ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis MSc Data Science & Marketing Analytics

The power of consumer-provided data: an analysis of the combination of customer reviews and online search trends in market response modeling.

Name student: Robin Zierlein

Student ID Number: 457166

Supervisor: Michiel van Crombrugge

Second assessor: Jeffrey Durieux

Date final version: 04-08-2023

Contents

1. Introduction 2
2. Literature review 4
  2.1 Market response models 4
  2.2 Search trend analysis 5
  2.3 Review analysis 6
  2.4 Conceptual framework 7
3. Data 10
  3.1 Datasets 10
  3.2 Variables 11
    3.2.1 Sales 11
    3.2.2 Price 11
    3.2.3 Review trends 12
    3.2.4 Search trends 13
    3.2.5 Seasonality 14
  3.3 Descriptive statistics 15
4. Methodology 18
  4.1 Method 18
  4.2 Assumptions 18
  4.3 Model & variable selection 19
  4.4 Final model specification 19
5. Results 21
  5.1 Topic interpretation 21
  5.2 VAR results interpretation 22
  5.3 Predictive performance 24
6. Conclusions 26
References 28
Appendices 31

1. Introduction

In many markets, consumer preferences change over time. For example, while 20 years ago consumers did not place a big emphasis on sustainability when buying a car, nowadays this feature has gained importance significantly. Because of these changes, businesses are consistently examining new ways to understand consumer preferences and predict market response as best as possible. With this information, companies can adapt their offerings and shift R&D budgets to create a product that suits customer needs perfectly. Moreover, market response information can be used to invest or save money when expecting positive or negative trends in their sales. As markets are rapidly evolving and competition is often fierce, being able to predict customer preferences and market response correctly can be a great asset to a company, possibly giving it an edge over competitors.

In recent years, online search trends are used increasingly to obtain these preferences and predict market response. For example, Du et al. (2015) used search trends from Google in order to gain insights in feature importance and predict sales. Since the publication of this article, leveraging search trends and other forms of readily available consumer data to gain insights has become a prevalent topic within marketing (Du et al., 2021; Lamberton & Stephen, 2016; Shankar et al., 2021).

While leveraging search trends as predictors in a model proved useful, this approach also had its limits. First, search trends would only be relevant predictors when consumers actively search for specific keywords related to a product. When this does not happen, search trends are expected to have low explanatory power. Second, keywords used by researchers to obtain search trends were often chosen subjectively (Du et al., 2015). This selection method may lead to bias, as possibly important keywords that can improve a market response model may be overlooked. A third limitation of the search trend approach is that searches for product features may not always perfectly align with feature importance (Du et al., 2015). Searching for something online is considered a low effort nowadays, which may lead to bigger effect sizes from search trends than in reality.

Aside from search trends, customer reviews have been used to gain consumer insights by marketeers more often as well (Floyd et al., 2014). For example, Archak et al. (2011) used text mining of customer reviews to incorporate these reviews in the choice modeling of product features for digital cameras and camcorders. The authors show that text of customer reviews can be used to model feature importance in a customer choice model (Archak et al., 2011). In another study, Floyd et al. (2014) found that online product reviews have a significant effect on retail sales. This shows that customer reviews also contain valuable insights into consumer preferences and market response.

However, a limitation of the method using customer reviews is that it’s often more suitable for vertically differentiated products compared to horizontally differentiated products, as the latter are usually not associated with differences in quality of features, but instead with different characteristics and personal preferences (Archak et al., 2011). Therefore, reviews may lack information about product information and results could be biased. A second limitation of the study by Archak et al. (2011) pertains to bias in reviews, as data of only one retailer was used, which could bias feature importance.

From the previous paragraphs, we can derive that both search trends and customer reviews may be viable in market response modeling, as both can act as a proxy for product feature importance and both metrics may influence sales. However, no present literature has tried to combine both constructs to create a possibly superior market response model. One could theorize that predictive performance could be enhanced when combining both search trends and customer reviews into one market response model. As a result, this research aims to add customer reviews to a market response model that utilizes search trends. With these ideas in mind, the central research question is:

To what extent can utilizing feature mentions of customer reviews enhance the predictive performance of market response models that employ search trends?

By combining both search trends and customer reviews to create a market response model, the present paper contributes to gaps in existing academic literature in multiple ways. First, this research aims to address limitations of the customer review approach by I) utilizing reviews to extract product features in a market of vertically differentiated products and II) combining review data with search trend data in order to eliminate bias in feature importance. Second, taking the limitations of the search trend approach into account, utilizing customer reviews may improve the method in three ways: I) by examining product features mentioned in reviews, one can construct an objective way of generating keywords used to obtain search trends, which may lead to more or different search trends used in analysis which may increase accuracy, II) consumer reviews may be a more valid indicator of product feature importance compared to online searches, as the mental cost of writing a review may be higher compared to searching online and III) by combining feature importance analysis from reviews with the regular feature importance analysis via trends, a more accurate market response model may be constructed overall.

The present paper can be considered both academically and socially relevant. The academic relevance can be derived from the previous paragraphs. Moreover, this topic is socially relevant as combining search trends with customer review data may lead to more accurate predictions of market response and product feature preferences. Utilizing this information, products can be tailored to customer needs as best as possible and a company’s capability to forecast sales may improve, possibly resulting in a stronger competitive position.

The remainder of the study is structured as follows: first, extant literature about market response models, search trends and review analysis is synthesized. Thereafter, the data used in the empirical analysis is expanded upon, after which the method used in the empirical analysis is justified. Logically, the chapter following the data and method covers the results from the analysis and includes some preliminary insights. In the final chapter, conclusions are drawn from the results and implications and limitations are discussed.

2. Literature review

The aim of this research is to incorporate and combine search trends and customer reviews into market response models. Therefore, the present chapter focuses on describing the current state of knowledge regarding these themes. Within this chapter, literature will be synthesized regarding market response models, search trend analysis and customer review analysis respectively. Moreover, the coherence between these three themes will be addressed, providing the basis for the rest of this paper.

2.1 Market response models

In general, market response models aim to predict a market response (e.g. sales, market share) based on several input factors (Neslin, 1990). One of the earliest works on market response models was done by Telser in 1962 (Neslin, 1990). Telser (1962) aimed to predict a company's market share for different products, for example coffee, based on the market share of a previous period and the price difference between two companies. Results indicated that the incorporated predictors of Telser's model held some explanatory power (Telser, 1962). Based on this approach, Telser could effectively predict market response based on price, competition and lagged market response (Telser, 1962). Some years into the future, Neslin (1990) expanded on Telser's model. This study focused on the effects of coupon promotions on a company's market share of instant coffee. Aside from the predictors from the research of Telser (1962), the author added retailer promotions and competitive couponing as predictors of market response (Neslin, 1990). Therefore, Neslin expanded the traditional market response model by adding a company's marketing activities as predictors. Results indicated this addition to be relevant in prediction of market response (Neslin, 1990). The relevance of marketing activities was supported further by a study from Dekimpe and Hanssens (1995). Here, the authors examined the long-term effects of marketing efforts on sales. Results indicated that different kinds of marketing had different kinds of short- and long-run impacts, but that marketing efforts in general had a strong relationship with sales (Dekimpe & Hanssens, 1995).

Taking the results from the previous literature into account, relevant predictors of market response so far have proven to be lagged market response and the pricing and marketing activities of both a company and its competitors. After the publication of Neslin (1990) and Dekimpe and Hanssens (1995), Ailawadi et al. (2001) examined the relevance of all these predictors further by analyzing the effect of long-term changes on market response for multiple products. The results reinforced the relevance of these predictors, as they were found to be relevant for analyzing long-term changes (Ailawadi et al., 2001). Another study by Danaher et al. (2008) examined the effect of competitive advertising on sales of packaged goods. Their market response model contained pricing and marketing activities of a company and its competitors as predictors. Results again showed that pricing and marketing activities of companies and competitors are valuable in market response modeling (Danaher et al., 2008). Aside from the aforementioned predictors, the paper by Ailawadi et al. (2001) also identified structural and external effects as predictors of market response; predictors that previous studies failed to mention. In the end, the above mentioned publications proved themselves to be relevant, as the predictors used in these studies were often found to be significant in many other markets such as clothing & apparel retail, cosmetics, groceries and ice cream (Dinner et al., 2013; Kumar et al., 2015; Kumar et al., 2017). Therefore, existing studies have substantiated that market response can be predicted through lagged response, price, marketing activities from a company and its competitors and structural & external effects.

2.2 Search trend analysis

After elaborating on the state of knowledge regarding market response models, present literature concerning the use of search trends in predictive modeling is harmonized to gain a better understanding of this specific topic.

While the market response models mentioned in the previous paragraph did not use search trends, many studies have already examined the effect of search trends on some form of market response, such as sales or market share. For example, Choi and Varian (2012) analyzed the effects of using search data from Google Trends as a predictor for automobile sales. Using an autoregressive model containing response lags and search trends about product names, their results indicated a significant improvement of model fit when including these search trends in the model (Choi & Varian, 2012). Within the same market, the predictive power of search trends was substantiated by Du and Kamakura (2012). In their study, the authors predicted sales based on a relatively bigger assortment of Google Trends searches (Du & Kamakura, 2012). The authors used a structural dynamic factor model that condensed many search trends into factors. In addition to finding a way to combine a multitude of searches into factors, Du and Kamakura (2012) again showed that utilizing search trends improved model fit. Aside from the automobile market, the use of search trends has proven relevant in predicting sales for multiple other markets, such as the housing market, retailing and travel (Boone et al., 2018; Wen et al., 2021; Wu & Brynjolfsson, 2009).

Aside from plainly using search trends to improve predictive performance, recent developments on market response models by Du et al. (2015) showed that search trends specifically can also be used to measure product feature importance, aside from only predicting market response. In their research, Du et al. (2015) were able to obtain search trends for different product features and use them in a market response model. Results showed that the model was more accurate when search trends of product features were incorporated, instead of only search trends about product names (Du et al., 2015).

Based on the previous paragraphs, one can conclude that previous literature suggests that search trends are a significant addition in the modeling of market response. However, as can be surmised from the previous paragraph, mentioned studies have taken different approaches in the exact use of these search trends in analysis. For example, Choi and Varian (2012) and Du and Kamakura (2012) only used search trends as predictors of market response. In contrast, Du et al. (2015) and Boone et al. (2018) also added variables concerning pricing, marketing activities or seasonality as predictors. Moreover, another difference between the mentioned studies pertains to the exact search trends used. Most literature only used keywords that were directly relevant to the product at hand (e.g. car, BMW, travel) (Boone et al., 2018; Du & Kamakura, 2012; Wen et al., 2021). However, Du et al. (2015) focused on specific product features (e.g. engine, fuel, interior). Finally, it is interesting to note that many of the mentioned studies selected search trend keywords subjectively (Boone et al., 2018; Choi & Varian, 2012; Wen et al., 2021). While this method may lead to more creatively found keywords that can improve model performance, some studies already noted that this method of keyword selection is arbitrary and possible better methods could be applied (Du &

Kamakura, 2012; Du et al., 2015). Through this arbitrary selection method, possible important keywords may be overlooked and irrelevant keywords may be included in the model, which in turn can bias results (Du et al., 2015).

In summary, current literature shows that search trends are promising predictors of market response indicators. While the use of search trends has gained increased attention in academic literature in recent years, the exact use in modeling is still relatively uncertain and subject to change.

2.3 Review analysis

Aside from search trends, another predictor that has gained increased attention in recent years is customer reviews. As the present paper wants to incorporate customer review data into a market response model that uses search trends, literature is synthesized concerning this topic.

Studies regarding the use of customer reviews to predict a market response indicator is readily available. However, many early studies focused on metrics that can be directly derived from the review (e.g. star-rating, volume, valence), instead of examining the actual text contained within. For example, research by Chevalier and Mayzlin (2006) investigated the effects of review star-ratings on book sales. Their results showed that the volume and average star-rating of reviews had a significant effect on sales (Chevalier & Mayzlin, 2006). Moreover, multiple studies focused on the effects of volume and valence of reviews on box office revenue (Dellarocas et al., 2007; Liu, 2006). Both concluded showing that valence and volume of customer reviews hold significant explanatory power (Dellarocas et al., 2007; Liu, 2006). However, while the mentioned articles seem to lay a strong foundation for volume and valence of reviews as relevant predictors of revenue, other studies contest these claims. For example, studies by Chen et al. (2004) and Duan et al. (2008) found review valence to have no significant effect on sales. Eventual meta-analysis by Floyd et al. (2014) synthesized all existing literature on this topic and concluded that both valence and volume of reviews affect sales significantly, with valence having the largest effect.

