BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE AND ENGINEERING

DEPARTMENT OF STATISTICS AND MATHEMATICS

BY

ETHEL A MAHORI

REGISTRATION NUMBER B192346B

A DISSERTATION SUBMITTED TO BINDURA UNIVERSITY IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE BACHELOR OF SCIENCE HONOURS DEGREE IN STATISTICS AND FINANCIAL MATHEMATICS.

Topic: The impact of product bundling on store sales. A case study of Simbisa

SUPERVISOR: MR KANJODO

CORE SUPERVISOR: MR B KUSOTERA

JUNE 2023

DECLARATION

I declare that this work is my own, which was done by me without copying or extracting from previous sources without proper acknowledgment of sources.

Ethel A Mahori ………………………….. …12/06/23

B192346B signature date

Certified by

Mr Kanjodo …………………………… …12/06/23

Supervisors' signature date

Mr Kusotera ……………………………… …12/06/23

Core supervisors' signature date

Mr Magodora Magodora …03/10/23

Chairperson signature date

DEDICATION

I dedicate this project to my family.

ACKNOWLEDGEMENTS

I am grateful for all the support that I have received during the work of this research project, without the support I would never have been able to finish this book. Firstly, I would like to thank my supervisor Mr Kanjodo, this research would not have been possible without his knowledge, advice and detailed comments on earlier drafts of this book. He has created for me the best environment to develop a research on my own and getting experience with statistics. Mr B Kusotera, has contributed greatly to this research by generating many of the ideas that are central to this thesis. Moreover, I would like to thank all my lecturers for their immersed contribution in making this programme successful. I am grateful to Innscor Fast Foods Restaurant managers and senior statisticians for providing data used in this research. My colleagues and friends have been a great support, I am grateful for their help and friendship. My family also supported me and I am grateful for all their support and I want them to know that they have helped me more than they believe. Above all, we give glory to God for allowing me to complete the programme at Bindura University of Science Education.

ABSTRACT

Bundling is a marketing strategy that involves an offer of vacation packages, selling at least two separate products at one single price, Martins et al., (2021). Bundling, a practice of marketing two or more products or services as a specially priced package Guo et al., (2021). Businesses engage bundling strategies in order for them to sell many products at lower costs, to contract consumer surplus, and to create value for consumers. VAR and VECM models were implemented in the analysis of data using Eviews software. The research used monthly time series data from December 2010 to Dec 2021. The main objectives of the study were to determine the long run and short run impact of bundled sales, non- bundled sales and price on sales. The findings of the study revealed that overall store sales are highly and positively influenced by bundled products sales, and this is in agreement with Dedernger and Kumar (2013), where he stated that bundling is an effective tool used to increase sales and profit. Results have also revealed that bundled sales and non-bundled sales have a positive impact on sales whereas price has a negative impact on total sales in the long run. The study recommends that Chicken Inn shops consider offering promotions such as bundling to boost sales revenue. They should also be cautious in setting prices to avoid negatively impacting sales revenue and experiment with different bundling strategies to find the most effective approach. The study further recommends that Chicken Inn shops should continually monitor and adjust their strategies to remain competitive in the industry. Future research could examine the impact of other variables such as costs and inventory levels on sales revenue and profitability, explore the effectiveness of different bundling strategies in different contexts, and examine the impact of bundling on consumer behavior and decision-making.

ACRONYMS

OSR- Quick Service Restaurant

CBD- Central Business District

GAAP- Generally Accepted Accounting Principles

AR- Autoregressive

ADF- Augmented Dickey Fuller

DF- Dickey Fuller

SC- Schwarz Criterion

LR- Likelihood Ration

FPE- Final Prediction Error

AIC- Akaike Information Criterion

HQ- Hartmann Quinn

VAR- Vector Autoregressive

VECM- Vector Error Correction Model

LM – Langrage Multiplier

VIF- Variance Inflation Factor

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

1.0 Introduction

The fast-food sector is getting more and more competitive, so companies are looking for fresh ways to persuade people to buy their products and boost sales. Because of the recent health food craze, it's easy to think that fast food consumption is declining, but most people aren't aware that it's actually increasing. This thesis aims to investigate the impact of product bundling on store sales at Chicken Inn. This chapter centers on the background, problem statement, significance, limitations, and shortcomings of the study. Additionally, it provides definitions for terms used in the research and this chapter's conclusion.

1.1 Background of study

The fast-food industry has experienced significant growth in recent years due to the liberalization of trade in the sector during the early 1990s. This policy change allowed for increased competition and the emergence of new players, including black entrepreneurs who were able to establish their own fast food outlets (Dzama, 2013). In this highly competitive industry, the ability to create value is essential for success. Innscor Africa Limited is the leading player in the Zimbabwean fast food market, with a diverse range of strategic business units that generate an average turnover of $814 million and an average operating profit before tax of $18 million. The company is a conglomerate with operations primarily focused on the food sector, both in Zimbabwe and in several other African countries. It holds strong positions in manufacturing, distribution, and retail sectors, and has a portfolio of well-known household brands, as well as a strong capacity to generate cash.

Innscor Africa Limited's subsidiary, Simbisa Brands Limited (Simbisa), is a publicly traded company that owns, operates, and franchises a variety of Quick Service Restaurant (QSR) brands. Innscor Fast Foods is a part of the Fast Moving Consumer Goods industry, offering food that is freshly prepared and customized to meet the specific needs of customers who prefer quick and convenient service.

Product bundling is often implemented with the aim of increasing productivity, but it can have negative impacts on profits while increasing expenses. Bundling tends to encourage customers to spend more on the day of the promotion, but then less on subsequent days, resulting in reduced sales revenue and profits on a weekly, monthly, and annual basis. Additionally, promotion costs associated with product bundling can lead stores to buy more stock, which can ultimately result in waste.Despite these potential drawbacks, Simbisa's introduction of product bundling resulted in a 3% increase in revenues to $79.1 million, with significant contributions coming from operations in Zimbabwe, which provided $48.9 million (62%) of the total revenue. This figure was 1% lower before bundling. However, cash flows from operations decreased by 9% due to an increase in receivables and stock levels (tellimer.com, 2018).

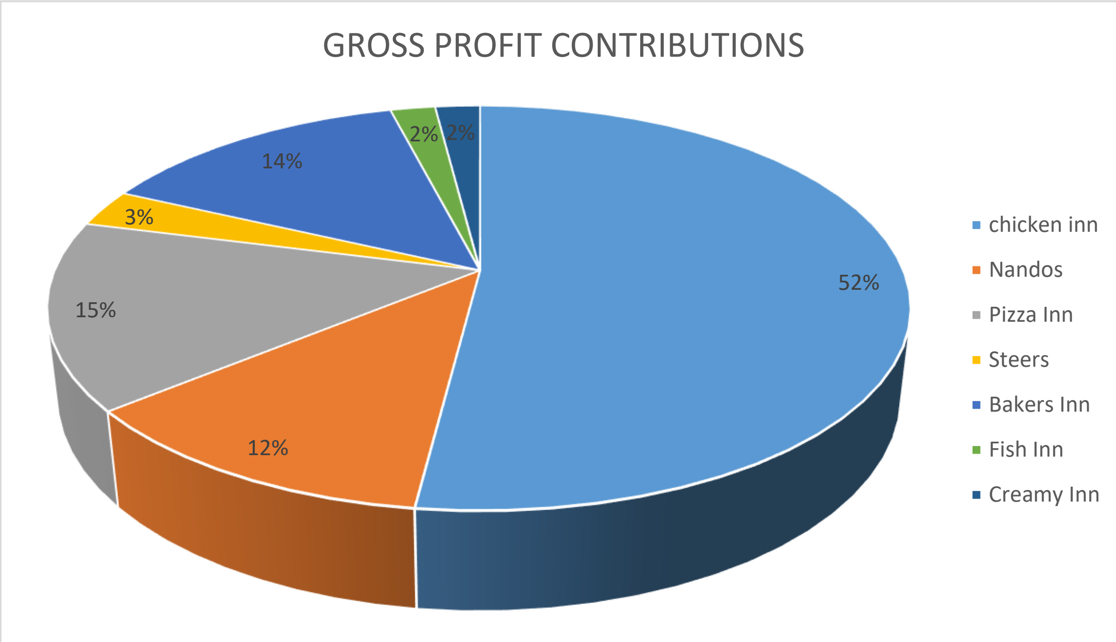
According to a study by Laube (2013), product bundling has been found to have a positive impact on sales revenue and profitability in the fast-food industry. Another study by Feng, Li and Zhang, (2019) found that bundling can help to increase customer loyalty and repeat purchases, which can further boost sales revenue in the long run.

However, not all studies have found that bundling is an effective strategy in the fast-food industry. For example, a study by Babutsidze and Cowan (2009) found that bundling can have a negative impact on sales revenue if it leads to cannibalization of non-bundled products.

Given the mixed findings on the effectiveness of bundling in the fast-food industry, it is important to further investigate the impact of bundling on store sales. This study aims to contribute to the existing literature by examining the impact of product bundling on sales revenue and profitability in Chicken Inn shops, a fast-food chain in Harare CBD. By analyzing the sales data of Chicken Inn shops, this study seeks to provide insights into the effectiveness of bundling as a marketing strategy in the fast-food industry.

In summary, while previous studies have provided some insights into the impact of bundling on sales revenue and profitability in the fast-food industry, further research is needed to examine the effectiveness of bundling in different contexts. This study aims to contribute to the literature by examining the impact of bundling on store sales in Chicken Inn shops.

Figure 1: Gross Profit Contributions



In general, Chicken Inn, one of Innscor's seven brands, generates close to 50% of the fast-food industry's total profit. The remaining percentage of the profits is made up by the other brands. The brand that produces the least gross profit for Simbisa Limited is Creamy Inn and Fishy Inn.

1.2 Problem Statement

Simbisa Brands Chicken Inn uses product bundling as a sales tactic with the goal of convincing customers to buy goods they might not otherwise buy if the goods were made available separately in order to boost sales and gross profit. The management is unsure as to whether or not this kind of technique is having the desired effect of increasing monthly sales. Due to this unawareness, it is necessary to assess whether or not concentrating on product bundling is important.

1.3 Objectives

The objectives of the research are:

1. To determine the short run impact of price, bundled sales and non-bundled sales on sales.

2. To determine the long run impact of price, bundled sales and non-bundled sales on sales

1.4 Research Questions

The methodology is determined by the questions of the research, and guide all stages of inquiry, analysis, and reporting. To know more of a situation that needs to be solved or addressed, it begins with a research question. I have found the following to be the research questions of the project.

1. Is there any short run relationship between price, bundled sales and non bundled sales with sales?
1. Is there any long run relationship between price, bundled sales and non bundled sales with sales?

1.5 Significance of Research

1.5.1To the researcher

The research enables the student to gain experience on how a research is done, including the process involved. The project also enhances the student's research intellectual abilities and getting insight into factors influencing bundled products in Zimbabwe.

1.5.2To the university

The research adds to the university's literature and becomes a source of secondary data to other researchers and students.

1.5.3To Simbisa Brands

The objective of this thesis is to check the efficacy of product bundling approach in boosting sales of Chicken inn shops in Harare. Having the academic interests in applying statistical models for this research, it is hoped that this research shall help Chicken inn management to make product bundling more profitable. The research is to help operation managers in decision making, control and planning bundled products sales. The emphasis is on measuring the effects of bundling products on sales. The results may then be used in future research to evaluate whether the current bundling techniques are profitable or not.

