University of Cape Town

Faculty of Commerce, School of Management Studies

AN EXPLORATIVE STUDY OF LINEAR VS NON-LINEAR HEDONIC PRICING OF USED CARS IN SOUTH AFRICA

By

Stuart Hendry

This thesis was prepared under the supervision of Professor David Priilaid in the School of Management Studies, University of Cape Town, in fulfilment of the requirements for the degree Master of Philosophy in Applied Management.

December 2024

ABSTRACT

This study investigates the pricing dynamics of the top four entry-level used cars in South Africa, namely the Ford Fiesta, Hyundai i10, Toyota Etios, and VW Polo Vivo. Using a hedonic pricing approach to develop four linear and four non-linear regression models the study analyses a comprehensive dataset of 4,386 second-hand vehicles collected from the Cars.co.za website in 2018 and focuses on eight key vehicle attributes, namely age, mileage, region (province of origin), transmission type, fuel type, colour, engine size, and sub-model variant.

The use of hedonic pricing theory enables the deconstruction of overall vehicle price into marginal contributions of the eight attributes, thereby offering a detailed understanding of their marginal, relative value in determining fair market prices. For instance, characteristics such as age and mileage exhibited diminishing returns, highlighting that their impact on price decreases disproportionately as values increase.

The research identifies asymmetrical relationships between vehicle price and its attributes through a comparative analysis of linear and non-linear regression models. The non-linear models, enhanced by the inclusion of dummy variables, outperformed linear models in terms of accuracy, offering a more precise tool for identifying value for money in the second-hand car market. This study integrates strategic marketing, and applied management elements to provide actionable insights for stakeholders across the used-car sector, including consumers, dealers, insurers, and manufacturers. The statistical analysis within an applied management framework is exploratory and the findings equip consumers with knowledge for making more informed purchasing decisions and assists dealers and insurers in refining pricing strategies and product offerings. Manufacturers are also advised to prioritize attributes significantly impacting resale value during the design and production phases.

In light of increasing digitalization and technological disruptions, the study advocates for the development of a consumer-facing digital platform that incorporates hedonic pricing models. This would offer consumers a reliable tool to evaluate and compare used-car prices in an increasingly complex market and promote trust between buyers and sellers.

DEDICATION

Ad Deo Gloria.

This master's study is dedicated to my wife, Clare, whose unwavering support has been the cornerstone of my academic pursuits. Her tireless efforts and incredible sacrifices supporting this journey have been an enduring source of inspiration.

And to my three wonderful children, Chrissie, Robbie, and Joel, whose entrepreneurial spirit constantly fuels my determination to do work that matters.

I would also like to thank my supervisor, Professor David Priilaid, not only for his vision and support, but also for his incredible patience.

Thank you all, for walking this journey with me and for being my greatest sources of strength and inspiration.

Stuart

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

"Strategically, the South African automotive industry is critical
to the economic future of South Africa."

(National Association of
Automobile Manufacturers
of South Africa, 2021)

1.1 Introduction

Etymologically, the term 'hedonics' is derived from the Greek word *hedonikos*, which simply means 'pleasure' (Triplett, 1986). In economics, the word 'hedonic' refers to the pleasure or utility derived from the consumption of a particular good or service (Priilaid & Van Rensburg, 2006). According to Goodman (1997) and later Stapleford (2011), the pioneering work on hedonic price analysis dates to a 1939 article by Andrew Court, who termed his price indexes as 'hedonic' because he saw buyer utility as an important factor in determining price differentials. Two of the biggest contributors to the development of hedonic price modelling theory are Lancaster (1966) and Rosen (1974). Both theorised that goods are made up of utility-affecting characteristics or attributes and any value derived from these goods, by either producer or consumer, is driven by the inherent value of the underlying attributes (Chin & Chau, 2003). Akerlof's (1970) seminal contribution to information asymmetry is a foundational element of this study, which builds upon and extends his Nobel Prize-winning insights.

A substantial body of literature, based primarily on empirical research, has contributed to the modelling of hedonic price functions and their analysis of what drives price and value across many different product categories (Arrondo, Garcia & González, 2018). By using statistical analysis tools such as regression analysis, these characteristics can be empirically determined, and their implicit, marginal, relative prices can be estimated. In such a way, based on the relationship between the observed prices of differentiated goods and their attributes, hedonic pricing may prove a helpful analytic for estimating the overall price of a product by deconstructing the relative price importance of extrinsic and intrinsic price cues (Erdem &

Şentürk, 2009). From this perspective, a product may therefore be viewed not as one concrete entity, but rather as a collection of characteristics, thereby enabling substitute products to be differentiated and compared (Ohta & Griliches, 1976).

Cars are a primary example of goods that can be qualified as a function of their defining characteristics (Andersson, 2005), and because the South African automotive industry is strategically critical to the economic future of South Africa (Naamsa [National Association of Automobile Manufacturers of South Africa], 2021), this study focused on developing and comparing eight hedonic price regression models for the top four used cars at the market-entry level. For the first time, eight independent characteristics were regressed against price, using four linear and then four non-linear regression models, with the outcomes compared.