Some of the aforementioned research included other variables, aside from the review data, that typically make up market response models. For example, Chevalier and Mayzlin (2006) used price as a predictor variable, aside from the review data. Moreover, Liu (2006) incorporated production budget and lagged response as predictors of revenue. A similar approach was taken by Dellarocas et al. (2007), who incorporated production budget and marketing budgets in their model. These models already contain some similarity to market response models coined earlier in this paper, such as models from Neslin (1990) and Danaher et al. (2008). This already lays a foundation for the argument that customer reviews are a worthwhile addition to market response models.

From the previous section, one can derive that literature regarding effects of review metrics on sales is well represented. However, a study by Archak et al. (2011) argued that information captured in a review cannot be reduced into a single value such as a rating or valence. Instead, the authors decompose reviews into segments that contain information about product features (Archak et al., 2011), a comparable approach to the study of Du et al. (2015). This multifaceted approach resulted in a way to measure customers' preference for different aspects of a single product, going beyond the line of measuring preference for the product as a whole (Archak et al., 2011). Since the emergence of this study, this approach has been adopted by others. For example, while Fan et al. (2017) utilized this method to model automobile sales, Li et al. (2019) showed the relevance of the method for tablet sales.

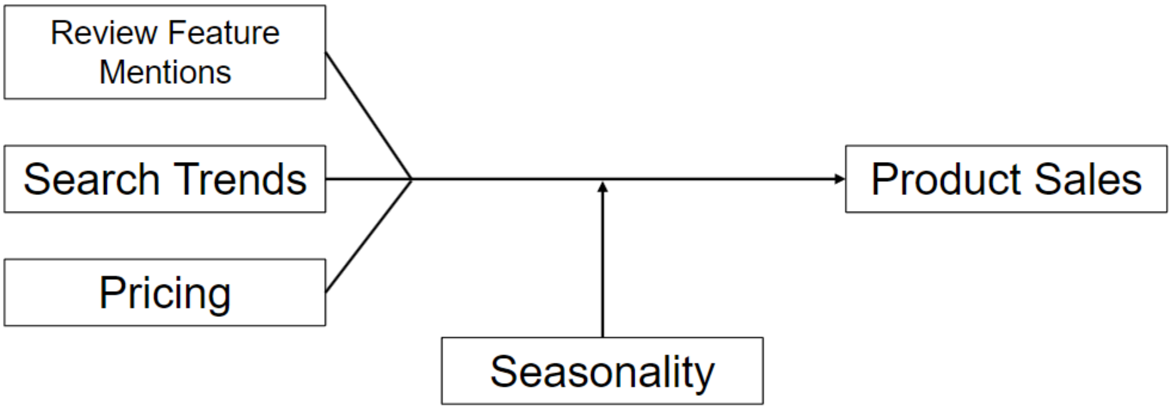
In summary, current literature suggests that customer review data can be leveraged as a relevant predictor of market response indicators. While earlier research focused mostly on metrics that can be directly derived from a review, later studies incorporated the actual text of a review to gain deeper and more diverse insights.

2.4 Conceptual framework

From the literature mentioned in the chapter thus far, many predictors of market response can be derived. Aside from the more common predictors typically found in market response models, such as price and marketing activities, this literature review also substantiated search trends and customer reviews to be relevant predictors of market response indicators. The following paragraph is dedicated to choosing which variables will be included in the market response model of this study.

As stated before, the aim of this research is to combine both customer reviews and search trends in a market response model. As a result, search trends and review feature mentions will be included in the models of this research. Aside from choosing these two data sources as input variables of our model, other variables will also be used in order to improve predictive performance and construct an academically valid model. Consequently, product pricing and seasonality will also be included in the models of this study, seeing as both variables have proven themselves to be of significant value in many studies mentioned in this review. While pricing will serve as an input variable, seasonality is considered a control variable as it may influence the relationship between search trends, review feature mentions, pricing and market response. Aside from deciding on relevant input and control variables, an outcome variable mimicking market response was chosen. Literature present in this review have often modeled market response through either a company's market share or their sales. This research utilizes sales as the outcome variable of the model, as it can be argued that sales are more directly influenced by predictors such as price and marketing activities compared to a company's market share. Summarizing, this study will examine the effect of search trends, review data, product pricing and seasonality on sales. A visualization showing the relationships between these variables is shown in Figure 1.

Figure 1: Visualization of the conceptual framework of the present study



In the remainder of this research, four models will be constructed:

1. A model with only pricing and search trends as predictors of market response. Search trends are chosen subjectively.
2. A model with only pricing and search trends as predictors of market response. Search trends are chosen by text mining of reviews.
3. A model with pricing, search trends and feature mentions in reviews as predictors of market response. Search trends are chosen subjectively.
4. A model with pricing, search trends and feature mentions in reviews as predictors of market response. Search trends are chosen by text mining of reviews.

To ease further reading of this paper, these four models will be referred to as model 1, model 2, model 3 and model 4, respectively. Based on the literature review, three propositions about the expected results of these models can be made:

p1: The model that incorporates only search trends, but are chosen by text mining, will outperform the model with only subjectively chosen search trends.

p2: The models that incorporate search trends and review feature mentions will outperform the models with only search trends.

p3: The model that incorporates search trends and review feature mentions and are chosen by text mining will have the best predictive performance.

The first proposition is substantiated through section 2.2 of the literature review. As concluded, subjectively chosen keywords could bias results, as important keywords may be overlooked and irrelevant keywords may be included. By utilizing review data as a basis for keyword selection, this bias can be eliminated, as keyword selection is based on customer data. The second proposition is substantiated through section 2.2 and 2.3 of the literature review. The significance of both search trends and review data as predictors of market response has been thoroughly substantiated by current literature. Therefore, a model that utilizes both search trends and review feature mentions as predictors will be expected to outperform a model that only uses search trends. The predictions of market response may improve, as more relevant data points are available. Lastly, the third proposition is substantiated through the combination of the previous propositions. As I) a model that uses objectively found trends is expected to outperform models with subjectively chosen trends and II) a model that uses review feature mentions is expected to outperform models without these feature mentions, a model that uses search trends and review feature mentions which are chosen objectively would have the relative best predictive performance.

3. Data

Based on the conceptual framework, the relationships between search trends, customer review mentions, price, seasonality and sales were tested. In order to do this, data regarding these concepts was gathered and an empirical context was selected that was suitable for the question at hand. As the present paper aimed to examine the effect of search trends on sales, it was necessary that the empirical context contained a product which features can be easily thought of and searched for online by consumers. As a result, the automotive industry was chosen as the empirical context of this research. The industry was chosen as cars have product features that are relatively easily identifiable by consumers, such as a car's handling, fuel consumption, interior or price. Moreover, purchasing a car is generally considered a high-involvement decision. A high-involvement decision is a buying decision that is characterized by a more extant process of information searching before the actual purchase. Buying a car is generally prefaced by a relatively big search for information and consideration, which makes it a good industry to investigate search trends in. Lastly, previous research has also utilized the automotive industry when examining the effect of search trends and review data on sales and found relevant insights (Choi & Varian, 2012; Du & Kamakura, 2012; Du et al., 2015; Fan et al., 2017). In the following paragraphs, the sources and descriptions of these datasets will be touched upon. Thereafter, relevant data pre-processing tasks will be substantiated and variables will be operationalized, after which some descriptive statistics are presented.

3.1 Datasets

This research utilized multiple external datasets: one containing car sales over time, one containing the Consumer Price Index (CPI) of cars over time, one containing customer reviews about cars and one aggregated dataset containing search trends for a multitude of keywords related to cars. It's important to note that this research used time series data for analysis. These datasets contain information paired with a specific date. Based on the date of observation, these different datasets could be combined to gain insights.

The first dataset was gathered from the Federal Reserve of Economic Data (FRED). FRED is an online database that aggregates time series data about economic activities. The used dataset contained monthly sales for light weight vehicles (e.g. cars and light trucks) in the United States. The time series contained 566 observations ranging from January 1976 until March 2023.

The second dataset contained the average monthly Consumer Price Index for light weight vehicles in the United States and was also sourced from FRED. This time series contained 364 observations ranging from January 1993 until April 2023.

The third dataset contained automobile reviews from customers between 2002 and 2018 and was sourced from Kaggle. The reviews pertained to five car brands: Audi, BMW, Infiniti, Lexus and Mercedes-Benz. Aside from the review text and the date the review was posted, the reviews also contained a star-rating ranging from one to five and the year a car was first manufactured. In total, the dataset contained 31,918 reviews.

The final datasets used in this research were self-constructed and sourced from Google Trends. Google Trends is a service that provides time series statistics about search queries made on Google.com. By inputting keywords, one can derive how much a keyword has been searched for over time. After inputting keywords, Google outputs a dataset containing the specific month and an index ranging from 0 to 100, indicating the search popularity of the combination of those keywords in that month. As this study aimed to construct different models based on subjectively chosen keywords and keywords chosen by keyword extraction from reviews, two different Google Trends datasets emerged containing different keywords. These datasets contained the amount of monthly searches for these keywords from January 2004 until April 2023. To ensure data quality, Google Trends was tuned to only count searches when they were in the explicit context of automobiles. This way, keywords such as “auto” will only count when the search is in the context of automobiles, and not in an unrelated context (e.g. automation).

3.2 Variables

From the previous section, it can be derived that the collected datasets contained time series data. Before performing empirical analysis, a suitable time dimension and date range had to be selected. As all datasets contained data from January 2004 until September 2018, this date range was chosen for the present study. All observations within the data pertained to one month in a specific year (e.g. 2012-02). Aside from identifying a date range, concepts mentioned in this paper so far needed to be operationalized into variables before adequate empirical analysis could be performed. Proper operationalization ensures that these theoretical concepts can be measured accurately and are reflected validly in the empirical research. Within the paper, the following concepts were operationalized: sales, price, review trends, search trends and seasonality. An overview of all resulting variables can be found in Table 1. The remainder of this paragraph will cover how the concepts of this study were translated into these variables.

3.2.1 Sales

As this research aimed to examine the effect of search trends and review feature mentions on sales, sales was selected as the dependent variable. In order to make this concept measurable, sales was represented by the monthly automobile unit sales over time from the dataset sourced from FRED. This operationalization was chosen, as many previous studies included unit sales as a proxy for market response (Danaher et al., 2008; Du et al., 2015; Kumar et al., 2017) and unit sales data was readily available. In this study, unit sales were measured in thousands of units sold and ranged from 9,223 to 21,135. No further data cleaning was necessary to use this data in analysis. The resulting variables containing the date of observation and monthly unit sales were named *date* and *sales*, respectively.

3.2.2 Price

After operationalizing the dependent variable, three independent variables were introduced. The first independent variable in this study was price. For the present study, price was operationalized as the monthly Consumer Price Index for motor vehicles. The CPI is an index indicating changes in prices consumers pay for goods or services. Because of this, the CPI for motor vehicles reflects changes in consumer prices for this market, creating a good proxy for price in the present research. The resulting variable was named *cpi* and ranged from 91,562 to 101,568.

3.2.3 Review trends

Aside from price, review trends were also used as an independent variable. In this study, the review data served two roles: I) use in keyword extraction and II) use as a direct predictor. First, keyword extraction was considered the process of obtaining words from the review data that can be used as keywords in Google Trends or to create review trends. Second, a review trend as a direct predictor was operationalized as the average number of mentions of specific keywords relating to a product feature or topic for a given month. More specifically, a review trend is considered a value ranging from 0 and upwards, indicating the popularity of a combination of keywords within reviews in a certain month. Higher values indicate more popularity for these keywords, which in turn may relate to a product feature or topic. The remainder of this subparagraph will first touch upon how keyword extraction was performed, after which the preprocessing to obtain the direct predictors is explained.

Due to the unstructured nature of text data, cleaning was necessary in order to identify and extract keywords more easily. First, all letters in the review text were set to lowercase and punctuation and stop words were removed to ease further cleaning. Second, Part-Of-Speech tagging (POS) was used to tag all words and their subsequent part of speech. POS is a method that can identify the part of speech of different words in a corpus. Here, POS was applied to more easily identify product features and keywords, as these words are often nouns (e.g. fuel, interior, price). Third, the dataset was filtered to only keep nouns. Moreover, the word "car" was also removed, as this word occurred in almost all reviews and could therefore bias results from keyword extraction. The final cleaned review data contained 26,132 reviews consisting of 677 different words.