1.6 Assumptions

Assumptions in a research are those things that are somewhat out of our control, but if they don't appear the study would lose its relevance. Leedy and Ormrod (2010) suggest that assumptions are essential to defining a research problem, to the extent that the problem would not even exist without them. For this study we are assuming that:

1. Every customer has access to bundled products.

2. Data collected from the company is accurate.

3. The tools being used for data collection are valid and reliable.

1.7 Limitations

These are factors influencing the research which the researcher can't control. Also defined as short coming conditions that the researcher can't control and they place restrictions on methodology and decisions. The study's limitations are as follows.

1. The models and method presented seek to find the effect of bundling on sales rather than pursue how to maximize sales.

2. The models are one-period models, considering only the current sales.

1.8 Delimitations

These are choices and boundaries that have been put for the study being conducted. They are as follows.

(a) Sample data is to be used representing the brand.

(b) The sample size can be large to increase our appreciation of products.

1.9 Definition of terms

Sales- related activities or the volume of goods or services sold in a specific time frame.

Customer- Someone who makes a purchase from a store or company.

Combo meal- A meal that combines multiple food items (a bundle).

Lag- predetermined period of time

Null Hypothesis- a hypothesis that posits that there is no noteworthy distinction between particular and any observed difference are due to sampling or experimental error.

Alternative Hypothesis-hypothesis used in hypothesis testing that contradicts the null hypothesis.

Correlation matrix - each cell of a correlation matrix, which displays the correlation coefficient between variables and displays the relationship between two variables.

Multicollinearity - According to Zuur (2009), it is a condition of high inter-association among explanatory variables, arises when there is an exact (or nearly exact) linear relationship between two or more explanatory variables.

Vector Autoregression model - A multivariate time series version of the univariate autoregression model, which treats all variables as response variables (Stock & Watson, 2001).

1.10 Chapter Summary

The research to model the impact of product bundling on store profits at Chicken inn particularly Harare. The objectives of the study were discussed as they are to pave a way in achieving the main goal of this study. The study indicates being of great significance from the way it was discussed. This chapter ended by stating limitations and delimitations of the research as well as definitions of terms

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

According to McCombes (2022), this section offers an overview of current knowledge, aids in the recognition of pertinent thesis, techniques, and gaps in the analysis that is already done, identifies those specialists who are researching the subject, and highlights the value of further research. The method for simulating how bundling affects retail sales is provided in this section. This chapter reviews and evaluates the researches available on the impact of product bundling on store sales. It also evaluates theories that are said to be behind sales volumes. Depth knowledge and understanding of the topic is obtained under this chapter

2.1 Product Bundling

The practice of bundling involves companies combining several of their products or services into one package, frequently at a lower price than they would charge customers to buy each item separately (Liberto, 2021). According to Liberto (2021), businesses sell the bundle for less than they would charge for the individual items under a bundle pricing scheme. By stimulating demand with discounts, businesses may be able to increase their sales volume by being able to sell products or services that they otherwise might not have been able to. Since getting less money for something means making less money off of it, this approach may eventually help make up for losses in per-item profit margins.

Some researchers claim that, in the highly competitive business environment of today, combining products or services for targeted sales is a strategy used by marketing decision-makers Tunali et al., (2021). Targeted sales can take many different forms, such as increasing the likelihood that a customer will buy, promoting certain products among a certain consumer segment, or enhancing customer experience.

Product bundling, according to Neil Kokemuller (2015), is a marketing strategy that entails combining several products or components into a single bundled solution. As businesses work to reduce acquisition costs, this tactic has grown in popularity at the start of the twenty-first century. Bundling has advantages for businesses and their clients when it is successfully implemented.

According to Knutsson (2011), the term "bundling" can be used in many different ways and for a vast range of offerings. There are also multiple ways to define it. It has been referred to as a simple method of combining new product offerings to complement the product line by Gerdeman, (2013) and is a marketing strategy that boosts sales by bundling separate items into a packed, typically cheaper bundle Banciu and Odegaard (2014).

According to Meissner et al. (2013), product bundling is the method of combining services and products to establish a new price point. The question of whether the buyer or the seller actually benefits from the bundling itself remains, despite the fact that this technique is widespread across many industries. Since the component products are already on the market, Derdenger and Kumar (2013) claimed that the process of combining two or more products into a single bundle is possibly the most pliable aspect of product strategy. Bundling, according to Binesh et al. (2021), is a way to alter how customers view the worth of the product. Bundling, according to Vithala et al. (2018), is the act of promoting two or more products simultaneously.

2.2 Types of product bundling

Potgieter and Howell (2021) claim that a variety of bundling strategies have emerged, each with a unique strategic justification and impact on the distribution and quantum of dynamic and fixed welfare. Pure bundling, also known as tying, occurs when products are only available as a bundle. Examples include a variety of articles and publications of advertisements in newspaper editions (or circulated online), TV programming that is bundled into a channel and unlimited broadband plans (effectively, access and usage costs are combined for a single price). When customers can purchase a product independently or as part of a bundle, whether it is sold by a single company or is put together from the component offerings of several rival companies, they are said to be engaged in mixed bundling. A prime example is the triple- and quadruple-play telecommunications packages that include television, Internet, and telephone (both fixed and mobile) services. Customers can select either the bundle or any individual part of the package.

According to Banciu and Odegaard (2014), pure bundling is defined as the ability to only purchase a bundle and not individual products, and mixed bundling is defined as the ability to purchase a bundle along with individual components of that bundle.

Pure bundling, according to Eghbali-Zarch et al., (2019), is the sale of bundles containing various products. This helps businesses to cut costs and substantially boost sales revenue and sales. Customers frequently prefer to purchase multiple items at once for a price that is typically less expensive than buying the components separately Chew et al., (2015). SAP (systems, applications, products) is an illustration of pure bundling by an incorporated company. Other instances include Microsoft selling software suites that contain a number of its programs Chakravarty et al., (2013) or app stores that let developers bundle apps to increase sales and foster customer loyalty Wan et al., (2017). Set menus offered at McDonald's include a variety of burgers, and serving of french fries, and soft drinks.

From the standpoint of the consumer, bundling, according to Krystallis et al., (2011), entails combining various food items with service components to provide meal options for various consumer groups and situations. According to Prasad et al., (2010), if the bundled products are sufficiently asymmetric in terms of product costs and network effects, mixed bundling is likely to be more profitable because the items are similar, pure bundling can be profitable.

According to Derdenger and Kumar (2013), there is proof that pure bundling prevails when incremental costs are low in propotion to consumer valuations, while mixed bundling is considered to be the most advantageous scenario when marginal costs are higher.

2.3 Techniques of Product Bundling

According to Journal of Public Policy and Marketing Vol. 29 (2), the fast-food industry currently employs four techniques, including:

(a) Bundles that are offered for sale at a reduced price compared to the totals of their parts' individual prices. Customers may begin to perceive bundles as price promotions as a result of this.

(b) Highlighting the bundle in the consumer-provided information. It is common knowledge that selected goods frequently increase sales despite remaining the same in price.

(c) Companies that feature the bundle in a variety of information formats, changing the level in which they highlight the product bundle and its price. An example is, in certain scenarios, businesses will position the bundle price and detail close to the price of the main item within the bundle, whereas in other instances, the information in this bundle will be placed in a different area from details about the product item. In keeping with Sharpe and Staelin (2010) study on the effects of structure of complementarity, which looked at how various menu designs can influence customer buying behavior.

(d) Combining a number of related items. This enables businesses to make it simpler for customers to locate, think about, and sequence the group of products rather than to find, choose, and order every product separately.

2.4 Objectives of Product Bundling

The primary goal, according to some researchers, is to entice or persuade customers to purchase goods they otherwise wouldn't have if they were sold separately Arora, (2011). All of this is done to boost sales and increase profit. According to the article "Overview of Consumer Trends in the Food Industry" (Bakos and Brynjolfsson, 2000), product bundling reduces costs by reducing the costs associated with distribution and proceedings.

These strategies, according to Cummins and Mullin, (2010), include growing volume, encouraging trial, encouraging repeat purchases, encouraging loyalty, expanding usage, creating interest, creating awareness, deflecting attention away from price and undermining price discrimination, providing intermediary support, favoring particular users, and reviving brand perception on service failure. By incentivizing customers to buy multiple products at once, product bundling lowers the cost of search and makes customer acquisition easier Eisenmann et al., (2011).

According to Sharpe and Staelin (2010), bundling is a useful strategy for increasing sales and profits while lowering carrying, production, and shipping expenses. The second goal is to differentiate prices so that sellers can divide the market based on consumer reservation prices, or the amounts that customers are prepared to pay Prasad et al., (2014), as well as to outperform rivals in the market Datta, (2010). Rafiei et al. (2013) conducted research on bundling's other goal and found that it increased customer willingness rather than attitude.

2.5 Theoretical Literature

Chen and Xie (2008) proposed a conceptual framework that explains how product bundling can affect store sales. The framework suggests that bundling can increase store traffic, lead to higher purchase incidence, and increase the average transaction value. The authors argue that these effects can be amplified when the bundled products are complementary or when the bundle price is perceived as a good value.

Villas-Boas and Winer (1999) developed a model that explains how product bundling can affect price competition and store sales. The model suggests that bundling can increase store traffic, reduce price competition, and increase store profits. The authors argue that these effects can be stronger when the bundled products are complementary and when the bundle price is lower than the sum of the individual prices.

Hui and Bradlow (2008) developed a model that explains how product bundling can affect store sales and customer preferences. The model suggests that bundling can increase store sales by increasing the variety of products offered and by enhancing customer preferences for the bundled products. The authors argue that these effects can be stronger when the bundled products are complementary and when the bundle price is perceived as a good value.

(Huang, 2007), the impact of product bundling on store sales is the price-quality inference theory. This theory suggests that consumers use the price of a bundle to infer the quality of the bundled products. If the bundle price is perceived as a good value, consumers may infer that the bundled products are of high quality, and this could lead to increased demand and sales for the store.

(Broniarczyk and Alba, 1994) the impact of product bundling on store sales is the complementarity theory. This theory suggests that product bundling can be more effective when the bundled products are complementary, meaning that they are used together or have a synergistic effect. When bundled products are complementary, consumers may be more willing to purchase the bundle, and this could lead to increased sales for the store.

(Kahn and Lehmann, 1991) suggests that the impact of product bundling on store sales is the variety-seeking theory. This theory suggests that product bundling can increase store sales by offering consumers a greater variety of products. When consumers are presented with a bundle that includes multiple products, they may be more likely to make a purchase, even if they only wanted one of the products. This could lead to increased sales for the store.

Overall, theoretical literature suggests that product bundling can have a positive impact on store sales by increasing demand, enhancing the perceived value of the bundled products, and offering consumers a greater variety of products. However, the effectiveness of product bundling may depend on various factors, including the complementary nature of the bundled products, the perceived value of the bundle, and the preferences of the target market.

2.6 Empirical Literature

Derdenger and Kumar (2013) used data from a hand held video game market made up of hardware consoles and software games to investigate the effects of firm bundling decisions on consumer choices and market outcomes in a setting with complementary goods. They created a dynamic model based on the preferences of individual consumers, in which consumers would gradually select consoles, bundles, and video games. They proposed a positive correlation between consumer valuation for component products and consumer population. The fact that they discovered a fresh stimulator of bundling effectiveness as a result of dynamic consumer segmentation is of utmost importance. Bundles entice some customers to make more purchases while luring others into the market who might not have otherwise done so. They found that the dynamic consumer segmentation theory makes bundling even more effective when there is a positive correlation between consumer valuations of the two products because it has the ability to convince customers with low valuations of both products to temporarily substitute purchase in order to enter the market sooner. Additionally, given that consumer valuations of component products are positively correlated, the conventional homogenization mechanism is probably only useful to a limited extent. The dynamic consumer segmentation mechanism that underlies the effectiveness of bundling is likely to be significant, particularly in markets where consumers place different values on different products and where there are significant intertemporal trade-offs, such as with durable goods or technology products.

