects of these weights and a pricing strategist determined these for our field experiment, while this can have a big impact on the results. Future research can assess the effects of changing these weights on the performance of our model.

Furthermore, one might want to consider the effect of promotions on price sensitivity and the effect of competitor prices on the effect of promotions. Previous research, such as Fok et al. (2007), have already shown the significant interaction between price elasticities and promotions. Additional information on those interactions in a competitive context as proposed in this paper, can lead to high competitive advantages and useful insights for marketeers.

A lot of research on pricing strategies concern learning algorithms (den Boer, 2013). Incorporating learning algorithms in our framework could prove a useful extension to our model, as the high variety of new price points can improve our understanding of the complex price sensitivities and improve the performance of the CPRM.

References

M. A. Babyak. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*, 66(3):411–421, 2004.

Y. Bakos. The emerging landscape for retail e-commerce. *The Journal of Economic Perspectives*, 15(1):69–80, 2001.

W. Bekker, A. R. Hartveld, T. van Noort, and A. Osinga. A new approach to the detection of product introduction success for supermarkets. Unpublished manuscript, Erasmus university Rotterdam, 2016.

D. E. Bell, R. L. Keeney, and J. D. Little. A market share theorem. *Journal of Marketing Research*, 12(2):136–141, 1975.

T. H. Bijmolt, H. J. v. Heerde, and R. G. Pieters. New empirical generalizations on the determinants of price elasticity. *Journal of marketing research*, 42(2): 141–156, 2005.

P. Boatwright, R. McCulloch, and P. Rossi. Account-level modeling for trade promotion: An application of a constrained parameter hierarchical model. *Journal of the American Statistical Association*, 94(448):1063–1073, 1999.

A. C. Cameron and P. K. Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005.

G. Casella and E. I. George. Explaining the gibbs sampler. *The American Statistician*, 46(3):167–174, 1992.

S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335, 1995.

W. G. Cochran. *Sampling techniques*. John Wiley & Sons, 1953.

E. Cope. Bayesian strategies for dynamic pricing in e-commerce. *Naval Research Logistics (NRL)*, 54(3):265–281, 2007.

G. S. Day. The product life cycle: analysis and applications issues. *The Journal of Marketing*, 45(4):60–67, 1981.

A. den Boer. *Dynamic pricing and learning*. PhD thesis, Vrije Universiteit Amsterdam, 2013.

P. Desmet and V. Renaudin. Estimation of product category sales responsiveness to allocated shelf space. *International Journal of Research in Marketing*, 15 (5):443–457, 1998.

J. Fan. Design-adaptive nonparametric regression. *Journal of the American statistical Association*, 87(420):998–1004, 1992.

D. Fok, P. H. Franses, and R. Paap. Econometric analysis of the market share attraction model. *Econometric Models in Marketing*, 16(10):223–256, 2002.

D. Fok, P. H. Franses, and R. Paap. Seasonality and non-linear price effects in scanner-data-based market-response models. *Journal of Econometrics*, 138 (1):231–251, 2007.

D. Gefen. Customer loyalty in e-commerce. *Journal of the association for information systems*, 3(1):27–51, 2002.

S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

J. Geweke. Bayesian comparison of econometric models. Technical report, Citeseer, 1994.

J. Geweke and C. Whiteman. Bayesian forecasting. *Handbook of economic forecasting*, 1:3–80, 2006.

E. Gijsbrechts. Prices and pricing research in consumer marketing: Some recent developments. *International Journal of Research in Marketing*, 10(2):115–151, 1993.

E. Greenberg. *Introduction to Bayesian econometrics*. Cambridge University Press, 2012.

W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.

J. P. Hobert and G. Casella. The effect of improper priors on gibbs sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91(436):1461–1473, 1996.

C. Horváth and D. Fok. Moderating factors of immediate, gross, and net cross-brand effects of price promotions. *Marketing Science*, 32(1):127–152, 2013.

C. Kocas. Evolution of prices in electronic markets under diffusion of price-comparison shopping. *Journal of Management Information Systems*, 19(3): 99–119, 2002.

R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International Joint Conference on Artificial Intelligence*, 14(2):1137–1145, 1995.

M. Lechner. *The estimation of causal effects by difference-in-difference methods*, volume 4. Now, 2011.

T. Mazumdar, S. Raj, and I. Sinha. Reference price research: Review and propositions. *Journal of marketing*, 69(4):84–102, 2005.

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

A. Nevo. A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics & Management Strategy*, 9(4):513–548, 2000.

J. Neyman. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *Journal of the Royal Statistical Society*, 97(4):558–625, 1934.

R. Paap and P. H. Franses. A dynamic multinomial probit model for brand choice with different long-run and short-run effects of marketing-mix variables. *Journal of Applied Econometrics*, 15(6):717–744, 2000.

R. L. Phillips. *Pricing and revenue optimization*. Stanford University Press, 2005.

V. R. Rao. Pricing research in marketing: The state of the art. *Journal of Business*, 57(1):S39–S60, 1984.

B. T. Ratchford. Online pricing: review and directions for research. *Journal of Interactive Marketing*, 23(1):82–90, 2009.

G. J. Tellis. The price elasticity of selective demand: A meta-analysis of econometric models of sales. *Journal of marketing research*, 25(4):331–341, 1988.

L. Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, 22(4):1701–1728, 1994.

H. J. Van Heerde, P. S. Leeflang, and D. R. Wittink. *Flexible decomposition of price promotion effects using store-level scanner data*. Marketing Science Institute, 2002.

E. Van Nierop, D. Fok, and P. H. Franses. Interaction between shelf layout and marketing effectiveness and its impact on optimizing shelf arrangements. *Marketing Science*, 27(6):1065–1082, 2008.

M. J. Zenor, B. J. Bronnenberg, and L. McAlister. The impact of marketing policy on promotional price elasticities and baseline sales. *Journal of Retailing and Consumer Services*, 5(1):25–32, 1998.