Once the dataset was cleaned, topic modeling was utilized for keyword extraction. The application of topic modeling in this research was done through Latent Dirichlet Allocation (LDA). Latent Dirichlet Allocation is an unsupervised machine learning method that aims to find latent topics within documents, summarizing what topics people talk about in these given documents. LDA poses that a document consists of a set of multiple latent topics, which in turn consist of a set of words that represent these topics well (Blei et al., 2003). For example, dining reviews may contain topics such as food quality and waiting times. In turn, these topics are represented by words that relate highly to each topic. The word "food" may, for example, relate highly to the topic about food quality, but less to the topic about waiting times. LDA tries to find the probabilities of each word occurring within a certain topic, after which it estimates the general probability of each topic occurring within a document. LDA lends itself well to keyword extraction through the assumptions mentioned above. For example, one could interpret the topics as product features and corresponding words as keywords that relate to these product features.

LDA model selection was done by comparing models with different amounts of topics (k), namely 10 until 50 topics with increments of 5. These models were trained on a subset of 35% of the review data due to hardware limitations. Results indicated that perplexity and coherence favored models with 50 and 20 topics, respectively. Graphs showing the perplexity and coherence for different amounts of topics can be found in Appendix C. Based on these metrics, a final model with 30 topics was chosen as a middle ground between the two measures. The final LDA model was trained on 80% of the review data, with 20% being used as a validation sample. The model outputted 30 topics and, for each topic, the ten words that related the most to this topic. The topics and their respective keywords can be found in Appendix D. After this step, keyword extraction was completed, as keywords were extracted from the reviews that could be inputted in Google Trends.

However, several steps still needed to be undertaken to reshape the reviews as direct predictor variables mentioned in the operationalization. Two datasets with review trends have arisen in the present research: one with review trends for the subjectively chosen keywords and one with review trends for keywords found through keyword extraction. First, the researcher saved all product features, topics and their related keywords. Second, for each topic or product feature, it was calculated how many of the relevant keywords were present in each review. This resulted in a score for each review for each topic. For example, when a review mentioned many keywords about driving, but only few concerning fuel consumption, this review would have a relatively high score in the topic containing many keywords about driving and relatively low score for the subsequent topic about fuel consumption. After calculating these scores, all reviews and scores were aggregated based on their year and month of observation. For each month, the mean of the scores of all reviews within this month was calculated. This resulted in the operationalization defined at the start of the paragraph, namely the average number of mentions of specific keywords relating to a product feature or topic for a given month (hereafter: review trends). For aggregation, the mean was chosen over the sum to correct for the total volume in reviews over time, as some periods contained more reviews than others. Lastly, the review trends were normalized to the range of 0 to 100 in order to coincide with the range of Google Trends data. The resulting review trend variables of subjectively found keywords were named *srt_name, srt_price, srt_fuel, srt_interior* and *srt_driving.* The review trends variables resulting from objectively found keywords were named *ort_1, ort_2, ort_3 … ort_30*. Here, the prefixes *srt* and *ort* are an abbreviation for subjective review trends and objective review trends, respectively. Moreover, keep in mind that the objective review trend variables contain a number instead of a product feature, as objective keyword extraction only provided 30 topics with associated keywords and no topic names. In the results section of this paper, meaning will be given to the final trends used in the models by examining the keywords.

3.2.4 Search trends

The second independent variable of this study is search trends. Within the present study, a search trend was translated into the Google Trends search index for a combination of keywords regarding a certain topic or product feature for a given month. More specifically, a search trend is seen as a value ranging from 0 to 100 indicating the popularity of a combination of keywords on Google in a certain month. The same logic concerning review trends applies here: a higher value indicates more popularity for these keywords, which in turn relate to a product feature or topic.

Similar to the review trends, two Google Trends datasets have arisen in this research: I) A dataset containing the search index from subjectively chosen keywords and II) a dataset containing the search index from keywords found through keyword extraction by LDA. The selection of subjectively chosen keywords by Du et al. (2015) was a starting point for the selection in the present research. Here, a product feature of a car was chosen and related keywords were brainstormed and input in Google Trends (Du et al., 2015). Moreover, Google Trends itself contained sections recommending similar topics and search terms. These sections were considered to find usable keywords. The subjectively gathered product features and keywords can be found in Appendix B. Moreover, minor preprocessing was required to make both datasets usable in analysis. First, composite queries of keywords were made for each product feature (in case of subjectively chosen keywords) and topic (in case of keyword extraction by LDA) and were inserted in Google Trends. The five resulting datasets with subjective keywords and 30 datasets with objectively found keywords were merged together based on the date, resulting in 35 variables containing the Google Trends search index for a multitude of keywords. The resulting variables from subjective keyword selection were named *sst_name*, *sst_fuel*, *sst_price*, *sst_interior* and *sst_driving.* Variables containing search trends from objective keyword extraction were named *ost_1, ost_2, ost_3 … ost_30.* Similar to the review trends, these variables contain abbreviations *sst* or *ost,* indicating subjective search trends and objective search trends, respectively. Again, the objectively found search trend variables contain numbers instead of words indicating the topic, as LDA does not provide names of the created topics itself. This interpretation is done by the researcher in the results-section.

3.2.5 Seasonality

Aside from the dependent and independent variables, this study made use of one control variable, namely seasonality. Within the present study, seasonality is operationalized as the date an observation was made. Through incorporating the data, possible lags and/or moving averages into our analysis, seasonality can be controlled for. As mentioned, the date range of this research is January 2004 until September 2018. The value is captured in the variable *date*.

Table 1: Variable names, types and descriptions

Variable name	Type	Description
$sales_t$	Integer	Amount of car sales, measured in thousands of units at month *t*
cpi_t	Integer	Consumer Price Index for motor vehicles at month *t*
sst_name_t	Integer	Search trend index of keywords concerning names of cars at month *t*, keywords chosen subjectively
sst_fuel_t	Integer	Search trend index of keywords concerning fuel consumption at month *t*, keywords chosen subjectively
sst_price_t	Integer	Search trend index of keywords concerning prices of cars at month *t*, keywords chosen subjectively
$sst\_interior_t$	Integer	Search trend index of keywords concerning car interior at month *t*, keywords chosen subjectively
$sst\_driving_t$	Integer	Search trend index of keywords concerning a car's driving attributes at month *t*, keywords chosen subjectively
$ost\_1_t \ldots ost\_30_t$	Integer	Search trend index of keywords concerning topic 1 through 30 at month *t* found by LDA, keywords chosen as the ten most occurring words within each topic
srt_name_t	Integer	Review trend index of keywords concerning names of cars at month *t*, keywords chosen subjectively
srt_fuel_t	Integer	Review trend index of keywords concerning fuel consumption at month *t*, keywords chosen subjectively
srt_price_t	Integer	Review trend index of keywords concerning prices of cars at month *t*, keywords chosen subjectively
$srt\_interior_t$	Integer	Review trend index of keywords concerning car interior at month *t*, keywords chosen subjectively
$srt\_driving_t$	Integer	Review trend index of keywords concerning a car's driving attributes at month *t*, keywords chosen subjectively
$ort\_1_t \ldots ort\_30_t$	Integer	Review trend index of keywords concerning topic 1 through 30 at month *t* found by LDA, index was calculated by aggregating the average of all keyword mentions per topic per month

3.3 Descriptive statistics

All resulting variable names, types and descriptions can be found in Table 1. After all variables have been operationalized, some descriptive statistics can be presented and are displayed in Table 2.

Table 2: Descriptive statistics regarding sales, cpi, search trend and review trend data

Variable	Mean	Median	Min	Max	Std. Dev.
date	02-05-2011	01-05-2011	01-01-2004	01-09-2018	–
sales	15,672	16,661	9,223	21,135	2,457.76
cpi	97,612	98,271	91,562	101,568	2,950.60
sst_name	71.98	70.00	53.00	100.00	10.00
sst_fuel	39.89	39.00	22.00	100.00	10.64
sst_price	52.32	50.00	29.00	83.00	13.68
sst_interior	61.41	60.00	42.00	90.00	10.74
sst_driving	72.31	72.00	51.00	93.00	9.82
ost_1	59.44	58.00	44.00	82.00	8.43
ost_2	62.45	61.00	44.00	92.00	9.86
ost_3	44.22	44.00	27.00	80.00	8.67
srt_name	47.95	46.00	0.00	100.00	16.88
srt_fuel	44.89	44.00	0.00	100.00	17.72
srt_price	33.44	31.00	0.00	100.00	19.05
srt_interior	33.89	32.00	0.00	100.00	17.74
srt_driving	23.19	22.00	0.00	100.00	11.06
ort_1	40.89	41.00	0.00	100.00	18.33
ort_2	26.65	23.00	0.00	100.00	15.81
ort_3	31.45	31.00	0.00	100.00	15.11

When interpreting the results from the descriptives, one must keep in mind that these averages, minima and maxima are specifically across the time period January 2004 until September 2018. Observing the statistics, multiple insights can be derived. First, it is interesting to note that the review trends have relatively lower averages compared to the search trends. This can be explained, as the reviews work with a relatively smaller sample compared to the search trends. This leads to keywords possibly being mentioned less often in reviews compared to in Google searches, leading to a lower average. Second, it is interesting to note that some search trends and all review trends have the same maximum value of 100. This can be explained, as Google Trends calculates an index from 0 to 100 to indicate the popularity of a combination of keywords. Moreover, the review trends were specifically rescaled to be on the scale of 0 to 100, always resulting in one observation that carries either the minimum or maximum value. For some search trend variables however, the point at which these keywords were most searched for (and thus had an index of 100) lay outside of the selected date range of this study. Because of this, some of these variables do not have the maximum value of 100.

Third, the subjectively found keywords about car names and driving attributes were searched for relatively the most on average across the selected date range, with mean values of 71.98 and 72.31 respectively. Fourth, within the date range provided, subjective keywords about driving attributes have occurred the least in reviews, with a mean value of 23.19. Fifth, from all search trends, the trend concerning car pricing had the highest standard deviation. The same can be said about the review trend concerning pricing. These relatively high values pose the argument that pricing is the product feature that had the most fluctuation in importance over the date range provided. Moreover, the review trends all had relatively high standard deviations as well. This is again supported by the fact that the review data spans less consumers compared to Google Trends. As a result, less data is available to calculate the average over a month, possibly leading to bigger differences in averages across different months.

While some interesting insights can already be derived through the descriptive statistics, further analysis is necessary in order to answer the main research question posed by this study. Consequently, the following chapter will elaborate on the research methods used in further analysis.

4. Methodology

After having operationalized the concepts of this study and having preprocessed the required data, the method used to answer the central research question is touched upon.

4.1 Method

This study aims to answer its research question by constructing multiple market response models through Vector Autoregressive (VAR) models. VAR-models are models in which a dependent variable is modeled as a linear combination of both past values of this same dependent variable and past values of other independent variables. For example, stock prices may be influenced by past values of the stock itself, but also by inflation and its past values. Through this assumption, VAR-models allow other variables and their past values to influence the outcome variable. This makes VAR-models suitable in modeling more dynamic relationships between dependent and independent variables. Moreover, this assumption makes VAR a valid method for the present paper, as one could argue that automobile sales are not only influenced by past sales, but also by search trends and its past values. This argument seems logical, as car purchases are not always made directly after searching for cars online. In addition, literature mentioned in the previous chapters of this paper often used (vector) autoregressive models to model the relationship between sales and trend searches (Boone et al., 2018; Choi & Varian, 2012; Dekimpe & Hanssens, 1995; Wu & Brynjolfsson, 2009).

4.2 Assumptions

Before being able to specify the final VAR-models, multiple assumptions were checked and model & variable selection was performed. To obtain valid results, two assumptions relating to VAR-models were checked. The first assumption pertains to the stationarity of the data. This was checked by performing an Augmented Dickey-Fuller (ADF) Test. When the resulting p-value was 0.1 or lower, the null-hypothesis that the given time series was non-stationary was rejected. After performing ADF-tests, it was found that the null hypothesis could not be rejected for a multitude of variables, indicating non-stationarity. As a result, first-order differencing was applied to all variables used in the four models. Performing ADF-tests on the differenced data, the null hypothesis could be rejected for all variables and the stationarity assumption was fulfilled. The results of the ADF-tests both before and after differencing can be found in Appendix E. Aside from stationarity, VAR-models also assume that large outliers are unlikely. To control for this, all variables were checked. First, boxplots of all variables were observed to gain initial insights into potentially troublesome values. Boxplots of all variables in the data can be found in Appendix F. Second, outliers were quantified by performing standard deviation analysis. When a value was more than four standard deviations outside of the mean of a variable, this value was deemed a large outlier and removed from the data. This process for finding and removing outliers was based on prior work by Aguinis et al. (2013). Within the paper, the authors recommend a visual method followed by a more quantitative method to identify and eliminate outliers (Aguinis et al., 2013). As a result, the boxplot and standard deviation analysis were used to detect and remove outliers. The specific cutoff of four standard deviations was chosen based on the first step of outlier detection. All datasets contained at least two and at most three values that exceeded this threshold. Two observations were removed from data used in the first model, while three observations were removed from the data used in the latter three models. The resulting datasets for the four models contained 175, 174, 174 and 174 observations, respectively.