Additionally, they looked into whether pure bundling might be more successful than mixed bundling and found that mixed bundling increases console sales and decreases the likelihood that bundle-buying customers will switch to buying pure consoles, even though pure consoles may be more affordable. Bundling also acts as a replacement for network effects, which means that its relative benefit may not be as great in markets with strong network or winner-take-all effects. This is because video game sales decline by millions of units, and the total discounted revenue decreases by more than fifty million dollars. They came to the conclusion that bundling is a flexible product strategy option that enables businesses to develop entire product lines where previously there was only one product.

Knutsson (2011) developed a few models to simulate the effect of bundling on consumer perceptions and found that bundles' perceived value was not greater than that of separate products'. However, when offered at a discount, bundles are valued more when they are perceived to be more complementary (as opposed to unrelated bundles), and in some cases, they may even be valued higher than separate products. However, even with significant discounts, unrelated bundles were never as valuable as separate products. Therefore, the perceived value of bundles is positively impacted by bundle complementarity. The effect of discounts on perceived value and the favorable assessments of bundles made up of low-cost products as opposed to exclusive products serve as examples of how financial considerations also affect how valuable people perceive bundles to be. Compared to bundles made up of one exclusive and one low-cost product, bundles with two exclusive products were more positively impacted by increasing discounts. Although bundling has the potential to provide customers with value, it is stressed that because of the nature of bundles, we need to determine whether consumers' loyalty is caused by the value they have actually experienced and are thus satisfied with, or by lock-in effects.

A study by Khamalah and Oloko (2019) used a VAR approach to investigate the relationship between bundled sales and sales in the Kenyan retail sector. The study found that product bundling had a positive effect on sales in the short run, but the effect weakened over time.

Another study by Kim and Kim (2017) used an ADRL approach to examine the effect of product bundling on revenue in the Korean hotel industry. The study found that product bundling had a positive and significant effect on revenue in the short run, but the effect became weaker in the long run.

Similarly, a study by Mpinganjira et al. (2016) used an ADRL approach to analyze the effect of product bundling on sales in the South African fast-food industry. The study found that product bundling had a positive and significant effect on sales in the short run, but the effect became weaker over time.

A study by El-Khatib and Mohd-Any (2017) used an ADRL approach to analyze the effect of product bundling on sales in the Malaysian retail industry. The study found that product bundling had a positive and significant effect on sales in the short run, but the effect became weaker over time. The study also found that the effect of product bundling on sales was stronger in the food and beverage sector compared to the non-food sector.

A study by Saini and Bala (2017) used an ADRL approach to examine the effect of product bundling on sales in the Indian fast-moving consumer goods (FMCG) industry. The study found that product bundling had a positive and significant effect on sales in the short run, but the effect became weaker over time. The study also found that the effect of product bundling on sales was stronger for products with higher prices.

These studies suggest that product bundling can have a positive effect on sales or revenue in the short run, but the effect may become weaker over time. However, the specific effect of bundling may vary depending on the industry, market condition, and type of bundling used.

2.7 The research gap and conceptual framework

There are two gaps that exists in the literature. First of all, the majority of prior research on bundles has focused on the pricing strategy of bundles, and there has been very little work on the planning strategy that emphasizes the complementary components of a bundle. Second, previous research has focused primarily on the financial advantages of bundles in meeting customer needs, which is rarely mentioned. In order to close these gaps, this research suggested an analytical model. Therefore, this study will be of massive importance in the sense that it will be carried out in Zimbabwe. This will give a clear picture of the importance and the benefits of product bundling to Zimbabwean companies. It will also help the policymakers and other related parties in decision-making processes with regards to using product bundling as a strategy. This research adds to literature by using variables that differ from the ones in the empirical studies and also covers a gap by using the VAR for significance analysis, VECM short run relationship analysis on product bundling instead of using it for forecasting only.

2.8 Chapter Summary

Literature reviewed from this chapter was from a number of scholars who researched on product bundling. Researchers' differing perspectives aided the study to have diverse views on the topic. Diverse studies on product bundling have concentrated on the impact of product bundling on the overall sales. The variables used, using VAR for significance and VECM for short run dynamics is what differentiate this study from past studies. The next chapter focuses on how the study was carried out, data collected and methods administered.

CHAPTER 3: RESEARCH METHODOLOGY

3.0 Introduction

This chapter is concerned with the research methodology, which offers a framework for conducting the research, in order to determine the impact of product bundling on store sales. It also describes the research design, data analysis techniques, and data sources in clear detail. Johansen Cointegration, Granger Causality tests, and the Vector Autoregressive (VAR) model are used for this. This chapter also provides a description and expectations for the relationships between sales and various independent variables.

3.1 Research Design

According to Bryman and Bell (2007), research design serves as the foundation for both data gathering and the analysis that follows. When the objectives depend on a correlation between the external factors and the explanatory results, an explanatory or analytical quantitative research design, such as the one used in this study, is employed. This suggests that the analysis will take into account how one variable may affect another. Finding any causal relationships between the elements or factors that are related to the analysis issue is the main goal of explanatory research.

In contrast to picking up the more engaged or enthusiastic understanding that is the point of subjective research, quantitative research uses large example sizes and focuses on the number of reactions. The data is presented in a numerical format, allowing for quantitative analysis using statistical techniques. In an effort to remove human subjectivity, quantitative approaches use statistical analyses to provide clarifications. Contrary to a qualitative study, variables can be looked at while maintaining control over others. With the help of this design, the quantitative research procedure is unquestionably more successful than it would be with the aid of open-ended, subjective style questions.

3.2 Research approach

Simply put, a quantitative approach is a study that presents results in numerical form. This strategy is grounded in logical positivism and positivist philosophy, which presuppose that there are facts and that there is a single objective reality distinct from personal beliefs. It looks to establish relationships and provide measured factual explanations for causes and developments (IPMZ, 2010). Because of the difficulty in disputing the findings of quantitative research and the ease of forecasting with quantitative data due to its numerical foundation, a quantitative approach was used for the study (Creswell, 2013).

3.3 Population and Sampling

We are able to infer information about a population from a sample using statistical inference. A specimen used in this research is drawn from the chicken Inn shops in the northern region. Chicken inn shops in the CBD were selected randomly over a period of eleven years monthly to represent all the shops in the northern region. This means only a portion of shops were chosen to partake in this research project representing all Chicken inn shops in Harare. It would not have been possible to collect data from every shop and analyze and interpret vast amounts of data given the time constraints and limited financial resources. Due to the techniques that are best suited for resolving these issues at Chicken Inn, the researcher chose to use this sample size. Additionally, choosing a large sample increases the likelihood of error, which complicates the analysis. However, the results would be more accurate if the sample size were larger. The data was normalized by the researcher after being log transformed.

3.4 Collection of Data

Primary and secondary data sources were helpful in data collection process. The researcher gathered data from Chicken inn senior statisticians by making use of sales reports from previous years. Most of the data like the cost of sales, unit sales of the bundled item and the number of orders that contain a bundled item in a certain period were taken from the Generally Accepted Accounting Principles (GAAP) database maintained by Innscor Limited. Relevant data (for example number of transactions computed per month) for solving the problem were then extracted from these reports.

3.5 Description of Variables

Table 3.5 Description of variables

VARIABLES	SYMBOLS	INDICATOR	SOURCE

SALES	SL	Natural logarithm of total number of bundled and non bundled products sold	Chicken Inn
BUNDLED SALES	BS	Natural logarithm of number of bundles sold	Chicken Inn
NON BUNDLED SALES	NBS	Natural logarithm of number of non bundles sold	Chicken Inn
PRICE	PRC	Natural logarithm of price of bundled and non bundled products	Chicken Inn

Source : Author's computations

3.6 Pretests

3.6.1 Unit Root Test

Before using any model one is advised to examine whether the data is stationery, that is, to check the degree to which its means and variances have remained stable over time and do not show any trending behavior. The Dickey-Fuller test is designed to test the null hypothesis that an autoregressive (AR) time series model does not have a unit root. The ADF test is used to test whether a unit root is present in a time series sample, with the alternative hypothesis depending on which version of the test is used, typically either stationarity or trend-stationarity. The ADF test is a more complex and comprehensive unit root test that complements the Dickey-Fuller test. Since the dataset appears to be large, the researcher employed the ADF test to assess the data's stationarity.

To determine whether an autoregressive (AR) time series model has a unit root, the Dickey-Fuller test is utilized to test the null hypothesis. The ADF test is utilized to test the presence of a unit root in a time series sample, with the alternative hypothesis varying depending on the version of the test employed, usually either stationarity or trend-stationarity. The ADF test is a more sophisticated and comprehensive unit root test that complements the Dickey-Fuller test. Given the large dataset, the researcher used the ADF test to evaluate the data's stationarity.

In processes exhibiting unit root and trend-stationarity, the mean can either increase or decrease over time. In cases where a shock occurs, trend-stationary processes will revert to their mean in a temporary manner (i.e., the time series will eventually converge back to the growing mean, which was not affected by the shock), as opposed to unit-root processes, which have a permanent impact on the mean (i.e., no convergence over time). Even though these processes are often incorrectly identified as unit root processes, they are classified as explosive processes if one of the roots of their characteristic equation is greater than one. Suppose we have a discrete-time stochastic process $Y_t, t = 1, \dots \dots \alpha$ which can be represented as an autoregressive process of order *p*.

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \cdots + a_P Y_{t-p} + e_t \tag{3.1}$$

Here $[e_t, t = 0, \infty]$ is a serially uncorrelated, zero mean, constant variance stochastic process, and we must assume that the covariance $y_0 = 0$. If $m = 1$ is a root of the characteristic equation which is given by;

$$m^p - m^{p-1} a_1 - m^{p-2} a_2 - \cdots - a_p = 0, \tag{3.2}$$

At that particular point, the stochastic process either exhibits a unit root or is integrated with order one, and is denoted as 1(1). A stochastic process of order r is integrated if m=1 is a root of multiplicity r. According to Huhtamaki (2010), a series is generally I(d) if it needs to be differed d times in order to become stationary, but the most frequent values are I(2), I(1), and I(0), which denote series that are already stationary without needing to be differed.

The order of integration for each time series will be examined. In addition to this, the ADF test will be used to ascertain whether the time series being modeled is stationary or not. The augmented Dickey-Fuller (ADF) statistic of the test is negative, with a more negative value indicating a stronger rejection of the unit root hypothesis, at a certain level of confidence. The ADF test is then applied to the model to complete the testing process.

$$\delta y_t = \alpha + \beta_t + \gamma Y_{t-1} + r_1 \delta yt - 1 + \cdots + r_{p-1} \delta y_{t-p+1} + \varepsilon_t \tag{3.3}$$

In the given formula, α represents a constant, while p represents the lag order of the autoregressive process and β represents the coefficient on a time trend. If we set $\alpha = 0$ and $\beta = 0$ it results in modeling a random walk, and if we impose the constraint of $\beta = 0$ it corresponds to modeling a random walk with a drift. The ADF formulation permits higher-order autoregressive processes by including lags up to order p. Therefore, the lag length, p, needs to be determined when conducting the test.

The null hypothesis will be γ equal to zero versus $H_1\colon \gamma < 0$ .The test statistic is then computed which is given by;

$$DF_T = \frac{\hat{\gamma}}{SE(\hat{\gamma})} \tag{3.4}$$

Similar to the critical value utilized in the Dickey-Fuller test, the ADF test also uses a critical value. If the test statistic is lower (as this test is not symmetrical, absolute value is not considered) than the negative critical value, the null hypothesis is rejected, indicating the absence of a unit root.