4.3 Model & variable selection

Aside from testing assumptions, variable selection was performed, as the second, third and fourth VAR-model all contained at least 11 predictor variables which could introduce overfitting. Variable selection was not performed for the subjectively found search trends, as the aim of these models was to provide a benchmark for reference. Observing that current literature did not perform variable selection for subjective search trends (Boone et al., 2018; Choi & Varian, 2012; Du et al., 2015), the present paper replicated this practice. For the remaining variables, selection was performed by performing Granger causality tests of an independent variable on *sales*. Only variables that led to a p-value of 0.1 or lower were kept in the final analysis, as only these variables granger-caused *sales* with some statistical significance. Results of the Granger causality tests can be found in Appendix G. After running Granger causality tests, the second, third and fourth model were condensed to 7, 8 and 8 independent variables, respectively. After performing variable selection, the optimal amount of lags for all four VAR-models were chosen. Four metrics were utilized to select this optimal amount. These metrics were Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Hannan-Quinn Criterion (HQ) and Akaike's Final Prediction Error (FPE). Reasons for using these four metrics is that HQ and BIC generally favor models that lead to more stable results, while AIC and FPE may favor models that do not guess the correct amount of lags, but do lead to low prediction errors (Ltkepohl, 2005). Therefore, a trade-off needed to be found between the optimal amount of lags suggested by all four metrics. After calculating all metrics, the preferred amount of lags based on each criterion can be found in Appendix I. Many of the four metrics favored either one or three as the optimal lag length. Based on these criteria, a lag length of one was chosen for all four models. This amount was chosen, as two of the four metrics indicated this lag length to be optimal for all four models. Moreover, choosing a lower lag length leads to less parameters being estimated by the models. With this knowledge, a lag length of one was chosen to reduce the possibility of overfitting.

4.4 Final model specification

As mentioned previously, four market response models were constructed for the empirical analysis. Taking the mathematical expression of VAR-models into account, the four final models can be written mathematically as:

Model 1:

$$ln\begin{bmatrix} sales_t + 0.01 \\ cpi_t + 0.01 \\ sst\_driving_t + 0.01 \\ sst\_price_t + 0.01 \\ sst\_fuel_t + 0.01 \\ sst\_driving_t + 0.01 \\ sst\_interior_t + 0.01 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_k \end{bmatrix} + \begin{bmatrix} \phi_{11} & \cdot & \cdot \\ \phi_{21} & \cdot & \cdot \\ \cdot & & \\ \cdot & & \\ \phi_{k1} & \cdot & \cdot \end{bmatrix} ln\begin{bmatrix} sales_{t-1} + 0.01 \\ cpi_{t-1} + 0.01 \\ sst\_driving_{t-1} + 0.01 \\ sst\_price_{t-1} + 0.01 \\ sst\_fuel_{t-1} + 0.01 \\ sst\_driving_{t-1} + 0.01 \\ sst\_interior_{t-1} + 0.01 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_k \end{bmatrix}$$

Model 2:

$$ln\begin{bmatrix} sales_t + 0.01 \\ cpi_t + 0.01 \\ ost\_1_t + 0.01 \\ ost\_3_t + 0.01 \\ ost\_4_t + 0.01 \\ ost\_21_t + 0.01 \\ ost\_25_t + 0.01 \\ ost\_28_t + 0.01 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_k \end{bmatrix} + \begin{bmatrix} \phi_{11} & \cdot & \cdot \\ \phi_{21} & \cdot & \cdot \\ \cdot & & \\ \cdot & & \\ \phi_{k1} & \cdot & \cdot \end{bmatrix} ln\begin{bmatrix} sales_{t-1} + 0.01 \\ cpi_{t-1} + 0.01 \\ ost\_1_{t-1} + 0.01 \\ ost\_3_{t-1} + 0.01 \\ ost\_4_{t-1} + 0.01 \\ ost\_21_{t-1} + 0.01 \\ ost\_25_{t-1} + 0.01 \\ ost\_28_{t-1} + 0.01 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_k \end{bmatrix}$$

Model 3:

$$ln\begin{bmatrix} sales_t + 0.01 \\ cpi_t + 0.01 \\ sst\_driving_t + 0.01 \\ sst\_price_t + 0.01 \\ sst\_fuel_t + 0.01 \\ sst\_driving_t + 0.01 \\ sst\_interior_t + 0.01 \\ srt\_fuel_t + 0.01 \\ srt\_driving_t + 0.01 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_k \end{bmatrix} + \begin{bmatrix} \phi_{11} & \cdot & \cdot \\ \phi_{21} & \cdot & \cdot \\ \cdot & & \\ \cdot & & \\ \phi_{k1} & \cdot & \cdot \end{bmatrix} ln\begin{bmatrix} sales_{t-1} + 0.01 \\ cpi_{t-1} + 0.01 \\ sst\_driving_{t-1} + 0.01 \\ sst\_price_{t-1} + 0.01 \\ sst\_fuel_{t-1} + 0.01 \\ sst\_driving_{t-1} + 0.01 \\ sst\_interior_{t-1} + 0.01 \\ srt\_fuel_{t-1} + 0.01 \\ srt\_driving_{t-1} + 0.01 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_k \end{bmatrix}$$

Model 4:

$$ln\begin{bmatrix} sales_t + 0.01 \\ cpi_t + 0.01 \\ ost\_1_t + 0.01 \\ ost\_3_t + 0.01 \\ ost\_4_t + 0.01 \\ ost\_21_t + 0.01 \\ ost\_25_t + 0.01 \\ ost\_28_t + 0.01 \\ ort\_4_t + 0.01 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \cdot \\ \cdot \\ c_k \end{bmatrix} + \begin{bmatrix} \phi_{11} & \cdot & \cdot \\ \phi_{21} & \cdot & \cdot \\ \cdot & & \\ \cdot & & \\ \phi_{k1} & \cdot & \cdot \end{bmatrix} ln\begin{bmatrix} sales_{t-1} + 0.01 \\ cpi_{t-1} + 0.01 \\ ost\_1_{t-1} + 0.01 \\ ost\_3_{t-1} + 0.01 \\ ost\_4_{t-1} + 0.01 \\ ost\_21_{t-1} + 0.01 \\ ost\_25_{t-1} + 0.01 \\ ost\_28_{t-1} + 0.01 \\ ort\_4_{t-1} + 0.01 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdot \\ \cdot \\ \varepsilon_k \end{bmatrix}$$

From these equations, one can derive multiple things. First, all models are specified in a log-log format. This was done so the effects of individual predictors could be easily interpreted as elasticities. Moreover, observe that the specification exactly is $ln(x + 0.01)$. This was done, as the data contained values of 0, which would transform into negative infinity without the added $+0.01$. Reasoning behind this method of eliminating log-zeros lies in the fact that adding a constant is considered a popular fix for this specific problem by literature (Bellégo & Pape, 2022). Moreover, a constant of 0.01 was chosen, as this value eliminates the problem whilst introducing relatively little differences to the original measurement scale (West, 2021). Second, one may observe that the first model only incorporates *sales*, *cpi* and all subjectively found search trends. The second model incorporates *sales* and *cpi* as well, but switches the subjectively found search trends for the search trends found through keyword extraction. The third model is similar to the first model, but also utilizes subjectively found review trends. Finally, the fourth model incorporates *sales*, *cpi*, objectively found search trends and objectively found review trends.

5. Results

After elaborating on the data and methods used in the present study, this section of the paper shows the process followed during the empirical analysis and its eventual results. As mentioned before, four different models were constructed during analysis. All four models contained *sales* as a dependent variable and *cpi* as an independent variable. Aside from this however, the models differ by using different independent variables, namely I) only subjectively found search trends, II) only objectively found search trends (through keyword extraction) III) subjectively found search trends and resulting review trends and IV) objectively found search trends and resulting review trends. As stated before, these four models will be referred to as model 1, model 2, model 3 and model 4 in the remainder of this paper.

5.1 Topic interpretation

Before going into the results from the VAR-analysis, recall that the data preparations and model specification spawned six search trends through objectively found keywords, (o*st_1, ost_3, ost_4, ost_21, ost_25, ost_28)*. As mentioned before, these variables were represented by numbers, as they were not created through previously stated product features. To ease further interpretation of results related to these variables, the variables were given meaning through the 10 keywords related to each of them. An overview of the relevant keywords and the resulting topic name can be found in Table 3.

Table 3: 10 Most relevant keywords from the topics used in the final models, found by LDA. The second row indicates a label given to the topic based on the keywords.

ost_1	***ost_3***	***ost_4***	***ost_21***	***ost_25***	***ost_28***
General	*Interior*	*Driving*	*Service*	*Fuel*	*Extra features*
suv	seat	drive	miles	gas	control
vehicle	space	engine	dealer	mileage	safety
ride	trunk	transmission	warranty	city	driver
love	ride	speed	transmission	hybrid	wheel
family	mileage	power	engine	miles	time
luxury	love	fun	months	ride	technology
row	gas	mode	repair	vehicle	vehicle
space	plenty	sport	buy	tank	parking
truck	road	acceleration	service	premium	cruise
cargo	comfort	feel	cost	quiet	tech

The resulting associated topics were named “general”, “interior”, “driving”, “service”, “fuel” and “extra features”. As can be seen from the table, some topics have multiple words hinting towards this common topic. For example, the topic tagged as “service” contains keywords such as “dealer”, “warranty”, “months”, “repair” and “service”. However, for some topics, such as “general”, this was deemed harder. The resulting interpretations were incorporated in the variable names to ease interpretability of the results from the VAR-models. The variables were renamed o*st_general, ost_interior, ost_driving, ost_service, ost_fuel* and o*st_extrafeatures*, respectively. As the review trend spawning from *ost_driving* is also present in the final model, the review trend *ort_4* was renamed to *ort_driving*.

5.2 VAR results interpretation

After giving meaning to the objectively found search trends, all four VAR-models were trained. However, variable coefficients derived directly from a VAR-model do not provide interpretable insights, as all variables are dependent on one another. As a result, impulse response functions were used to gain more insights into the effects of individual predictors on *sales*. The impulse response functions created in the present research described the impact of a change in a single variable on sales one month into the future. Because of the log-log specification of the VAR-model, the coefficients from the impulse response functions represent elasticities at month *t.* For interpretation, one month in the future was chosen, as this showed the direct, short-term effect of a shock in a predictor on sales. Moreover, the main goal of the present research was to assess general predictive performance. As a result, analyzing long-term effects and calculating statistical significance of individual predictors was deemed outside of the scope of this paper. However, both short- and long-term effects will be considered when checking predictive performance in the next paragraph. The sales elasticities of the individual predictors can be found in Table 4.

Table 4: Elasticities of a shock in one of the independent variables on sales for one month in the future.

Variable	Model 1	Model 2	Model 3	Model 4
cpi	0.0080	0.0079	0.0075	0.0077
sst_name	0.0037	-	0.0039	-
sst_fuel	-0.0048	-	-0.0048	-
sst_price	-0.0041	-	-0.0042	-
sst_interior	0.0036	-	0.0031	-
sst_driving	0.0041	-	0.0041	-
ost_general	-	0.0052	-	0.0052
ost_interior	-	-0.0078	-	-0.0078
ost_driving	-	0.0069	-	0.0069
ost_service	-	0.0052	-	0.0052
ost_fuel	-	-0.0034	-	-0.0034
ost_extrafeatures	-	0.0051	-	0.0051
ort_driving	-	-	-	-0.0000
srt_fuel	-	-	-0.0022	-

srt_driving	-	-	0.0058	-

Taking the elasticities into account, it could first be noted that all values are positive except for search and review trends regarding pricing and fuel and the objectively found search trends about interior. For many search and review trends, this positive elasticity was in line with expectation; more searches for a product or its product features may indicate an increase in popularity, leading to more sales. Interestingly enough, the coefficient for *cpi* is also positive in all four models. This may seem counterintuitive, as a higher CPI usually means higher prices, which in turn should lead to lower sales. However, while CPI indicates higher prices, the metric does not contain information about purchasing power of consumers. Therefore, the elasticity may be positive due to variables not included in the model, such as wages. Moreover, the negative elasticities of pricing, fuel and interior could be explained as well. For example, more searches for car pricing may indicate consumers becoming more price sensitive, possibly lowering sales. Furthermore, more searches for fuel consumption may indicate higher prices for oil, also possibly leading to more price sensitivity and lower willingness-to-pay. However, the negative elasticity of *ost_interior* is interesting, as the subjectively found search trend about interior (*sst_interior)* has a positive elasticity in model 1 and 3. This difference may be explained by the keywords related to the objective search trends, as this trend also contains keywords such as “mileage” and “love” which may have skewed the elasticity.

Diving deeper into the elasticities, it is interesting to see that CPI has a relatively large impact on sales across the four models, with an elasticity ranging from 0.0075 to 0.0080. This value suggests that a 1% increase in CPI leads to a 0.008% increase in sales the following month ceteris paribus, according to model 1. To get a better understanding of the magnitude of this effect, the mean unit sales from our data (15,671,950) can be used. Taking this into account, a 1% increase in CPI will lead to an increase in unit sales of 0.008 x 15,671,950 ≈ 125,376 units the following month ceteris paribus, according to model 1. Aside from CPI, the subjective search trends with the largest elasticity are the ones concerning pricing, fuel and driving attributes, with elasticities of -0.0048, -0.0041 and 0.0041 in model 1 respectively. Interpreting the elasticities for search trends, the value for pricing suggests that a 1% increase in online searches for keywords related to car pricing leads to a 0.0041% decrease in sales the following month ceteris paribus, according to model 1. In other words, a 1% increase in online searches for keywords related to car pricing leads to a decrease in unit sales of 0.0048 x 15,671,950 ≈ 75,225 units. Taking the relatively large elasticities of pricing and fuel into account, it was found that this is in line with the earlier interpretation about price sensitivity of consumers. Price can be considered as one of the most important aspects of a buying decision, leading to a high elasticity. Considering the large elasticity of the search trend concerning driving attributes, we theorize that driving attributes may also be relatively important in driving sales, as this can also be seen as a key attribute of an automobile. Interestingly enough, search trends concerning names of cars have a relatively low elasticity in model 1 and 3. At first glance this seems illogical, as general names and terms concerning cars are usually a starting point for Google searches. However, while general names about cars may pose a starting point for Google searches, these searches are usually followed up by more intensive searches for a car’s product features, seeing as purchasing a car is a high-involvement decision. As a result, it may be that the other search trend variables crowd out the effect of the name trend, leading to a low elasticity.

5.3 Predictive performance

After analyzing the details of the four created models more closely, general predictive performance was examined. In order to determine which model was superior in predicting market response, two performance metrics were used, namely the Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The MAE reports the average absolute difference between the predicted and true values. The RMSE does roughly the same, but penalizes big differences between predictions and true values more harshly through its squared-component. A lower RMSE and MAE indicate better predictive performance. As mentioned in the previous paragraph, both short- and long-term predictive performance will be considered. As a result, sales were predicted for five months into the future, after which the RMSE and MAE were calculated by comparing the predictions with the true values. By using this time period in the calculation of RMSE and MAE, the models' ability to predict both short- and long-term effects on sales is reflected in the metrics.

The RMSE and MAE for all models can be found in Table 5. In general, predictive performance was deemed acceptable, with the largest RMSE being 264.97. Seeing as sales ranged from 9,233 to 21,135 and had a standard deviation of 2,468, this error was deemed acceptable by the researcher.

Table 5: General model properties and performance of all four models averaged over 5 months in the future, shown through RMSE and MAE.

	Model 1	**Model 2**	**Model 3**	**Model 4**
Search trends	Subjective	Objective	Subjective	Objective
Review trends	-	-	Subjective	Objective
RMSE	264.97	246.73	71.60	245.81
MAE	191.12	187.40	67.20	186.93

Comparing the different models, it was observed that the model that used price, subjectively found search trends and the subjectively found review trends (model 3) outperformed the remaining three models, with the RMSE being 71.60. This evidence supported the second proposition from the literature review, positing that using review trends as active predictors would improve predictive performance of market response models that use search trends. Moreover, the support of this proposition shows that text from customer reviews has relevant predictive power while utilized as direct predictors. Unfortunately, the model that used both objectively found search and review trends (model 4) had an RMSE of 245.81, which is similar to the RMSE of model 2. Analyzing the elasticities of individual predictors, this makes sense, as the elasticities of model 2 and 4 were relatively similar and the single variable added in model 4 had an elasticity of -0.0000. As a result of this, proposition three from the literature review was not supported by the empirical analysis, indicating that a model that uses both objectively found search and review trends is not the best model for predicting sales. Finally, the model that only used objectively found search trends (model 2) outperformed the model that only used subjectively found search trends (model 1), with the RMSE being 246.73. This supports the first proposition from the literature review, positing that objectively found keywords for search trends may lead to a more accurate model. Moreover, the support of this proposition indicates that text from customer reviews also holds predictive power while used as input for search trends. Summarizing the results concerning predictive performance, it was concluded that the first and second proposition from the literature review were confirmed, while the third proposition could not be confirmed given the empirical results.

6. Conclusions

The aim of this research was to assess the predictive performance of market response models that use both search trends and review data. After creating multiple VAR-models, examining the effects of individual predictors on sales and comparing predictive performance, an adequate answer to the central research question can be given. Circling back to the start of this paper, the central question posed was:

To what extent can utilizing feature mentions of customer reviews enhance the predictive performance of market response models that employ search trends?

After synthesizing the results, it can be concluded that feature mentions of customer reviews can enhance predictive performance of market response models that employ search trends in two ways, namely I) by using the feature mentions as keywords for obtaining search trends in an objective way or II) by reformatting the feature mentions into active predictors usable in a market response model. These two conclusions were drawn as a result of the confirmation of the first and second proposition in the empirical results. If customer reviews are used either I) to extract keywords from for use in search trends or II) to add review trend variables in subjective search trend analysis, one may expect the predictive performance of a market response model to improve. However, the combination of both of these methods did not yield significant improvements to sales predictions.

The conclusions of this research spawn multiple implications for practice. First and foremost, companies that use subjectively found search trends to predict sales may want to adapt their current method to either I) incorporate review data as active predictors more often or II) replace subjectively found search trends with search trends found by keyword extraction. The upside of the first method is that the present research showed a relatively big improvement in predictive accuracy. However, an upside of the second method is that it may be considered a relatively easier or faster way to increase predictive performance, as less data preprocessing is required. A general upside of both methods is that extraction of keywords or creation of review trends can be done more often easily once the required framework has been constructed. Second, to accommodate this shift in market response modeling, companies may want to encourage customers to write reviews more often, as this leads to a bigger sample that can be used in analysis. Moreover, companies may want to consider creating their own review page on their website or make agreements with independent review sites to gain better access to customer reviews. As a result of this, the quality of the available data may increase, possibly leading to more accurate predictions.

While the conclusions and implications of the present research may seem promising, this research has limitations, restraining its generalizability. The first limitation pertains to the compatibility of the different datasets. While CPI, sales and search trends consider all light weight vehicles in the U.S, the review dataset contained reviews concerning only five car brands: Audi, BMW, Mercedes-Benz, Lexus and Infiniti. Because of this smaller sample in the review data, the keywords extracted from the reviews and the created review trends may not be perfectly aligned with the rest of the gathered data, possibly introducing bias in results. This can already be seen in the present paper, as only three review trends significantly granger-caused sales, while 35 review trends were created. The second limitation pertains to the variable operationalization of pricing. In the present paper, CPI for automobiles was used as the operationalization for pricing. While CPI was deemed an acceptable proxy, the positive elasticity of CPI on sales does raise some doubt as to the suitability of this metric, as a negative elasticity was expected. A final limitation of the present research concerns omitted variables. While this research established that search trends, pricing and review trends have an effect on sales, many other variables that affect sales and vary over time have been omitted. For example, the present paper did not include marketing activities, competitor actions or relevant external effects in modeling. As a result, the created models could have contained biased coefficients and possible inaccurate relationships could have been fitted.

Combining the implications and limitations of the present paper, multiple suggestions for future research arise. A first direction could be to explore ways to combine objectively found review trends with objectively found search trends in a better way. Main reasons for this direction lie in that the empirical results from the present paper were against expectations set by prior literature. Specifically, literature suggested that objectively found search trends combined with objectively found review trends should improve performance, while the empirical analysis showed the inverse. An extension of this direction could be to investigate how variable selection for objective search and review trends can be done as best as possible. For example, this research created 30 objective review trends in model 4. However, only one of these trends granger-caused sales significantly. Consequently, the single remaining review trend had an elasticity of approximately 0. Future research may examine better ways to perform variable selection for these trends. A second direction for future research could be to explore the performance of market response models that use search trends, review data and more other predictors that influence sales to see whether the benefit of review and search trend data diminishes. This may be an interesting angle, as other variables may possibly crowd out the effect reviews and search trends pose on sales. As a result, it is interesting to examine the use of search trends and review data in a more holistic model. Fourth, future research could replicate the existing paper with a bigger review dataset, as the review corpus was relatively limited in the present research. Utilizing more reviews may confirm the third proposition posed in the literature review, as it may lead to better keywords for search trends and more valid review trends, hopefully increasing predictive performance. Finally, future research may aim to analyze the magnitude of effects and statistical significance of individual predictors more closely when combining customer reviews and online search trends. While the main focus of the present paper was to examine general predictive performance, comparing effect sizes and statistical significance of predictors between different models may shed more light on the dynamics at play when these two constructs are combined in market response modeling.

In any case, the present research showed that search trends and review data can be combined to enhance predictive performance of market response models. While many companies put significant effort into creating the most accurate model to forecast sales and understand consumer preference as best as possible, this research showed that some of that consumer preference is already revealed through the way someone writes a review and Google's for products. The future is most certainly bright for market response models based on readily available consumer-provided data.

References

Ailawadi, K. L., Lehmann, D. R., & Neslin, S. A. (2001). Market Response to a Major Policy Change in the Marketing Mix: Learning from Procter & Gamble's Value Pricing Strategy. *Journal of Marketing*, *65*(1), 44–61. https://doi.org/10.1509/jmkg.65.1.44.18130

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-Practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the Pricing Power of Product Features by Mining Consumer Reviews. *Management Science*, *57*(8), 1485–1509. https://doi.org/10.1287/mnsc.1110.1370

Bellégo, C., & Pape, L. (2022). Dealing with the Log of Zero in Regression Models. *Social Science Research Network*. https://doi.org/10.2139/ssrn.3444996

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.5555/944919.944937

Boone, T., Ganeshan, R., Hicks, R. A., & Sanders, N. R. (2018). Can Google Trends Improve Your Sales Forecast? *Production and Operations Management*, *27*(10), 1770–1774. https://doi.org/10.1111/poms.12839

Chen, P., Wu, S., & Yoon, J. (2004). The Impact of Online Recommendations and Consumer Feedback on Sales. *International Conference on Information Systems*, 711–724. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1146&context=icis2004

Chevalier, J. A., & Mayzlin, D. (2006). The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research*, *43*(3), 345–354. https://doi.org/10.1509/jmkr.43.3.345

Choi, H., & Varian, H. R. (2012). Predicting the Present with Google Trends. *Economic Record*, *88*, 2–9. https://doi.org/10.1111/j.1475-4932.2012.00809.x

Consumer Price Index for All Urban Consumers: New and Used Motor Vehicles in U.S. City Average . (2023). Federal Reserve Economic Data. https://fred.stlouisfed.org/series/CUSR0000SETA

Danaher, P. J., Bonfrer, A., & Dhar, S. K. (2008). The Effect of Competitive Advertising Interference on Sales for Packaged Goods. *Journal of Marketing Research*, *45*(2), 211–225. https://doi.org/10.1509/jmkr.45.2.211

Dekimpe, M. G., & Hanssens, D. M. (1995). The Persistence of Marketing Effects on Sales. *Marketing Science*, *14*(1), 1–21. https://doi.org/10.1287/mksc.14.1.1

Dellarocas, C., Zhang, X., & Awad, N. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, *21*(4), 23–45. https://doi.org/10.1002/dir.20087

Dinner, I. M., Van Heerde, H. J., & Neslin, S. A. (2013). Driving Online and Offline Sales: The Cross-Channel Effects of Traditional, Online Display, and Paid Search Advertising. *Journal of Marketing Research*, *51*(5), 527–545. https://doi.org/10.1509/jmr.11.0466

Du, R. Y., Hu, Y., & Damangir, S. (2015). Leveraging Trends in Online Searches for Product Features in Market Response Modeling. *Journal of Marketing*, *79*(1), 29–43. https://doi.org/10.1509/jm.12.0459

Du, R. Y., & Kamakura, W. A. (2012). Quantitative Trendspotting. *Journal of Marketing Research*, *49*(4), 514–536. https://doi.org/10.1509/jmr.10.0167Lamberton

Du, R. Y., Netzer, O., Schweidel, D. A., & Mitra, D. (2021). Capturing Marketing Information to Fuel Growth. *Journal of Marketing*, *85*(1), 163–183. https://doi.org/10.1177/0022242920969198

Duan, W., Gu, B., & Whinston, A. B. (2008). Do Online Reviews Matter? - An Empirical Investigation of Panel Data. *Decision Support Systems*, *45*(4), 1007-1016.