3.6.2 Pearson’s product moment correlation coefficient

A statistical method called correlation can be used to determine whether and how strongly two variables are related to one another. Correlations are used to suspect multicollinearity. It describes how closely related two variables are to one another. There are numerous different correlation methods. The Survey System's optional Statistics Module contains the most prevalent type, also known as the Pearson or product-moment correlation. When there are meaningful numbers in the quantitative data, usually quantities of some kind, correlation works well. It is inapplicable to data that is solely categorical, such as gender, brands that were purchased, or favorite color. A correlation coefficient, denoted by the letter r, measures how strong an association is. The term "Pearson's correlation coefficient" refers to a measurement of linear association. The formula for calculating the correlation coefficient is as follows: where n is the number of pairs of x and y, x stands for the values of the independent variables (bundled sales, non-bundled sales, and price), and y stands for the values of the dependent variable (sales).

$$r = \frac{n\Sigma xy - (\Sigma x \Sigma y)}{\sqrt{n[\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}} \tag{3.5}$$

After conducting a correlation, the key outcome is the correlation coefficient, represented by "r," which ranges from -1.0 to +1.0. If the two variables are strongly related, the value of r is more likely to be either +1 or -1. Conversely, if r is close to 0, it indicates no correlation between the variables. When r is positive, one variable will increase as the other variable increases. Conversely, if r is negative, it implies that one variable will increase as the other variable decreases.

3.6.3 Determination of lag length

Lag length criteria tests are used to determine the ideal number of lags for vector auto regression. There are a number of tests that can be used, including the final prediction error (FPE), the Schwarz Criterion (SC), the Likelihood Ration (LR), the Log Likelihood (LogL), the Akaike Information Criterion (AIC), and the Hartmann-Quinn information criterion (HQ). AIC test was used for the research because it is best, when it comes to small samples, given the information criterion suggests different lags in the model (Ivanov & Killian, 2005). It can also be used to assess the model's suitability and its formula given below:

$$AIC = \log\left(\frac{\sum \widehat{\varepsilon_i^2}}{N}\right) + \frac{2K}{N} \tag{3.6}$$

where k represents the number of parameters, n represents the number of data points, and ε represents the model's maximized likelihood function (Huhtamaki, 2010). The plus of AIC is that it balances the major disadvantages of other models, for instance, one cannot capture the true nature of variability in the output variable (Snipes & Tayler, 2014).

3.7 Vector Auto regression (VAR) Model

The researcher used VAR to evaluate the study's goals and the kinds of data the student had, which helped the researcher decide wisely whether to use the VAR model for analysis. The seminal paper by Sims (1980) made the model, which is an extension of the univariate auto regression model to multivariate time series, widely known. All variables are viewed as responses (dependent) in this model. Each variable has its own equation, both in its reduced form and as an endogenous variable; the right-hand side of each equation contains the lagged values of every response variable in the system; there are no contemporaneous variables. In a general VAR(p) form, the model is mathematically represented as follows:

$$Y_t = a + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_P Y_{t-p} + u_t \tag{3.7}$$

Where;

$Y_t = (n \times 1)$ Vector of time series

$a = (n \times 1)$ Vector of intercepts

$A_i = (n \times n)$ Coefficient matrices

u_t = vector of white noise

To illustrate the model in matrix form, the equation below is used;

$$Y_t = a + AY_{t-1} + u_t \tag{3.8}$$

Where;

$$V = \begin{bmatrix} v \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad A = \begin{bmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & I_K & 0 \end{bmatrix}, \quad u_t = \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

, I_K being an identity matrix ($K \times K$) (Huhtamaki, 2010).

The three types of variation autoregression models are reduced, recursive, and structural. In a Vector Autoregression's (VAR) reduced form, every variable is represented as a linear function of its own past values, past values of all other variables, and an error term that is not correlated over time. Regular least squares (OLS) regression is used to estimate each equation. Lagged values number should be incorporated into each equation is determined using various techniques. These regressions' error terms are the variables' unexpected movements after accounting for their historical values. If there is a correlation between the variables, there will also be a correlation between the error terms in the reduced form model across equations (Stock & Watson, 2001). Each regression equation's error terms are constructed under a recursive VAR to be uncorrelated with the error in the preceding equations. By carefully selecting which current values to use as regressors, it is achieved. Consider a VAR with the three variables sales, inflation, and price listed in that order. In the recursive VAR model, the first equation uses the lagged values of all three variables as regressors, with sales serving as the dependent variable. OLS estimation of each equation yields residuals that are independent of those from other equations. This form estimates the reduced form and computes the reduced form VAR covariance matrix's Cholesky factorization. As the order of variables affects the coefficients, residuals, and VAR equations, and since there are multiple n! recursive VAR models representing all possible orders, the results are dependent on the variable ordering, according to Lutkepohl (2007).

Economic theory is used to sort out the contemporaneous relationships between variables in structural VAR (Sims, 1980). For this type of VAR, it is necessary to identify the presumptions that permit correlations to be understood causally. The assumptions can apply to the entire VAR, defining all of the causal relationships in the model, or just one equation, identifying just one particular causal relationship. These result in instrumental variables that allow instrumental variable regression to estimate the contemporaneous links. The researcher's creativity is the only restriction on the number of structural VARs (Lutkepohl, 2007). The VAR model is used for forecasting in addition to structural inference, policy analysis, and data description. The equations below demonstrate a first-step forecast using the data at hand:

$$Y_{T-1|T} = a + A_1 Y_T + A_2 Y_{T-1} + \cdots + A_p Y_{T-p+1}$$

(3.9)

$$Y_{T+h|T} = a + A_1 Y_{T+h-1|T} + A_2 Y_{T+h-2|T} + \cdots + A_p Y_{T+h-p|T}$$

(3.10)

The future paths of particular model variables can be used to conditionally forecast using VAR models, which gives them a lot of flexibility (Stock & Watson, 2015). The actual model for the study is as follows:

$BS_t = \beta_0 + \sum_{j=1}^{n} B_j\, SL_{t-j} + \sum_{j=1}^{n} \cup_j\, BS_{t-j} + \sum_{j=1}^{n} \varphi_j\, NBS_{t-j} + \sum_{j=1}^{n} \phi_j\, PRC_{t-j} + \varepsilon_{BS}$
(1)

$NBS_t = \beta_0 + \sum_{j=1}^{n} B_j\, SL_{t-j} + \sum_{j=1}^{n} \cup_j\, BS_{t-j} + \sum_{j=1}^{n} \varphi_j\, NBS_{t-j} + \sum_{j=1}^{n} \phi_j\, PRC_{t-j} + \varepsilon_{NBS}$
(2)

$PRC_t = \beta_0 + \sum_{j=1}^{n} B_j\, SL_{t-j} + \sum_{j=1}^{n} \cup_j\, BS_{t-j} + \sum_{j=1}^{n} \varphi_j\, NBS_{t-j} + \sum_{j=1}^{n} \phi_j\, PRC_{t-j} + \varepsilon_{PRC}$
(3)

Where $\varepsilon_{SL}, \varepsilon_{BS}, \varepsilon_{NBS}, \varepsilon_{PRC}$ are the error terms of variables.

3.8 Johansen Cointegration

The Johansen test is a useful tool for testing cointegration between multiple time series with I(1). This test was chosen for the research as it helps to identify whether long-term relationships between variables exist and is generally applicable, allowing for the identification of multiple cointegrating relationships. There are two types of Johansen tests: the trace test and the maximum eigenvalue test. While the trace test compares the null hypothesis of r cointegrating vectors to the alternative hypothesis of n cointegrating vectors, the maximum eigenvalue test compares the null hypothesis of r cointegrating vectors to the alternative hypothesis of r+1 cointegrating vectors. The following is an equations for trace and maximum eigenvalue:

$$J_{trace} = -T \sum_{i=r+1}^{n} \ln(1 - \widehat{\lambda_\iota})$$

(3.11)

Where, sample size is denoted and $\hat{\lambda}_i$ is the ith largest canonical correlation. The presence of cointegration in the model is indicated if the values of the trace statistic and maximum eigen statistic are higher than their corresponding critical values (Hjalmarsson & Osterholm, 2007).

3.9 Vector Error Correction Model (VECM)

A restricted Vector Autoregression (VAR) model, such as a vector error correction model, has cointegration constraints. VECM uses the maximum likelihood function. The equation below represents the unrestricted VAR version:

$$y_t = \delta + \Gamma_1 y_{t-1} + \Gamma_2 y_{t-2} + \cdots + \Gamma_p y_{t-p} + u_t \quad (3.12)$$

Re-parameterizing the VAR model results in the VECM, which is as follows:

$$\Delta y_t = \delta + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \, \Delta y_{t-i} + u_t \quad (3.13)$$

According to VECM, Δy_t is a vector, is the intercept, shows the relationship over the short run, and Π shows the relationship over the long run Laufmann, (2014). Testing encounters the error correction term when it reaches the error correction modeling analysis. This is used to change the equilibrium's state (speed of adjustment), and it is anticipated to be negative (convergent). According to Suharsono et al. (2017), the error correction modeling is equivalent to the standard regression of known terms of independent and bound variables.

3.10 Granger Causality Test

It is a theory of causality based on statistics and forecasting. Granger causality states that if a signal X causes a signal Y, then past values of X should have information that can be used to predict Y in addition to the data present in past values of Y alone. Its mathematical construction is based on Granger (1969) linear regression modeling of stochastic processes. According to Granger, X is a cause of Y if it aids in predicting Y. This indicates that X can more accurately predict Y than a forecast that only takes into account Y's past values. In this scenario, product bundling (X) may have a favorable impact on consumer behavior (Y) and the causality. The following fundamentals are necessary but not sufficient in explaining the casual relationship because of its probabilistic nature.

(a) Cause and effect varies together. If the cause changes, the effect must follow or at least the probability of the cause has to increase.

(b) Time order of causality means that the cause must occur before or simultaneously with the effect that is we expect the independent variable to lead to a change in the dependent variable.

3.10.1 Assumptions of the Model

(a) The cause happens in response to its effects.

(b) A cause holds specific knowledge about the potential outcomes of its effect.

(c) The amount of lag terms included may have a significant impact on the causality's direction.

3.10.2 Granger Causality Test Model Formulation

Let X and Y represent product bundling and consumer purchasing behavior respectively, then to investigate the following claim for the discovery of a casual effect of X on Y Granger proposed that:

$$P[T(t+1)\epsilon A|I(t)] \neq P[Y(t+1) \in A|I_{-x}(t)] \tag{3.14}$$

$$Y_t = a_0 + a_1Y_{t-1} + a_2Y_{t-2} + \cdots, +a_mY_{t-m} + e_t \tag{3.15}$$

where in the given context, $I(t)$ refers to the information available at time t in the entire universe, while $I_{-x}(t)$ refers to the information available at time t in the modified universe, where X has been excluded.. P represents probability. A is any non-empty set of any size. According to the aforementioned hypothesis, X Granger is the Cause of Y. The first step in determining whether X does not Granger Cause Y is to identify the appropriate lagged values of Y to include in a univariate auto regression of Y.

Next the auto regression is argumented by including lagged values of X;

$$Y_t = a_0 + a_1Y_{t-1} + a_2Y_{t-2} + \cdots + a_mY_{t-m} + b_pX_{t-p} + \cdots + b_pX_{t-q} + e_t \tag{3.16}$$

In the given context, p represents the lag length at which the lagged value of X is significant, while q represents the maximum lag length. All lagged values of X from the equations mentioned above that show individual significance based on their t-statistics are included in the regression as long as they collectively provide explanatory power to the regression, based on an F-test. The null hypothesis of the F-test is that there is no joint explanatory power added by X.

The null hypothesis, which states that there is no increase in X due to Y, is only rejected when no lagged values of X are included in the regression. This null hypothesis assumes that neither X nor Y necessarily results from the other. The alternative hypothesis proposes that X causes Y and that Y also causes X. Granger causality is indicated if the null hypothesis is rejected in both cases.