Fan, Z., Che, Y., & Chen, Z. (2017). Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis. *Journal of Business Research*, *74*, 90–100. https://doi.org/10.1016/j.jbusres.2017.01.010

Floyd, K., Freling, R., Alhoqail, S. A., Cho, H., & Freling, T. H. (2014). How Online Product Reviews Affect Retail Sales: A Meta-analysis. *Journal of Retailing*, *90*(2), 217–232. https://doi.org/10.1016/j.jretai.2014.04.004

Kumar, V., Choi, J., & Greene, M. (2017). Synergistic effects of social media and traditional marketing on brand sales: capturing the time-varying effects. *Journal of the Academy of Marketing Science*, *45*(2), 268–288. https://doi.org/10.1007/s11747-016-0484-7

Kumar, V., Sunder, S., & Sharma, A. (2015). Leveraging Distribution to Maximize Firm Performance in Emerging Markets. *Journal of Retailing*, *91*(4), 627–643. https://doi.org/10.1016/j.jretai.2014.08.005

Lamberton, C., & Stephen, A. T. (2016). A Thematic Exploration of Digital, Social Media, and Mobile Marketing: Research Evolution from 2000 to 2015 and an Agenda for Future Inquiry. *Journal of Marketing*, *80*(6), 146–172. https://doi.org/10.1509/jm.15.0415

Li, X., Wu, C., & Mai, F. (2019). The effect of online reviews on product sales: A joint sentiment-topic analysis. *Information & Management*, *56*(2), 172–184. https://doi.org/10.1016/j.im.2018.04.007

Light Weight Vehicle Sales: Autos and Light Trucks. (2023). Federal Reserve Economic Data. https://fred.stlouisfed.org/series/ALTSALES

Liu, Y. (2006). Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue. *Journal of Marketing*, *70*(3), 74–89. https://doi.org/10.1509/jmkg.70.3.74

Ltkepohl, H. (2005). New Introduction to Multiple Time Series Analysis. In *Springer eBooks*. https://doi.org/10.1007/978-3-540-27752-1

Neslin, S. A. (1990). A Market Response Model for Coupon Promotions. *Marketing Science*, *9*(2), 125–145. https://doi.org/10.1287/mksc.9.2.125

Reviews of 5 Car Brands. (2021). Kaggle. https://www.kaggle.com/datasets/ashisparida/reviews-of-5-car-brands

Shankar, V., Grewal, D., Sunder, S., Fossen, B. L., Peters, K., & Agarwal, A. (2021). Digital marketing communication in global marketplaces: A review of extant research, future directions, and potential approaches. *International Journal of Research in Marketing*, *39*(2), 541–565. https://doi.org/10.1016/j.ijresmar.2021.09.005

Telser, L. G. (1962). The Demand for Branded Goods as Estimated From Consumer Panel Data. *The Review of Economics and Statistics*, *44*(3), 300. https://doi.org/10.2307/1926401

Wen, L., Liu, C., Song, H., & Liu, H. (2021). Forecasting Tourism Demand with an Improved Mixed Data Sampling Model. *Journal of Travel Research*, *60*(2), 336–353. https://doi.org/10.1177/0047287520906220

West, R. (2021). Best practice in statistics: The use of log transformation. *Annals of Clinical Biochemistry*, *59*(3), 162–165. https://doi.org/10.1177/00045632211050531

Wu, L. & Brynjolfsson, E. (2009). The Future of Prediction: How Google Searches Foreshadow Housing Prices and Quantities. *International Conference on Information Systems*, 89–118. http://digital.mit.edu/research/papers/2012.03_Wu_Brynjolfsson_The%20Future%20of%20Prediction_299.pdf

Appendices

Appendix A: Overview of literature used in the literature review

Topic/theme	Authors & title	Method	Data	Relevant findings
Market Response Models & Trend Searches	Du et al (2015). *Leveraging Trends in Online Searches for Product Features in Market Response Modeling*	Log-log autoregressive model	US Automotive data on advertisement spending and sales, Google Trends-data.	Product feature search trends can be proxies of feature importance. Models with search trends have significantly better performance. (Du et al, 2015)
Market Response Models	Ailawadi et al. (2001) *Market Response to a Major Policy Change in the Marketing Mix: Learning from Procter & Gamble's Value Pricing Strategy*	Log-linear OLS model	Price, promotion, market share, PEN, SOR, USE for P&G and 4 competitors for a multitude of product categories	Own and competitors price and advertising have significant effects on share. Structural and firm-specific effects are also significant. (Ailawadi et al., 2001)
Market Response Models	Neslin (1990) *A Market Response Model for Coupon Promotions*	OLS model	Scanner panel data and coupon distribution data for instant coffee	Model with price, advertising and lagged response for company and competitors. Results indicate significant effects of coupons on market share (Neslin, 1990)
Market Response Models	Danaher et al. (2008) *The Effect of Competitive Advertising Interference on Sales for Packaged Goods*	Log-Log OLS model	Sales and advertising data from a Chicago supermarket	Competitive advertising effects on sales are strong. When competitors advertise in the same week as a company, elasticity decreases (Danaher et al., 2008)
Market Response Models	Telser (1962). *The Demand for Branded Goods as Estimated From Consumer Panel Data*	OLS model	Household panel data from four product categories	Company and competitors prices and lagged market share a relevant predictors of market response (Telser, 1962)
Market Response Models	Kumar et al.(2017) *Synergistic effects of social media and traditional*	Time-varying effect model (TVEM)	Sales, pricing, advertising and seasonality info from US Ice-cream brand	TVEM-approach outperformed benchmarks. Marketing and social media ads are time-

	marketing on brand sales: capturing the time-varying effects			variant (Kumar et al., 2017)
Market Response Models	Dinner et al(2013) *Driving Online and Offline Sales: The Cross-Channel Effects of Traditional, Online Display, and Paid Search Advertising.*	Log-Log Regression	Sales, pricing, advertising and seasonality info from clothing retailer in the US	Display advertising online and Search engine advertising are better than traditional forms of advertising. (Dinner et al., 2013)
Market Response Models	Kumar et al. (2015) *Leveraging Distribution to Maximize Firm Performance in Emerging Markets*	Log-Log Regression	Sales, pricing, advertising, competition, seasonality from multiple markets, such as groceries and cosmetics	Marketing mix and store format is dependent on the product at hand. Price and advertising elasticities may differ between brands (Kumar et al., 2015)
Market Response Models	Dekimpe & Hanssens (1995) *The Persistence of Marketing Effects on Sales.*	ARMA, VAR models	Sales, advertising, seasonality of home-improvement store	Different kinds of advertising have different short and long run impacts. In general, advertising is very relevant in modeling sales. (Dekimpe & Hanssens, 1995)
Trend Searches	Wu & Brynjolfsson (2009) *The Future of Prediction: How Google Searches Foreshadow Housing Prices and Quantities*	AR1-Model	Sales, prices and Google Trends for the housing sector	Search trends relate to future sales and prices in housing. Web search can be utilized to predict future numbers in the economic sector. (Wu & Brynjolfsson, 2009)
Trend Searches	Choi & Varian (2012) *Predicting the Present with Google Trends*	AR-1 model	Automobile sales, unemployment claims, travel destination arrivals	Found significant effects for all datasets, indicating search trends as a valid predictor (Choi & Varian, 2012)
Trend Searches	Du & Kamakura (2012) *Quantitative Trendspotting*	Structural DFA Model	Car sales and Google Trends band/car name searches	Used SDFA to condense many trends into factors. Predictive performance was better than

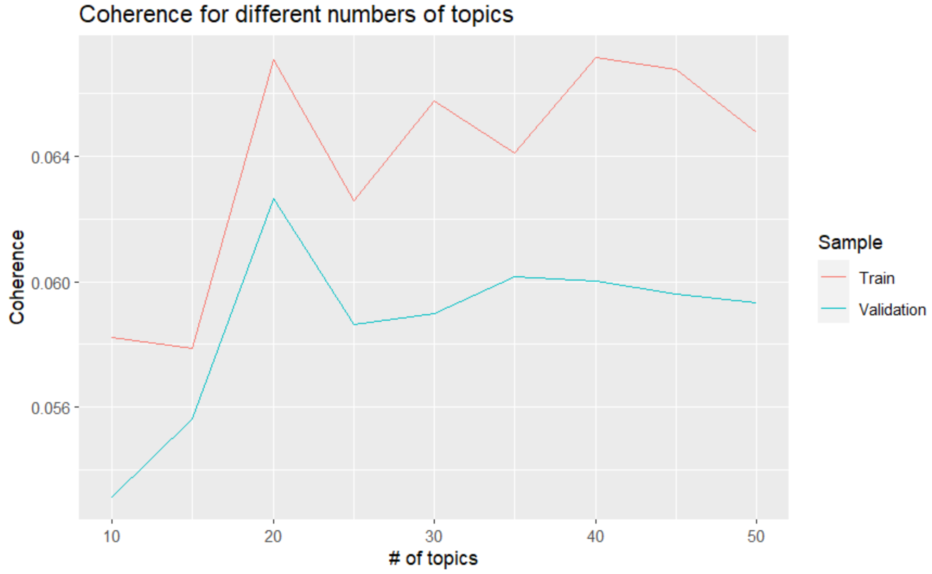
				benchmark without search trends. (Du & Kamakura, 2012)
Trend Searches	Wen et al (2021) *Forecasting tourism demand with an improved mixed data sampling model*	Improved MIDAS model that accounts for moving averages	Tourist arrivals in Hong Kong and Google Trends	Improved MIDAS-SARIMA beats SARIMA and MIDAS benchmarks in performance (Wen et al., 2021)
Trend Searches	Boone et al (2018) *Can Google Trends Improve Your Sales Forecast?*	ARIMA(4,0,0) model	Sales, price of SKU's and Google Trends	Adding search trends to time series models reduces prediction error (Boone et al., 2018)
Reviews	Archak et al (2011) *Deriving the Pricing Power of Product Features by Mining Consumer Reviews*	Log-linear OLS model	Price, sales and reviews from camcorders & digital cameras	Text in reviews has significant predictive power, aside from only using review metrics in modeling. (Archak et al., 2011)
Reviews	Floyd et al (2014) *How Online Product Reviews Affect Retail Sales: A Meta-analysis*	Meta-analysis	Data from all synthesized studies	Meta-analysis on volume and valence of reviews on sales. Results indicate both to be significant, but valence to be most important (Floyd et al., 2014)
Reviews	Duan et al. (2008) *Do online reviews matter? — An empirical investigation of panel data*	3SLS & OLS models	Movie box office revenue, budget, marketing costs & reviews	Significant effect of review volume on movie sales, while valence was not significant. (Duan et al., 2008)
Reviews	Chevalier & Mayzlin (2006) *The Effect of Word of Mouth on Sales: Online Book Reviews*	log-log OLS model	Price, sales ranking & reviews of books	Volume of reviews and average star rating has a significant effect on differences in sales between companies. (Chevalier & Mayzlin, 2006)
Reviews	Liu (2006) *Word of mouth for movies: Its dynamics and impact on box office revenue*	log-log OLS model	Movie budgets, production expenditure & reviews	Volume and valence of reviews have significant effects on box office revenue, with the biggest effect being volume. (Liu, 2006)

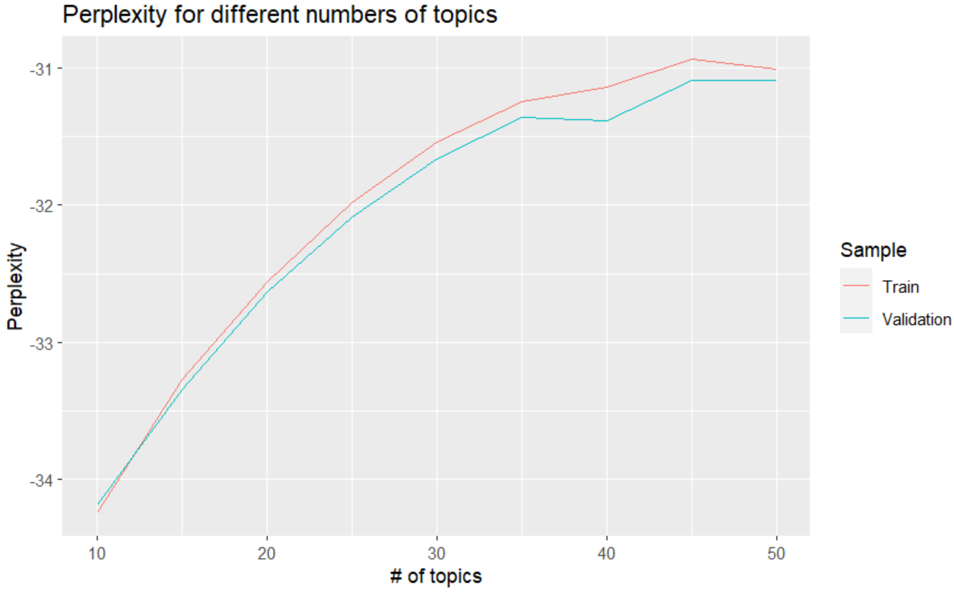
Reviews	Dellarocas et al. (2007) *Exploring the value of online product ratings in revenue forecasting: The case of motion pictures*		Movie budgets, production expenditure & reviews	Expands on the model of Liu. Volume and valence of reviews are significant in predictions. (Dellarocas et al., 2007)
Reviews	Chen et al. (2004) *The impact of online recommendations and consumer feedback on sales*	log-log OLS model	Price, sales & reviews of books	Found a significant effect of review volume on movie sales, while valence was not significant. (Chen et al., 2004)
Reviews	Li et al (2019) *The effect of online reviews on product sales: A joint sentiment-topic analysis*	Joint Sentiment-Topic model & Log-linear OLS model	Price, sales & reviews of tablets	Sentiment information contained in reviews have predictive power in sales modeling. (Li et al., 2019)
Reviews	Fan et al (2017) *Product sales forecasting using online reviews and historical sales data: A method combining the Bass model and sentiment analysis*	Naive Bayes for sentiment extraction, Bass/Norton model for forecasting	Sales and review data from automobiles	Bass/Norton model combined with NB outperforms regular B/N model in terms of accuracy. (Fan et al., 2017)

Appendix B: Subjectively chosen keywords

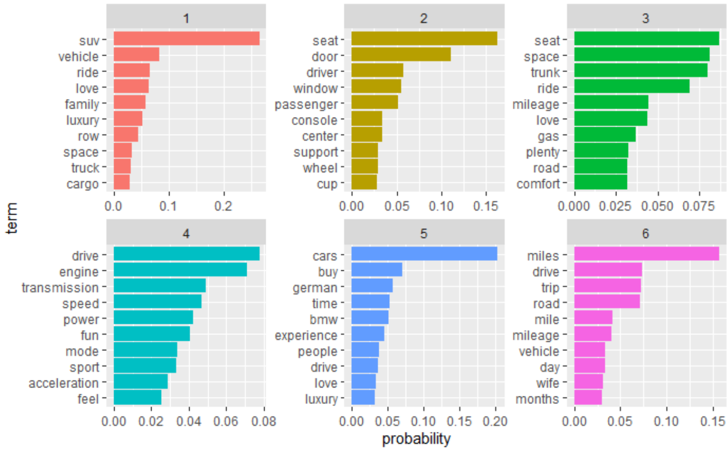
Product feature	**Keywords input**
Product names	car + auto + automobile + motor + vehicle
Price	price + cashback + pricing + cost + costs + value + finance rate + cash back + msrp
Fuel consumption	fuel + mpg + fuel efficiency + city mileage + fuel efficient + mileage + miles per gallon + fuel consumption
Interior	seating + seat + seats + color + dashboard + comfort + material + steering wheel + steer + windows
Driving	acceleration + acc + top speed + miles per hour + mph + handling + gear + gears

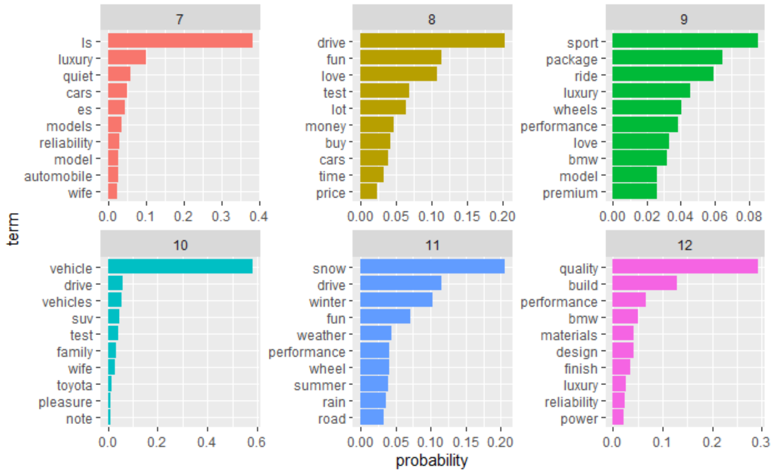
Appendix C: Perplexity and coherence plots for model selection of Latent Dirichlet Allocation.

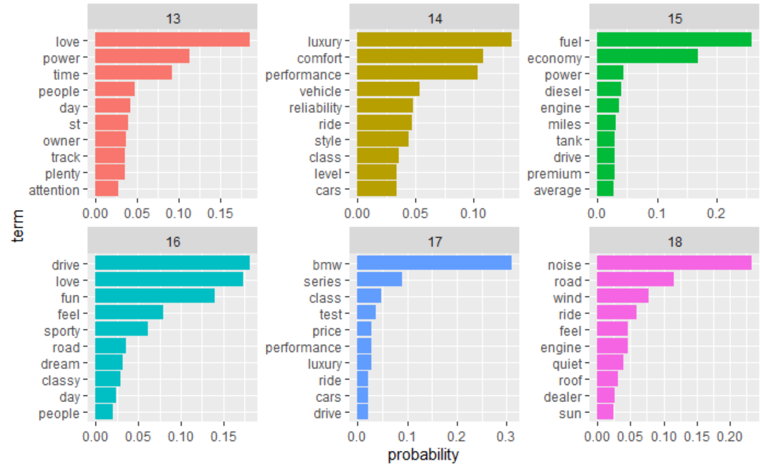


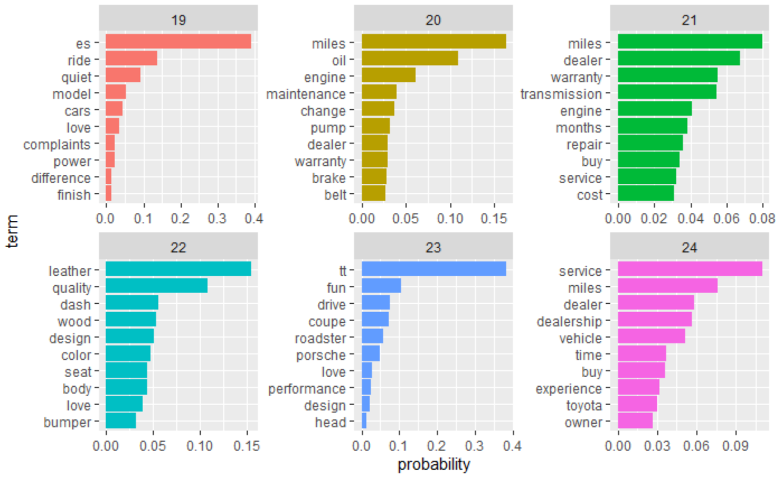


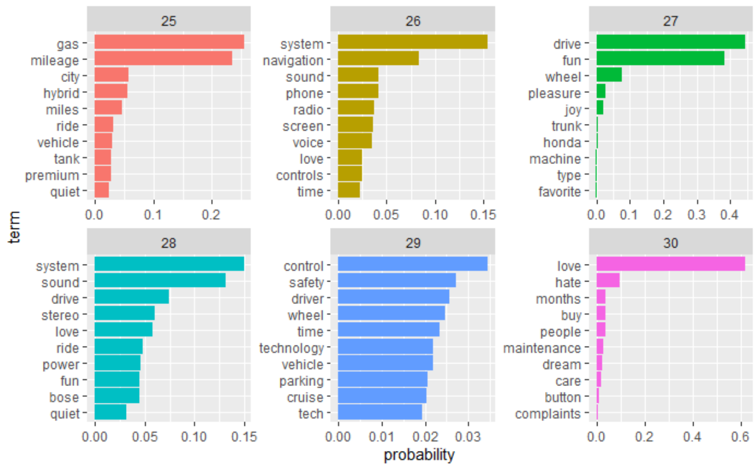
Appendix D: Topics found through LDA and their 10 most related keywords




term
probability
7
ls
luxury
quiet
cars
es
models
reliability
model
automobile
wife
8
drive
fun
love
test
lot
money
buy
cars
time
price
9
sport
package
ride
luxury
wheels
performance
love
bmw
model
premium
10
vehicle
drive
vehicles
suv
test
family
wife
toyota
pleasure
note
11
snow
drive
winter
fun
weather
performance
wheel
summer
rain
road
12
quality
build
performance
bmw
materials
design
finish
luxury
reliability
power


term
probability
13
love
power
time
people
day
st
owner
track
plenty
attention
14
luxury
comfort
performance
vehicle
reliability
ride
style
class
level
cars
15
fuel
economy
power
diesel
engine
miles
tank
drive
premium
average
16
drive
love
fun
feel
sporty
road
dream
classy
day
people
17
bmw
series
class
test
price
performance
luxury
ride
cars
drive
18
noise
road
wind
ride
feel
engine
quiet
roof
dealer
sun





Appendix E: Stationarity checks before and after differencing, performed through Augmented Dickey-Fuller (ADF) Tests

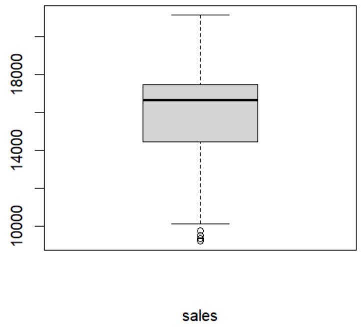
Variable	P-value (before differencing)	P-value (after first-order differencing)
sales	0.811	0.01*
cpi	0.851	0.01*
sst_name	0.427	0.01*
sst_fuel	0.01*	0.01*
sst_price	0.01*	0.01*
sst_interior	0.01*	0.01*
sst_driving	0.01*	0.01*
ost_1	0.544	0.01*
ost_2	0.019	0.01*
ost_3	0.01*	0.01*
ost_4	0.063	0.01*

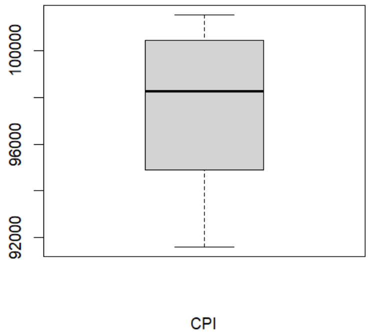
ost_5	0.01*	0.01*
ost_6	0.01*	0.01*
ost_7	0.081	0.01*
ost_8	0.01*	0.01*
ost_9	0.043	0.01*
ost_10	0.331	0.01*
ost_11	0.267	0.01*
ost_12	0.018	0.01*
ost_13	0.01*	0.01*
ost_14	0.062	0.01*
ost_15	0.021	0.01*
ost_16	0.01*	0.01*
ost_17	0.01*	0.01*
ost_18	0.020	0.01*
ost_19	0.037	0.01*
ost_20	0.044	0.01*
ost_21	0.01*	0.01*
ost_22	0.157	0.01*
ost_23	0.028	0.01*
ost_24	0.092	0.01*
ost_25	0.01*	0.01*
ost_26	0.01*	0.01*
ost_27	0.01*	0.01*
ost_28	0.069	0.01*
ost_29	0.319	0.01*
ost_30	0.01*	0.01*
srt_name	0.012	0.01*
srt_fuel	0.023	0.01*
srt_price	0.044	0.01*
srt_interior	0.047	0.01*