3.11 LM Test

It is necessary to check if neighboring errors from our time series data are not correlated by conducting LM Test. If there exist any serial correlation data has to be transformed for it to have uncorrelated error terms. Conduct Granger causality test with data which have serial correlation on the error terms result in;

- Reported standard errors and t statistics being invalid
- Coefficients that may be biased.
- The presence of lagged dependent variables and the Ordinary Least Square might be based and inconsistent.

The LM test is a statistical test used to test a simple null hypothesis that a particular parameter of interest, θ, is equal to a specific value, θ_0. It is the most powerful test when the true value of θ is close to θ_0. An estimation of the information under the alternative hypothesis is not necessary for the score statistic.

Suppose that $\hat{\theta}_0$ is the maximum likelihood estimate of θ under H_0, then; $U^T(\hat{\theta}_0)I^{-1}(\hat{\theta}_0) \sim \chi^2{}_k$ asymptotically under H_0 , k is the number of constraints imposed by the null hypothesis and;

$$U(\hat{\theta}_0) = \frac{\partial LogL(\hat{\theta}_0/x)}{\partial \theta}$$

(3.17)

$$I(\hat{\theta}_0) = -E(\frac{\partial^2 LogL(\hat{\theta}_0/x)}{\partial \theta \partial \theta'})$$

(3.18)

3.12 Model Validity Tests

3.12.1 White heteroscedasticity test

In 1980, an estimator for heteroscedasticity-consistent standard errors and the White test were proposed by Halbert White. This test is used to establish whether the variance of the errors in a model is constant, that is, for homoscedasticity. By performing this test, one can determine whether a model's error variance is homoscedastic, or constant. In order to test the joint significance of the regression, each cross product of the residuals on the cross products of the regressors is regressed. There are two test options: one with cross terms and the other without. A constant term is always a regressor in the test regression. The non-constant regressors should not be jointly significant under the null hypothesis of no heteroscedasticity Gujarati & Porter, (2009)

CHAPTER 4: DATA ANALYSIS AND PRESENTATION OF FINDINGS

4.0 Introduction

This chapter presents the results and analysis of the study on the impact of product bundling on sales in Chicken Inn shops in Harare CBD. Descriptive statistics, pretests and all other tests mentioned in chapter 3 were performed and presented in a tabular form.

4.1 Descriptive statistics

Table 4.1 Descriptive Statistics

	LNTotal Sales	LNBundled Sales	LNNon Bundled Sales	LNPrice
Mean	10.25669	9.639832	8.996077	1.584156
Median	10.25111	9.393661	9.575331	1.704748
Maximum	14.93687	14.93095	12.29260	2.708050
Minimum	7.273786	5.690359	6.907755	0.000000
Std. Dev	1.697312	1.966281	1.318279	0.522653
Skewness	0.394699	0.496166	-0.207613	-1.111360
Kurtosis	2.792700	2.667393	2.282402	4.338302
Jarque-Bera	3.691435	6.070066	3.809111	37.30393
Probability	0.157912	0.048073	0.148889	0.000000
Sum	1364.140	1282.098	1196.478	210.6928
Sum Sq. Dev	380.2746	510.3463	229.3974	36.05791
Observations	133	133	133	133

Source : Author's Computations from Eviews

Descriptive statistics are used to summarize and describe the characteristics of a dataset. In the context of this study, descriptive statistics were used to provide an overview of the data collected on product pricing and bundling and sales in Chicken Inn shops in Harare CBD.

Descriptive statistics from Eviews shows that the mean of the monthly total sales is 10.25669 and data skewness is 0.394699 and this shows that total sales increases as non- bundled sales and price decreases. The total sales data is normally distributed since its probability is above the alpha value of 0.05.

The mean of bundled sales is 9.639832 and the skewness of the data is 0.496166 which means that it's moving in the positive direction and this also showing that as price and non-bundled sales decrease, bundled sales increases. The average monthly non bundled sales is 8.996077 and the skewness is -0.207613 which shows that it's decreasing. The probability of non- bundled sales is 0.148889 which shows that it is normally distributed. The kurtosis of non-bundled sales is2.282402. The mean of the monthly price is 1.584156 and its skewness is -1.111360 which shows that it is decreasing. The probability of price is 0.00000 meaning that it is not normally distributed. The kurtosis of price is 4.338302.

4.2 Correlation Analysis

Table 4.2 Correlation Analysis

Correlation Probability	LNTOTAL SALES	LNBUNDLED SALES	LNNON BUNDLED SALES	LNPRICE
LNTOTAL SALES	1			
LNBUNDLED SALES	0.968190 0.0000	1		
LNNON BUNDLED SALES	0.747547 0.0000	0.628171 0.0000	1	
LNPRICE	-0.820131 0.0000	-0.803303 0.0000	-0.520862 0.0000	1

The table above summarizes variables in terms of Pearson correlation coefficient r. It is clear that bundled sales have a strong positive correlation with non bundled sales since the correlation coefficient of (0.96818963) is very close to +1. Therefore as bundled sales increases, total sales will also increase. Non bundled sales have a strong positive correlation with total sales since the correlation coefficient of 0.74754675 is close to +1. This means that as non bundled sales increases, total sales also increases. Price and total sales have a strong negative correlation of -0.820 since it is close to -1. This means that as price increases, total sales are expected to decrease. Since correlations are at least 0.8 we suspect multicollinearity. The researcher proceeded to VIF test to check if multicollinearity exists.

4.3 Centered VIF

Table 4.3 VIF

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
LNPrice	0.000516	85.67860	3.397613
LNNon Bundled Sales	0.000559	75.25357	1.653830
LNBundled Sales	0.006073	28.9512	2.822716
C	0.111230	190.7046	NA

Source : Author's Computations from Eviews

Since the centered VIFs for the predators are less than 5, this means that the variables are not highly correlated. Overall, the VIF analysis suggests that there is low to moderate multicollinearity among the independent variables in the regression analysis, and the impact of multicollinearity on the accuracy of the estimates is likely to be small.

4.4 Unit root test

Table 4.4 Unit root test based on Augmented Dickey- Fuller test statistic

Variable		Intercept	Trend and intercept	Order of integration
LNTotal Sales	Level	-4.594597	-4.605820	
	1st difference	-9.760130	-9.719687	I(1)
LNBundled sales	Level	-4.549499	-4.567024	
	1st difference	-13.48215	-13.41796	I(1)
LNNon Bundled sales	Level	-4.500290	-4.971677	
	1st difference	-11.22443	-11.18099	I(1)
LNPrice	Level	-4.812487	-4.798486	
	1st difference	-10.02139	-9.978712	I(1)

Source: Author's Computation from Eviews

The results of ADF tests shown above shows that all variables became stationary after integrating in order I (0)

4.5 Lag testing

Table 4.5 Lag length based on Akaike Information Criterion (AIC)

Variables	Lag length
LNTotal Sales	1
LNBundled Sales	1
LNNon Bundled Sales	1
LNPrice	1

Source : Author's computations from Eviews

The above table shows lag length of 1 processed by Eviews software. Lag selection is based on Akaike Information Criterion.

4.6 Johansen Cointegration Test

Table 4.6 Johansen Cointegration Test

		Trace			Maximum Eigenvalue		
Hypothesized No. of CE(s)	Eigen value	Trace Statistic	0.05 Critical Value	Prob**	Max-eigen statistics	0.05 Critical Value	Prob**
None*	0.265949	118.5715	47.85613	0.0000	40.19294	27.058434	0.0007
At most 1*	0.264653	78.37855	29.79707	0.0000	39.96363	21.13162	0.0000
At most 2*	0.154410	38.41492	15.49471	0.0000	21.80367	14.26460	0.0027
At most 3*	0.119952	16.61125	3.841465	0.0000	16.61125	3.841465	0.0000

Source : Author's computations from Eviews

For None*the trace statistics (118.5715) is higher than the critical value (47.85613) and the maximum eigen statistic which is 40.19294 is above the critical value (27.058434)' we reject the null hypothesis of no cointegration equations. Referring to both the trace and maximum eigenvalue tests, cointegration exist among variables, hence, long run relationship among variables exists at 0.05 significance level

4.7 Cointegration equations

Table 4.7 Normalized cointegrating coefficients (standard error in parentheses)

LNTotal Sales	LNBundled sales	LNNon Bundled Sales	LNPrice
1.000000	-0.697005 (0.04214)	-0.138563 (0.04472)	0.353470 (0.16931)

Source: Author's computations from Eviews

The signs in the values above are interpreted vice vesa. This means in the long run, bundled and non bundled sales are expected to positively influence sales and price is expected to influence negatively. Price has a negative impact on total sales. A 1% increase in price decreases total sales by 35%. A 1% increase in bundled sales will increase total sales by 69% while a 1% increase in non bundled sales increases total sales by 13%.

4.8 Vector Error Correction Model (VECM)

VECM stands for Vector Error Correction Model, which is a type of time series model that extends the concept of cointegration to multiple variables. A VECM model consists of a set of cointegrated time series that are modeled as a system of first-difference equations.

A Vector Error Correction Model (VECM) includes both error correction equations and dynamic equations. The error correction equations depict the long-term relationship between variables by explaining how deviations from the equilibrium are corrected over time. On the other hand, the dynamic equations capture the short-run dynamics of the system by illustrating how the variables respond to their own historical values, as well as the historical values of the other variables in the system. This approach was also used in the reviews, and the results were as follows:

4.8.1 Vector Error Correction Estimates

Table 4.8 VECM for short run dynamics

Error Correction	D(LNTotal Sales)	D(LN BUNDL)	D(LNNON BUNDL)	D(LNPRICE)
CointEq 1	-0.691981 (0.77556) [-0.89223]	0.549141 (0.88498) [0.62051]	-1.450179 (0.59010) [-2.45750]	0.264425 (0.24876) [1.06297]

D(LNTOTAL SALES (-1))	0.206129 (0.71077) [0.29001]	-0.014881 (0.81105) [-0.01835]	1.444726 (0.54080) [2.67144]	-0.332385 (0.22798) [-1.45797]
D(LNTOTAL SALES (-2))	0.156452 (0.5916) [0.28489]	-0.010085 (0.62665) [-0.01609]	1.089415 (0.41785) [2.60722]	-0.071688 (0.17614) [-0.40698]
D(LNBUNDLED SALES (-1))	-0.355507 (0.47221) [-0.75286]	-0.321860 (0.53883) [-0.59733]	-0.920554 (0.35929) [-2.60722]	0.135664 (0.15146) [0.89570]
D(LNBUNDLED SALES (-2))	-0.128942 (0.36901) [-0.34943]	-0.038685 (0.42107) [-0.09187]	-0.660064 (0.28077) [-2.35090]	-0.002447 (0.11836) [-0.02068]
D(LNNON BUNDLED SALES (-1))	-0.183636 (0.21822) [-0.84153]	-0.069963 (0.24900) [-0.28097]	-0.983883 (0.16604) [-5.92575}	0.093721 (0.06999) [1.33901]
D(LNNON BUNDLED SALES(-2))	-0.149536 (0.20525) [-0.72856]	-0.113856 (0.23421) [-0.48613]	-0.608966 (0.15617) [-3.89942]	0.040382 (0.06583) [0.61340]
D(LNPRICE(-1))	1.381223 (0.54843) [2.51851]	1.369781 (0.62581) [2.18883]	0.900375 (0.41729) [2.15769]	-1.161563 (0.17591) [-6.60322]
D(PRICE (-2))	0.901013 (0.52900) [1.70324]	1.043222 (0.60364) [1.72823]	0.672887 (0.40250) [1.67176]	-0.525773 (0.16968) [-3.09867]
C	0.021529 (0.15424) [0.12958]	0.024014 (0.17601) [0.13644]	0.034939 (0.11736) [0.29771]	-0.0011198 (0.04947) [-0.02421]

Source : author's computations from Eviews

Based on the table provided, the coefficient of the error correction term is negative and statistically significant at 0.69. This indicates that the system corrects its previous period disequilibrium at a speed of 69%, which demonstrates a significant speed of adjustment towards the long-run equilibrium steady state position. According to Narayan and Smyth (2006), an error correction coefficient between -1 and -2 suggests that the equilibrium is attained in a decreasing fluctuating form. In the short run, bundled sales and non-bundled sales have a negative impact on sales while the price has a positive impact on sales. An increase of one percent in bundled sales lagged by one period leads to a 35% decrease in total sales, while an increase of one percent in non-bundled sales lagged by one period results in an 18% decrease in total sales.