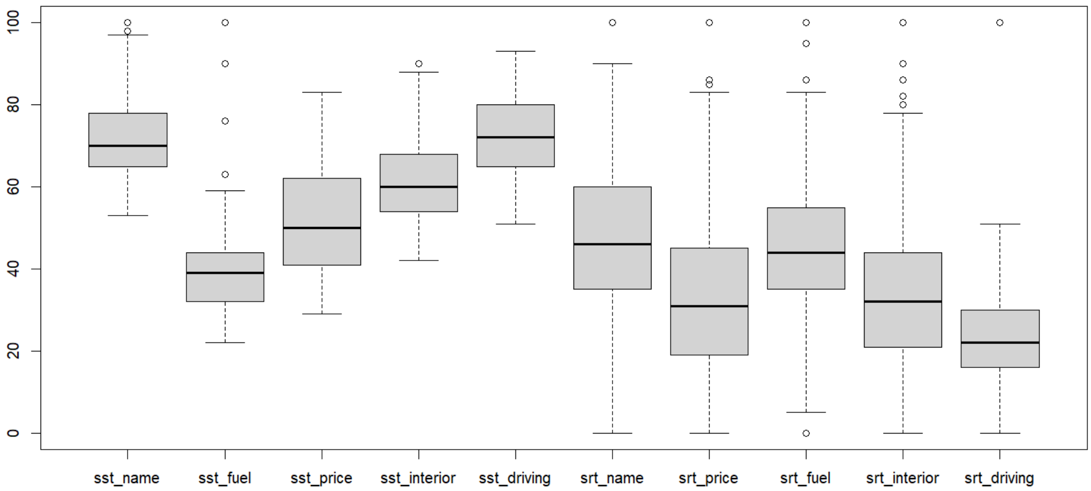
srt_driving	0.01*	0.01*
ort_1	0.137	0.01*
ort_2	0.01*	0.01*
ort_3	0.091	0.01*
ort_4	0.013	0.01*
ort_5	0.227	0.01*
ort_6	0.023	0.01*
ort_7	0.01*	0.01*
ort_8	0.021	0.01*
ort_9	0.053	0.01*
ort_10	0.019	0.01*
ort_11	0.01*	0.01*
ort_12	0.182	0.01*
ort_13	0.423	0.01*
ort_14	0.071	0.01*
ort_15	0.216	0.01*
ort_16	0.025	0.01*
ort_17	0.025	0.01*
ort_18	0.021	0.01*
ort_19	0.075	0.01*
ort_20	0.583	0.01*
ort_21	0.359	0.01*
ort_22	0.043	0.01*
ort_23	0.046	0.01*
ort_24	0.064	0.01*
ort_25	0.036	0.01*
ort_26	0.145	0.01*
ort_27	0.01*	0.01*
ort_28	0.014	0.01*
ort_29	0.020	0.01*

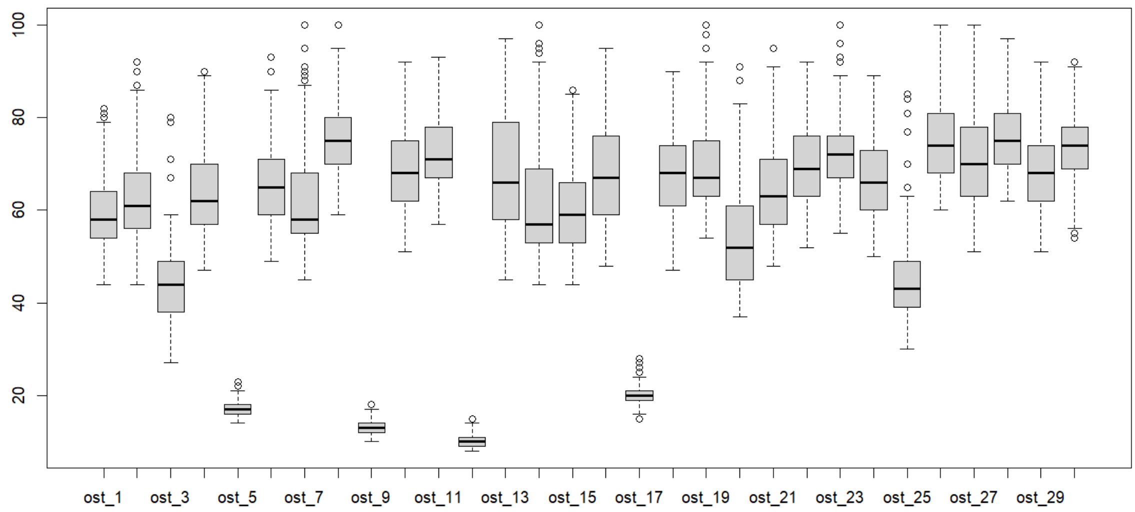
ort_30	0.069	0.01*

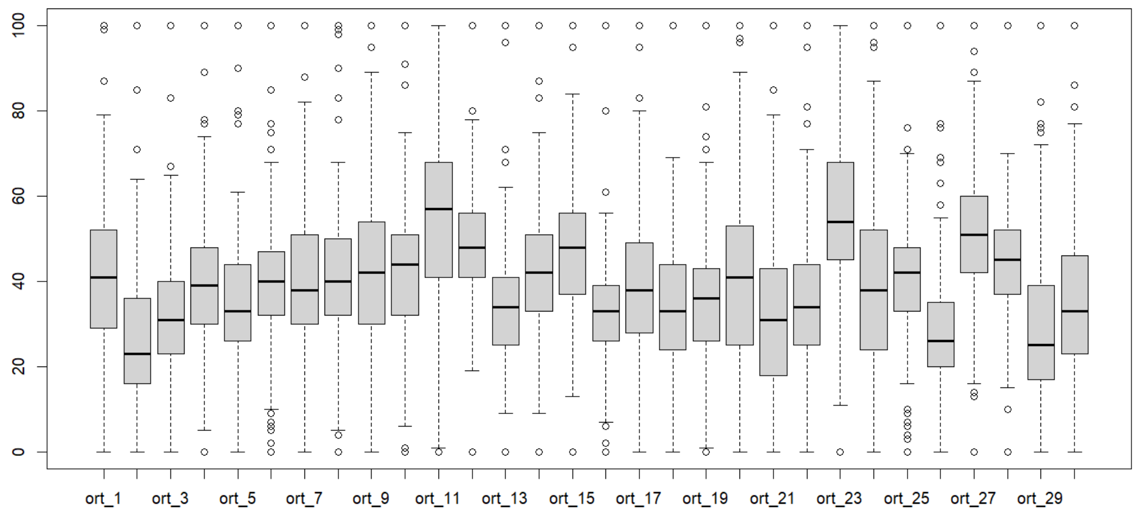
Appendix F: Boxplots of the distribution of all variables, used to check for outliers











Appendix G: Granger-causality tests on all independent variables on *sales*.

Variable	P-value (model 1)	P-value (model 2)	P-value (model 3)	P-value (model 4)
cpi	<u>0.033*</u>	<u>0.032*</u>	<u>0.034*</u>	<u>0.031*</u>
sst_name	0.874	-	0.881	-
sst_fuel	0.250	-	0.238	-
sst_price	0.346	-	0.392	-
sst_interior	0.126	-	0.152	-
sst_driving	0.343	-	0.387	-
ost_1	-	<u>0.037*</u>	-	<u>0.037*</u>

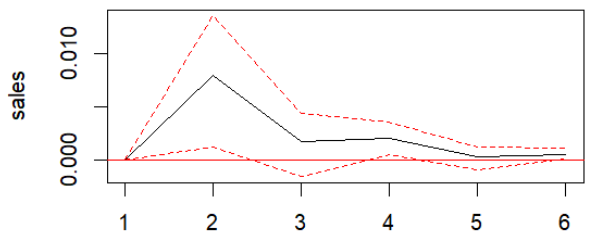
ost_2	-	0.131	-	0.131
ost_3	-	<u>0.095*</u>	-	<u>0.095*</u>
ost_4	-	<u>0.026*</u>	-	<u>0.026*</u>
ost_5	-	0.615	-	0.615
ost_6	-	0.938	-	0.938
ost_7	-	0.798	-	0.798
ost_8	-	0.458	-	0.458
ost_9	-	0.673	-	0.673
ost_10	-	0.881	-	0.881
ost_11	-	0.141	-	0.141
ost_12	-	0.786	-	0.786
ost_13	-	0.146	-	0.146
ost_14	-	0.925	-	0.925
ost_15	-	0.272	-	0.272
ost_16	-	0.193	-	0.193
ost_17	-	0.783	-	0.783
ost_18	-	0.497	-	0.497
ost_19	-	0.856	-	0.856
ost_20	-	0.184	-	0.184
ost_21	-	<u>0.083*</u>	-	<u>0.083*</u>
ost_22	-	0.170	-	0.170
ost_23	-	0.336	-	0.336
ost_24	-	0.254	-	0.254
ost_25	-	<u>0.048*</u>	-	<u>0.048*</u>
ost_26	-	0.186	-	0.186
ost_27	-	0.651	-	0.651
ost_28	-	<u>0.035*</u>	-	<u>0.035*</u>
ost_29	-	0.226	-	0.226
ost_30	-	0.665	-	0.665
srt_name	-	-	0.850	-

srt_fuel	-	-	0.058*	-
srt_price	-	-	0.606	-
srt_interior	-	-	0.512	-
srt_driving	-	-	0.048*	-
ort_1	-	-	-	0.958
ort_2	-	-	-	0.466
ort_3	-	-	-	0.430
ort_4	-	-	-	0.099*
ort_5	-	-	-	0.473
ort_6	-	-	-	0.382
ort_7	-	-	-	0.268
ort_8	-	-	-	0.987
ort_9	-	-	-	0.927
ort_10	-	-	-	0.756
ort_11	-	-	-	0.415
ort_12	-	-	-	0.466
ort_13	-	-	-	0.618
ort_14	-	-	-	0.994
ort_15	-	-	-	0.487
ort_16	-	-	-	0.481
ort_17	-	-	-	0.776
ort_18	-	-	-	0.160
ort_19	-	-	-	0.484
ort_20	-	-	-	0.455
ort_21	-	-	-	0.346
ort_22	-	-	-	0.758
ort_23	-	-	-	0.557
ort_24	-	-	-	0.387
ort_25	-	-	-	0.999
ort_26	-	-	-	0.896

ort_27	-	-	-	0.756
ort_28	-	-	-	0.515
ort_29	-	-	-	0.399
ort_30	-	-	-	0.794

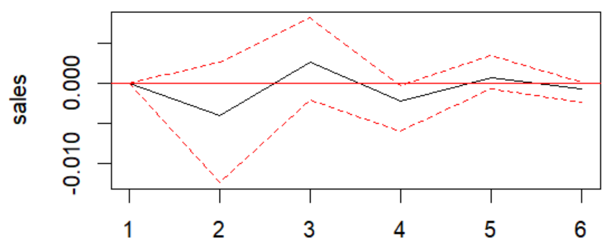
Appendix H: Plots of Impulse Response Functions of *cpi* and *sst_price* on *sales*, extracted from Model 1.

Orthogonal Impulse Response from cpi



95 % Bootstrap CI, 100 runs

Orthogonal Impulse Response from st_price



95 % Bootstrap CI, 100 runs

Appendix I: AIC, HQ, BIC and FPE as criteria for lag order selection for all four models

Model 1

# of lags	**1**	**2**	**3**	**4**
AIC	-45.63	-45.76	-45.79	-45.69

HQ	-45.21	-44.97	-44.64	-44.16
BIC	-44.59	-43.81	-42.94	-41.93
FPE	1.53e-20	1.35e-20	1.31e-20	1.47e-20

Model 2

# of lags	**1**	**2**	**3**	**4**
AIC	-53.19	-53.38	-53.52	-53.41
HQ	-52.65	-52.36	-52.01	-51.42
BIC	-51.85	-50.85	-49.81	-48.50
FPE	7.95e-24	6.59e-24	5.79e-24	6.62e-24

Model 3

# of lags	**1**	**2**	**3**	**4**
AIC	-46.00	-46.07	-46.02	-46.07
HQ	-45.32	-44.78	-44.13	-43.57
BIC	-44.33	-42.90	-41.36	-39.90
FPE	1.06e-20	9.95e-21	1.06e-20	1.05e-20

Model 4

# of lags	**1**	**2**	**3**	**4**
AIC	-53.67	-53.89	-54.10	-54.02
HQ	-53.00	-52.60	-52.20	-51.51
BIC	-52.00	-50.71	-49.42	-51.5
FPE	4.90e-24	3.97e-24	3.27e-24	3.69e-24