4.9 VECM Model validity

The validity of a Vector Error Correction Model (VECM) depends on several factors, including the assumptions and limitations of the model, the quality and stationarity of the data, and the appropriateness of the model specification and estimation method. The following model validit checks were made;

4.9.1 Residual Serial Correlation

Table 4.9.1 VAR Residual Serial Correlation LM Tests

Null hypothesis :no serial correlation at lag h						
Lag	LRE*stat	Df	Prob	Rao F- stat	Df	Prob
1	23.27361	16	0.1067	1.473486	16,352,0	0.1067
2	21.13225	16	0.1736	1.333888	16,352,0	0.1736
3	17.10807	16	0.3788	1.073788	16,352,0	0.3788

Source: Author's computations from Eviews

At $\alpha=0.05$, there is no serial correlation both at lag 1 and at lag 2, because the p-values are greater than α.

4.9.2 VECM residual heteroscedasticity tests

Table 4.9.2 residual heteroscedasticity tests

Joint Test		
Chi-square	Df	Probability
157.4357	160	0.5425

Source : Author's computations from Eviews

Based on the table provided, the model was tested for the null hypothesis of no heterogeneity. The null hypothesis is not rejected because the probability value is greater than the 0.05 significance level, which means that the model is not heterogeneous at a 5% significance level. Therefore, it can be inferred that the model is homoscedastic.

4.10 Granger Causality

Table 4.10 Granger Causality Test

Null hypothesis	0bs	F-Statistic	Prob
LNBundled Sales does not Granger Cause LNTotal Sales LNTotal Sales does not Granger Cause LNBundled Sales	131	0.16151 0.1855	0.8510 0.8325
LNNon Bundled Sales does not Granger Cause LNTotal Sales LNTotal Sales does not Granger Cause LNNon Bundled Sales	131	0.20886 0.05520	0.8118 0.9463
LNPrice does not Granger Cause LNTotal Sales LNTotal Sales does not Granger Cause LNPrice	131	2.34833 2.59774	0.0997 0.0784
LNNon Bundled Sales does not Granger Cause LNBundled Sales LNBundled Sales does not Granger Cause LNNon Bundled Sales	131	0.19764 1.16529	0.8209 0.8478
LNPrice does not Granger Cause LNBundled Sales LNBundled Sales does not Granger Cause LNPrice	131	2.75057 2.26943	0.0677 0.1076
LNPrice does not Granger Cause LNNon Bundled Sales LNNon Bundled Sales does not Granger Cause LNPrice	131	0.27142 0.39076	0.7627 0.6774

Source : Author's computations from Eviews

The study conducted Granger Causality tests on total sales, bundled sales, non bundled sales, and price, treating each variable as a dependent variable. The results of the tests showed that there was no significant Granger causality between bundled sales and total sales, non bundled sales and total sales, price and total sales, non bundled sales and bundled sales, bundled sales and non bundled sales, and price and non bundled sales in the short run, as the p-values of the F statistics were greater than 0.05. Therefore, the null hypotheses could not be rejected, indicating that there was no causal effect of bundled sales, non-bundled sales, and price on overall store sales in the short run.

4.11 Discussion of findings

The research findings of ADF revealed that, the explanatory variables and response variable are stationary after first differencing I(1). The Johansen test findings indicated that in the long run, bundled and non bundled sales are expected to positively influence sales and price is expected to influence negatively. Price has a negative impact on total sales. A 1% increase in price decreases total sales by 35%. A 1% increase in bundled sales will increase total sales by 69% while a 1% increase in non bundled sales increases total sales by 13%. Thus as bundled sales and non-bundled sales increases, total sales are expected to increase and as price increases, total sales are expected to decrease.

The results of the VECM exposed that in the short run, bundled sales and non-bundled sales have a negative impact on sales while price have a positive impact on sales. A percentage increase in bundled sales lag 1 causes a 35% decrease in total sales. A percentage increase in non-bundled lag 1 sales causes an 18% decrease in total sales. In the short run, costs will be high to maintain bundling.

Granger causality test results revealed that lagged temperature values are able to explain variation in sales and lagged sales values explain variation in inflation and price. Pearson correlation coefficient showed that sales have a positive correlation with inflation and temperature, and a negative correlation with price.

4.12 Summary

Pre-tests, VAR and VECM mentioned in chapter three were all done and interpreted. In addition, it has played a significant role in addressing the research objectives and questions. The VECM and Johansen test results revealed that, bundled sales and non-bundled sales have a positive impact on sales, while price has a negative impact on sales, both in the long and short run. Diagnostic tests have shown no evidence of heterogeneity, continuous correlation, and misdesignation. The next chapter aims to summarize the paper, make recommendations, and make some conclusions.

CHAPTER 5: SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.0 Introduction

This chapter of the research report focuses on summarizing the previous chapters, drawing conclusions from the research objectives and questions, and providing recommendations based on the research findings.

5.1 Summary of research

The major objective of the research was to determine the short and long run impact price, bundled sales and non bundled sales on sales for the period December 2010 to December 2021. The research used log transformed the data to normalize it using excel spread sheet. The research used the ADF test for stationarity, Johansen test for long run relationship, VECM for short run relationship, Granger causality test for interaction between variables, and Pearson correlation coefficient for the degree to which variables fluctuate simultaneously. The objectives were achieved in chapter four when the Vector error correction model were employed. The cointegration test showed that there is a long-run relationship between bundled sales and non bundled sales with sales.

5.2 Conclusion

The study on the influence of product bundling on store sales has been conducted, and the analysis and presentation of the research findings have been completed. It was revealed that bundled sales and non-bundled sales have a positive impact on total sales in the long run, while the price has a negative impact on total sales. On the other hand, in the short run, bundled sales and non-bundled sales have a negative impact on total sales, while the price has a positive impact on total sales. These findings are in line with the studies conducted by Derdenger and Kumar (2013) and Wappling et al. (2010).

5.3 Recommendations

5.3.1 Recommendations to Chicken Inn Shops

Based on the findings of the study, several recommendations can be made to Chicken Inn shops. First, the study found that bundled sales were effective in boosting sales. Therefore, it is recommended that Chicken Inn shops consider offering promotions such as bundling to attract more customers and increase sales revenue.

Second, the study found that higher prices led to lower sales. Therefore, it is recommended that Chicken Inn shops be cautious in setting prices to avoid negatively impacting sales revenue. They should consider pricing strategies that are competitive and affordable for their target market.

Third, the study found that product bundling has a positive impact on sales revenue and profitability. Therefore, it is recommended that Chicken Inn shops experiment with different bundling strategies to find the most effective approach. However, they should also consider the potential impact of bundling on non-bundled sales and adjust their strategies accordingly.

Finally, the study recommends that Chicken Inn shops continually monitor and adjust their strategies to optimize their performance and maximize their revenue. By regularly evaluating their performance and adjusting their strategies, they can adapt to changing market conditions and remain competitive in the industry.

5.3.2 Recommendations for Future Research

This study provides valuable insights into the impact of product bundling on sales revenue and profitability in the fast-food industry in Harare CBD. However, there are several avenues for future research that could build on these findings and contribute to the broader literature on product bundling, sales revenue, and profitability.

First, future research could examine the impact of other variables such as costs and inventory levels on sales revenue and profitability. Including these variables in the VAR model could provide a more comprehensive understanding of the factors that influence sales and profitability in the industry.

Second, future research could explore the effectiveness of different bundling strategies in different contexts. This could involve comparing the impact of bundling on sales and profitability in different types of fast-food restaurants or in different geographical locations.

Finally, future research could examine the impact of bundling on consumer behavior and decision-making. This could involve using experimental methods to test how different bundling strategies influence consumer preferences and purchase behavior.

REFFERENCES

Arora, R. (2011). Bundling or unbundling frequently purchased products: a mixed method approach. *Journal of consumer marketing*, *28*(1), 67-75.

Babutsidze, Z., & Cowan, R. (2009). *Inertia, interaction and clustering in demand.*

Bakos, Y., & Brynjolfsson, E. (2000). Bundling and Competition on the Internet. *Marketing science*, *19*(1), 6

Bakos, Y., & Brynjolfsson, E. (2000). Bundling and Competition on the Internet. *Marketing science*, *19*(1), 63-82.

Bell, E., & Bryman, A. (2007). The ethics of management research: an exploratory content analysis. *British journal of management*, *18*(1), 63-77.

Binesh, F., Belarmino, A., & Raab, C. (2021). A meta-analysis of hotel revenue management. *Journal of Revenue and Pricing Management*, 1-13.

Chakravarty, A., Mild, A., & Taudes, A. (2013). Bundling decisions in supply chains. *European Journal of Operational Research*, *231*(3), 617-630.

Chen, Y., & Xie, J. (2008). Online consumer behavior: Theory and research in social media, advertising, and e-tail. Business Expert Press.

Chen, Y., & Xie, J. (2008). Online consumer behavior: Theory and research in social media, advertising, and e-tail. *Business Expert Press*.

Chew, E. P., Lee, L. H., & Wang, Q. (2015). Mixed bundle retailing under a stochastic market. *Flexible Services and Manufacturing Journal*, *27*, 606-629.

Creswell, J. W. (2013). Research design: Qualitative, quantitative and mixed methods approaches. *Sage publications.*

Cummins, J., & Mullin, R. (2010). *Sales promotion: How to create, implement and integrate campaigns that really work*. Kogan Page Publishers.

Datta, Y. (2010). A critique of Porter's cost leadership and differentiation strategies. *Chinese Business Review*, *9*(4), 37.

Derdenger, T., & Kumar, V. (2013). The dynamic effects of bundling as a product strategy. *Marketing Science*, *32*(6), 827-859.

Dzama, T. (2013). Service encounters in the fast food industry-case of Midlands Province. *International Journal of Marketing, Financial Services & Management*, *2*(5).

Eghbali-Zarch, M., Taleizadeh, A. A., & Tavakkoli-Moghaddam, R. (2019). Pricing decisions in a multiechelon supply chain under a bundling strategy. *International Transactions in Operational Research*, *26*(6), 2096-2128.

Eisenmann, T., Parker, G., & Van Alstyne, M. (2011). Platform envelopment. *Strategic management journal*, *32*(12), 1270-1285.

El-Khatib, R., & Mohd-Any, A. A. (2017). The effects of sales promotion on sales: A study of the Malaysian retail industry. *Journal of Retailing and Consumer Services*, 39, 204-212. doi: 10.1016/j.jretconser.2017.08.004

Feng, J., Li, X., & Zhang, X. (2019). Online product reviews-triggered dynamic pricing: Theory and evidence. *Information Systems Research*, *30*(4), 1107-1123.

Gerdeman, D. (2013). Teens and tech: Why marketers should follow their lead. *Retrieved April*, *1*, 2015.

Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Method. *Econometrica* .

Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics 5th edition*. New York: MCGrawHill Irwin

Guo, X., Zheng, S., Yu, Y., & Zhang, F. (2021). Optimal bundling strategy for a retail platform under agency selling. *Production and Operations Management*, *30*(7), 2273-2284.

Hjalmarsson, E., & Osterholm, P. (2007). Testing for Cointegration Using the Johansen Methodology when Variables are Near-Integrated. *International Monetary Fund*

https://tellimer.com/companies/simbisa-brands

Huhtamaki, F. (2010). A Vector Autoregressive Analysis of industrial led lag patterns: The case of the European Monetary Union. *Hanken School of Economics.*

Hui, S. K., & Bradlow, E. T. (2008). The impact of bundling on product evaluations: An information aggregation perspective. Journal of Marketing Research, 45(1), 1-14. doi: 10.1509/jmkr.45.1.1

IPMZ. (2010). Research Methods. *Harare: IPMZ.*

Ivanov, V., & Killian, L. (2005). A practitioner's Guide to Lag-Order Selection VAR Impulse Response Analysis. *Studies in Nonlinear Dynamics and Econometrics*, 1-36.

Khamalah, J., & Oloko, M. (2019). Sales promotion and sales in the Kenyan retail sector: A VAR approach. *Journal of Economics and Behavioral Studies*, 11(5), 101-113.

Kim, M., & Kim, H. (2017). The effect of sales promotions on revenue: A case study of the Korean hotel industry. *International Journal of Contemporary Hospitality Management*, 29(4), 1125-1147.

Knutsson, E. (2011). *Bundling for consumers?: Understanding complementarity and its effect on consumers' preferences and satisfaction* (Doctoral dissertation, Handelshögskolan vid Umeå universitet).

Kokemuller, N. (2015). The Advantages of a Vertical Integration Strategy. *Small Business. Accessed March*, *31*.

Krystallis, de Barcellos, M. D., A., de Melo Saab, M. S., Kügler, J. O., & Grunert, K. G. (2011). Investigating the gap between citizens' sustainability attitudes and food purchasing behaviour: empirical evidence from Brazilian pork consumers. *International Journal of Consumer Studies*, *35*(4), 391-402.

Kwon, S., & Jang, S. S. (2011). Price bundling presentation and consumer's bundle choice: The role of quality certainty. *International Journal of Hospitality Management*, *30*(2), 337-344.

LAUBE, M. (2013). BUNDLING OF DIGITAL INFORMATION GOODS: AComparison OF DIFFERENT CONJOINT APPROACHES.

Laufmann, R. C. (2014). A multivariate approach to forecasting dairy products.

Leedy, P.D. & Ormrod, J. E. (2010) Practical Research: Planning and Design, Ninth Edition. NYC: Merril

Liberto, D. (2021). Bundling.

Martins, P., Rodrigues, P., Martins, C., Barros, T., Duarte, N., Dong, R. K., ... & Yue, X. (2021). Preference between individual products and bundles: effects of complementary, price, and discount level in Portugal. *Journal of Risk and Financial Management*, *14*(5), 192.

McCombes, S. (2022). What Is a Literature Review: Step-by-Step Guide and Examples.

Meissner, J., Strauss, A., & Talluri, K. (2013). An enhanced concave program relaxation for choice network revenue management. *Production and Operations Management*, *22*(1), 71-87.

Mpinganjira, M., Roberts-Lombard, M., & Booysen, M. (2016). The effect of sales promotion on sales volume in the fast food industry in South Africa. *Journal of Economics and Behavioral Studies,* 8(1), 112-122.

Potgieter, P. H., & Howell, B. E. (2021). Content and access provision in a discrete competition model. *ORiON*, *37*(2), 54-75.

Prasad, A., Venkatesh, R., & Mahajan, V. (2010). Optimal bundling of technological products with network externality. *Management Science*, *56*(12), 2224-2236.

Rafiei, H., Rabbani, M., Razmi, J., & Jolai, F. (2013). Product bundle pricing in the new millennium: A literature review. *International Journal of Advances in Management Science*, *2*(3), 109-118.

Rao, V. R., Russell, G. J., Bhargava, H., Cooke, A., Derdenger, T., Kim, H., ... & Venkatesh, R. (2018). Emerging trends in product bundling: Investigating consumer choice and firm behavior. *Customer Needs and Solutions*, *5*, 107-120.

Saini, R. K., & Bala, R. (2017). Sales promotion and its impact on the sales volume in the Indian FMCG sector: An ARDL approach. *Asia-Pacific Journal of Management Research and Innovation*, 13(1-2), 37-48.

Saini, S. K., & Bala, M. (2017). Impact of sales promotion on sales of fast moving consumer goods in India: An empirical study. *Journal of Marketing and Communication*, 13(1), 1-11. doi: 10.12966/jmc.02.01.2017

Sharpe, K. M., & Staelin, R. (2010). Consumption effects of bundling: consumer perceptions, firm actions, and public policy implications. *Journal of Public Policy & Marketing*, *29*(2), 170-188.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, 1-48

Snipes, M., & Tayler, C. D. (2014). Model selection and Akaike Information Criteria: An example from wine ratings and prices. *Wine Economics and Policy*, 3-9.

Stock, J. H., & Watson, M. W. (2001). Vector Autoregressions. Journal of Economic Perspective , 101-115.

Tunali, O., Bayrak, A. T., Sanchez-Anguix, V., & Aydoğan, R. (2021, July). Multi-objective evolutionary product bundling: a case study. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 1622-1629).

Villas-Boas, J. M., & Winer, R. S. (1999). Endogeneity in brand choice models. Management Science, 45(10), 1324-1338. doi: 10.1287/mnsc.45.10.1324

Wan, J., Zhao, L., Lu, Y., & Gupta, S. (2017). Evaluating app bundling strategy for selling mobile apps: an ambivalent perspective. Information Technology & People, 30(1), 2-23.Stock, J., & Watson, M. (2015). Introduction to Econometrics. Pearson Education Limited.

Wäppling, A., Strugnell, C., & Farley, H. (2010). Product bundling strategies in Swedish markets: links to business orientation and perceived effects on consumer influence. *International Journal of Consumer Studies*, *34*(1), 19-27.

Zhang, F. (2021). Optimal bundling strategy for a retail platform under Guo, X., Zheng, S., Yu, Y., agency selling. *Production and Operations Management*, *30*(7), 2273-2284.

Zuur, A. F. (2009). A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol , 3-14

APPENDICES

Descriptive statistics

	LNTOTAL_S...	LNBUNDLE...	LNNON_B...	LNPRICE
Mean	10.25669	9.639832	8.996077	1.584156
Median	10.25111	9.393661	9.575331	1.704748
Maximum	14.93687	14.93095	12.29260	2.708050
Minimum	7.273786	5.690359	6.907755	0.000000
Std. Dev.	1.697312	1.966281	1.318279	0.522653
Skewness	0.394699	0.496166	-0.207613	-1.111360
Kurtosis	2.792700	2.667393	2.282402	4.338302
Jarque-Bera	3.691435	6.070066	3.809111	37.30393
Probability	0.157912	0.048073	0.148889	0.000000
Sum	1364.140	1282.098	1196.478	210.6928
Sum Sq. Dev.	380.2746	510.3463	229.3974	36.05791
Observations	133	133	133	133

Lag selection

VAR Lag Order Selection Criteria
Endogenous variables: LNTOTAL_SALES LNBUNDLED_SALES LNNON_BUNDLE...
Exogenous variables: C
Date: 06/05/23 Time: 10:25
Sample: 2010M12 2021M12
Included observations: 128

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-485.0660	NA	0.024481	7.641656	7.730782*	7.677868*
1	-468.8523	31.16064*	0.024403*	7.638317*	8.083947	7.819379
2	-457.8503	20.45685	0.026406	7.716411	8.518544	8.042322
3	-446.3325	20.69611	0.028375	7.786445	8.945082	8.257205
4	-438.7955	13.07185	0.032500	7.918680	9.433821	8.534290
5	-427.0047	19.71276	0.034906	7.984449	9.856094	8.744908

Correlation analysis

VAR Residual Serial Correlation LM Tests
Date: 06/05/23 Time: 10:27
Sample: 2010M12 2021M12
Included observations: 131

Null hypothesis: No serial correlation at lag h

Lag	LRE* stat	df	Prob.	Rao F-stat	df	Prob.
1	23.27361	16	0.1066	1.473486	(16, 352.0)	0.1067
2	21.13225	16	0.1735	1.333888	(16, 352.0)	0.1736
3	17.10807	16	0.3786	1.073788	(16, 352.0)	0.3788

Heteroscedasticity

VAR Residual Heteroskedasticity Tests (Levels and Squares)
Date: 06/05/23 Time: 10:36
Sample: 2010M12 2021M12
Included observations: 131

Joint test:

Chi-sq	df	Prob.
157.4357	160	0.5425

Individual components:

Dependent	R-squared	F(16,114)	Prob.	Chi-sq(16)	Prob.
res1*res1	0.085370	0.665034	0.8226	11.18346	0.7980
res2*res2	0.077273	0.596677	0.8808	10.12276	0.8601
res3*res3	0.177119	1.533604	0.0998	23.20261	0.1084
res4*res4	0.122458	0.994267	0.4681	16.04196	0.4500
res2*res1	0.079387	0.614410	0.8668	10.39971	0.8449
res3*res1	0.127012	1.036621	0.4245	16.63853	0.4094
res3*res2	0.116890	0.943077	0.5231	15.31258	0.5019
res4*res1	0.102315	0.812080	0.6695	13.40322	0.6431
res4*res2	0.090352	0.707699	0.7812	11.83610	0.7552
res4*res3	0.128057	1.046410	0.4147	16.77553	0.4003

Correlation analysis

	LNTOTAL_S...	LNBUNDLE...	LNNON_B...	LNPRICE
LNTOT...	1	0.77631151...	-0.1790265...	0.00744646...
LNBU...	0.77631151...	1	-0.0610953...	-0.0115288...
LNNO...	-0.1790265...	-0.0610953...	1	-0.0456136...
LNPRICE	0.00744646...	-0.0115288...	-0.0456136...	1

Cointegration test

Date: 06/05/23 Time: 10:39
Sample (adjusted): 2011M03 2021M12
Included observations: 130 after adjustments
Trend assumption: Linear deterministic trend
Series: LNTOTAL_SALES LNBUNDLED_SALES LNNON_BUNDLED_SALES L...
Lags interval (in first differences): 1 to 2

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.265949	118.5715	47.85613	0.0000
At most 1 *	0.264653	78.37855	29.79707	0.0000
At most 2 *	0.154410	38.41492	15.49471	0.0000
At most 3 *	0.119952	16.61125	3.841465	0.0000

Trace test indicates 4 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.265949	40.19294	27.58434	0.0007
At most 1 *	0.264653	39.96363	21.13162	0.0000
At most 2 *	0.154410	21.80367	14.26460	0.0027
At most 3 *	0.119952	16.61125	3.841465	0.0000

Max-eigenvalue test indicates 4 cointegrating eqn(s) at the 0.05 level
* denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'*S11*b=I):

LNTOTAL_SA...	LNBUNDLED...	LNNON_BU...	LNPRICE
-5.036852	3.510711	0.697920	-1.780374
-0.717787	-0.734390	0.066011	-5.698622
1.591317	-0.359980	-1.188964	-0.781728
-3.290963	1.866985	1.996138	-1.985553

Unrestricted Adjustment Coefficients (alpha):

D(LNTOTAL_...	0.137384	0.200772	-0.505675	-0.343672
D(LNBUNDL...	-0.109025	0.188615	-0.588904	-0.393562
D(LNNON_B...	0.287914	0.281778	-0.100175	-0.355452
D(LNPRICE)	-0.052498	0.113905	0.145101	0.109368

1 Cointegrating Equation(s):		Log likelihood	-502.9718

Normalized cointegrating coefficients (standard error in parentheses)

LNTOTAL_SA...	LNBUNDLED...	LNNON_BU...	LNPRICE
1.000000	-0.697005	-0.138563	0.353470
	(0.04214)	(0.04472)	(0.16931)

Adjustment coefficients (standard error in parentheses)

D(LNTOTAL_...	-0.691981
	(0.77556)
D(LNBUNDL...	0.549141
	(0.88498)
D(LNNON_B...	-1.450179
	(0.59010)
D(LNPRICE)	0.264425
	(0.24876)

2 Cointegrating Equation(s):		Log likelihood	-482.9900

Normalized cointegrating coefficients (standard error in parentheses)

LNTOTAL_SA...	LNBUNDLED...	LNNON_BU...	LNPRICE
1.000000	0.000000	-0.119681	3.427215
		(0.12137)	(0.34221)
0.000000	1.000000	0.027090	4.409933
		(0.17218)	(0.48547)

Adjustment coefficients (standard error in parentheses)

D(LNTOTAL_...	-0.836092	0.334869
	(0.77783)	(0.54834)
D(LNBUNDL...	0.413755	-0.521271
	(0.88962)	(0.62716)
D(LNNON_B...	-1.652436	0.803847
	(0.58152)	(0.40996)
D(LNPRICE)	0.182666	-0.267956
	(0.24564)	(0.17317)

3 Cointegrating Equation(s):		Log likelihood	-472.0882
Normalized cointegrating coefficients (standard error in parenth...			
LNTOTAL_SA...	LNBUNDLED...	LNNON_BU...	LNPRICE
1.000000	0.000000	0.000000	3.989819
			(0.32515)
0.000000	1.000000	0.000000	4.282589
			(0.38643)
0.000000	0.000000	1.000000	4.700857
			(0.75223)
Adjustment coefficients (standard error in parentheses)			
D(LNTOTAL_...	-1.640781	0.516902	0.710365
	(0.77695)	(0.52538)	(0.20117)
D(LNBUNDL...	-0.523377	-0.309277	0.636546
	(0.88697)	(0.59978)	(0.22965)
D(LNNON_B...	-1.811847	0.839908	0.338646
	(0.60735)	(0.41069)	(0.15725)
D(LNPRICE)	0.413568	-0.320190	-0.201641
	(0.24750)	(0.16736)	(0.06408)

Unit root test

LNTotal sales level, intercept

Null Hypothesis: LNTOTAL_SALES has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.594597	0.0002
Test critical values:	1% level	-3.481217	
	5% level	-2.883753	
	10% level	-2.578694	

LNTotal sales level, intercept, linear trend

Null Hypothesis: LNTOTAL_SALES has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.605820	0.0015
Test critical values:	1% level	-4.030157	
	5% level	-3.444756	
	10% level	-3.147221	

LNTotal sales 1st difference, intercept.

Null Hypothesis: D(LNTOTAL_SALES) has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-9.760130	0.0000
Test critical values:	1% level	-3.481623	
	5% level	-2.883930	
	10% level	-2.578788	

LNTotal sales 1st difference, intercept, linear trend

Null Hypothesis: D(LNTOTAL_SALES) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-9.719687	0.0000
Test critical values:	1% level	-4.030729	
	5% level	-3.445030	
	10% level	-3.147382	

LnBundled sales level, intercept

Null Hypothesis: LNBUNDLED_SALES has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.549499	0.0003
Test critical values:	1% level	-3.481217	
	5% level	-2.883753	
	10% level	-2.578694	

LnBundled sales level, intercept linear trend

Null Hypothesis: LNBUNDLED_SALES has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.567024	0.0018
Test critical values:	1% level	-4.030157	
	5% level	-3.444756	
	10% level	-3.147221	

LnBundled sales 1st difference, intercept

Null Hypothesis: D(LNBUNDLED_SALES) has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-13.48215	0.0000
Test critical values:	1% level	-3.481217	
	5% level	-2.883753	
	10% level	-2.578694	

LnBundled sales 1st difference, intercept linear trend

Null Hypothesis: D(LNBUNDLED_SALES) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-13.41796	0.0000
Test critical values:	1% level	-4.030157	
	5% level	-3.444756	
	10% level	-3.147221	

LnNonBundled sales level, intercept

Null Hypothesis: LNNON_BUNDLED_SALES has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.500290	0.0003
Test critical values:	1% level	-3.481217	
	5% level	-2.883753	
	10% level	-2.578694	

LnNonBundled sales level, intercept linear trend

Null Hypothesis: LNNON_BUNDLED_SALES has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.971677	0.0004
Test critical values:	1% level	-4.030157	
	5% level	-3.444756	
	10% level	-3.147221	

LnNonBundled sales 1st diff, intercept

Null Hypothesis: D(LNNON_BUNDLED_SALES) has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.22443	0.0000
Test critical values:	1% level	-3.481623	
	5% level	-2.883930	
	10% level	-2.578788	

LnNonBundled sales 1st diff, intercept linear trend

Null Hypothesis: D(LNNON_BUNDLED_SALES) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.18099	0.0000
Test critical values:	1% level	-4.030729	
	5% level	-3.445030	
	10% level	-3.147382	

LnPrice level, intercept

Null Hypothesis: LNPRICE has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.812487	0.0001
Test critical values:	1% level	-3.481217	
	5% level	-2.883753	
	10% level	-2.578694	

LnPrice level, intercept, linear trend

Null Hypothesis: LNPRICE has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.798486	0.0008
Test critical values:	1% level	-4.030157	
	5% level	-3.444756	
	10% level	-3.147221	

LnPrice 1st difference, intercept

Null Hypothesis: D(LNPRICE) has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-10.02139	0.0000
Test critical values:	1% level	-3.481623	
	5% level	-2.883930	
	10% level	-2.578788	

LnPrice 1st difference, intercept, linear trend

Null Hypothesis: D(LNPRICE) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 2 (Automatic - based on AIC, maxlag=2)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-9.978712	0.0000
Test critical values:	1% level	-4.030729	
	5% level	-3.445030	
	10% level	-3.147382	

Granger-Causality test

v

Pairwise Granger Causality Tests
Date: 06/05/23 Time: 11:43
Sample: 2010M12 2021M12
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
LNBUNDLED_SALES does not Granger Cause LNTOTAL_SALES	131	0.16151	0.8510
LNTOTAL_SALES does not Granger Cause LNBUNDLED_SALES		0.18355	0.8325
LNNON_BUNDLED_SALES does not Granger Cause LNTOTAL_SALES	131	0.20886	0.8118
LNTOTAL_SALES does not Granger Cause LNNON_BUNDLED_SALES		0.05520	0.9463
LNPRICE does not Granger Cause LNTOTAL_SALES	131	2.34833	0.0997
LNTOTAL_SALES does not Granger Cause LNPRICE		2.59774	0.0784
LNNON_BUNDLED_SALES does not Granger Cause LNBUNDLED_SALES	131	0.19764	0.8209
LNBUNDLED_SALES does not Granger Cause LNNON_BUNDLED_SALES		0.16529	0.8478
LNPRICE does not Granger Cause LNBUNDLED_SALES	131	2.75057	0.0677
LNBUNDLED_SALES does not Granger Cause LNPRICE		2.26943	0.1076
LNPRICE does not Granger Cause LNNON_BUNDLED_SALES	131	0.27142	0.7627
LNNON_BUNDLED_SALES does not Granger Cause LNPRICE		0.39076	0.6774

VECM

Vector Error Correction Estimates
Date: 06/05/23 Time: 11:53
Sample (adjusted): 2011M03 2021M12
Included observations: 130 after adjustments
Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:	CointEq1			
LNTOTAL_SALES(-1)	1.000000			
LNBUNDLED_SALES(-1)	-0.697005			
	(0.04214)			
	[-16.5397]			
LNNON_BUNDLED_SA...	-0.138563			
	(0.04472)			
	[-3.09833]			
LNPRICE(-1)	0.353470			
	(0.16931)			
	[2.08765]			
C	-2.853620			
Error Correction:	**D(LNTOTAL...**	**D(LNBUND...**	**D(LNNON_...**	**D(LNPRICE)**
CointEq1	-0.691981	0.549141	-1.450179	0.264425
	(0.77556)	(0.88498)	(0.59010)	(0.24876)
	[-0.89223]	[0.62051]	[-2.45750]	[1.06297]
D(LNTOTAL_SALES(-1))	0.206129	-0.014881	1.444726	-0.332385
	(0.71077)	(0.81105)	(0.54080)	(0.22798)
	[0.29001]	[-0.01835]	[2.67144]	[-1.45797]
D(LNTOTAL_SALES(-2))	0.156452	-0.010085	1.089415	-0.071688
	(0.54916)	(0.62665)	(0.41785)	(0.17614)
	[0.28489]	[-0.01609]	[2.60722]	[-0.40698]
D(LNBUNDLED_SALES...	-0.355507	-0.321860	-0.920554	0.135664
	(0.47221)	(0.53883)	(0.35929)	(0.15146)
	[-0.75286]	[-0.59733]	[-2.56212]	[0.89570]
D(LNBUNDLED_SALES...	-0.128942	-0.038685	-0.660064	-0.002447
	(0.36901)	(0.42107)	(0.28077)	(0.11836)
	[-0.34943]	[-0.09187]	[-2.35090]	[-0.02068]
D(LNNON_BUNDLED_...	-0.183636	-0.069963	-0.983883	0.093721
	(0.21822)	(0.24900)	(0.16604)	(0.06999)
	[-0.84153]	[-0.28097]	[-5.92575]	[1.33901]
D(LNNON_BUNDLED_...	-0.149536	-0.113856	-0.608966	0.040382
	(0.20525)	(0.23421)	(0.15617)	(0.06583)
	[-0.72856]	[-0.48613]	[-3.89942]	[0.61340]
D(LNPRICE(-1))	1.381223	1.369781	0.900375	-1.161563
	(0.54843)	(0.62581)	(0.41729)	(0.17591)
	[2.51851]	[2.18883]	[2.15769]	[-6.60322]
D(LNPRICE(-2))	0.901013	1.043222	0.672887	-0.525773
	(0.52900)	(0.60364)	(0.40250)	(0.16968)
	[1.70324]	[1.72823]	[1.67176]	[-3.09867]
C	0.021529	0.024014	0.034939	-0.001198
	(0.15424)	(0.17601)	(0.11736)	(0.04947)

C	0.021529	0.024014	0.034939	-0.001198
	(0.15424)	(0.17601)	(0.11736)	(0.04947)
	[0.13958]	[0.13644]	[0.29771]	[-0.02421]

Pearson correlation test

Covariance Analysis: Ordinary
Date: 06/10/23 Time: 14:35
Sample: 1/01/2019 7/22/2019
Included observations: 133

Correlation Probability	LNTOTAL_S...	LNBUNDLE...	LNNON_B...	LNPRICE
LNTOTAL_SALES	1.000000			

LNBUNDLED_SA...	0.968190	1.000000		
	0.0000	-----		
LNNON_BUNDLE...	0.747547	0.628171	1.000000	
	0.0000	0.0000	-----	
LNPRICE	-0.820131	-0.803303	-0.520862	1.000000
	0.0000	0.0000	0.0000	-----

Variance Inflation Factors
Date: 06/10/23 Time: 15:21
Sample: 1/01/2019 7/22/2019
Included observations: 133

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
LNBUNDLED_SALES	0.000516	85.67860	3.397613
LNNON_BUNDLED_...	0.000559	79.25357	1.653830
LNPRICE	0.006073	28.95120	2.822716
C	0.111230	190.7046	NA