This chapter is organised as follows: The first section of this chapter has introduced the field of study, and section 1.2 now presents a definition and explanation of the problems investigated. Section 1.3 explains the contribution of this research to literature and, finally, the structure of this thesis is set out in section 1.4.

The automotive industry is the largest contributor to manufacturing in the South African economy, accounting for as much as 30,1% of the country's manufacturing output, contributing 6,4% to South Africa's gross domestic product (GDP) and representing 15% of South Africa's total exports in 2019 (Naamsa, 2021). From car production and assembly to component manufacturing, the industry employs more than 112 250 people directly and over 457 000 indirectly, generating revenues of R500 billion in 2019 ("WeBuyCars may get new owner", 2021). In 2019, South Africa exported a record 387 125 cars worth R148 billion, along with a record R53 billion in automotive components to 151 international markets.

Despite its substantial impact on the South African economy, there appears to be very limited research related to the definition and development of a statistical model that could underwrite the determination of fair value in the pricing of used cars in South Africa.

While the overall price of used cars is easy to ascertain, consumers struggle with several aspects of the sales transaction. In other words, when buyers are considering the purchase of a used car in South Africa, some of the key characteristics that may influence their decision are model, mileage and age of the car. However, the price-conscious consumer has little to go on when it comes to the intrinsic, relative value of each of these characteristics and by implication, their individual contribution to the overall price of the car.

The pricing problem is compounded by the following three problems that are all pervasive in today's used-car market:

- Market inefficiencies as highlighted by Akerlof's (1970) seminal and Nobel prize-winning contribution to information asymmetry. For example, information asymmetry where sellers have greater access to information or data about the quality of the cars, they wish to sell than the potential buyers of those cars (Bauer, Zavolokina & Schwabe, 2019).
- A lack of trust between buyers and sellers (Karloff, 1970)
- Quality uncertainty (Levin, 2001).

A recent consumer study in Germany ranked the automotive market among the top three markets with the lowest levels of trust (Bauer, Zavolokina & Schwabe, 2019). In essence, the lack of trust between buyers and sellers (the first point above) is constantly exacerbated by the other two points (i.e. information asymmetries and quality uncertainty). This, in turn, leads to market inefficiencies (Akerlof, 1970; Bond, 1984; Levin, 2001) where buyers and sellers are forced to adopt costly and often time-consuming strategies to overcome the lack of trust. Sellers, for example, offer guarantees, discounts and add-ons. Buyers may request specialist check-ups or spend time browsing reviews and conducting online price comparisons (Bauer, Zavolokina & Schwabe, 2019).

Market inefficiencies are as much a problem for the buyer as they are for the seller; where buyers are at risk of buying a 'lemon', sellers are forced to sell 'peaches for peanuts' due to their inability to prove the quality of the car (Akerlof, 1970). This mutual predicament is exacerbated by the imperfect, confusing and often overwhelming amount of information available to the market. In the absence of credible research and, by implication, credible pricing methods, sellers of used cars must often resort to crude 'blue-book' valuations or even personal perceptions. Such practice often leads to unfair or inaccurate car pricing (Van Rensburg & Priilaid, 2004).

In a 2019 research paper, Priilaid and Hendry adopted a linear regression model, implicitly presuming a straight-line relationship between price and the relative, implicit value of the underpinning explanatory variables. The linear regression model proved to be very useful in identifying the statistically most relevant explanatory variables, with the greatest impact on price. In so doing, the authors also determined that further analysis of the price/value continuum was required.

Given the pricing anomalies identified by Van Rensburg and Priilaid (2004) in their linear regression model, and as highlighted by Priilaid and Van Rensburg (2006) in their non-linear model, there seems to be a need for a comparative study.

While Priilaid and Hendry (2019) set out a predictive analytics valuation methodology based on a linear regression framework, there is, to date, no research study that compares linear and non-linear regression models to define the marginal contribution or relative value of individual characteristics in the pricing of used cars in South Africa.

Given the context outlined above, this study therefore addressed the following problems:

Problem Statement One: *There is no linear, hedonic pricing technique that provides a research-based, consumer-facing tool for evaluating the price of used cars in South Africa.* In other words, there is no quantitative valuation methodology that is convenient, accurate and useful for consumers trying to assess relative value when making purchasing decisions in the used-car market.

The benefits of conducting a hedonic price analysis by developing and applying a linear regression model are many and diverse. Firstly, for consumers, it would provide much greater insight into what constitutes real value for money, thereby ensuring a much more effective utilisation of time and money when it comes to purchasing a used car.

Secondly, for retailers, it would not only allow for more effective selling practices but may also provide input as to where higher pricing may be appropriately applied. In both instances, the model leads to more transparent, more effective and more trustworthy pricing.

Thirdly, for manufacturers, it provides an indication of the product characteristics that drive the resale price of cars in South Africa and which characteristics should be focused on, not only in the manufacturing, but also in the design stages of car production. This potentially saves manufacturers from producing unnecessary and/or unwanted characteristics or features.

Fourthly, insurance companies would be able to do more effective pricing if they had access to an effective evaluation tool.

Finally, and perhaps most importantly, the model suggests how informationally efficient pricing in the used-car industry and, perhaps, in other retail categories, might reasonably be achieved. By improving price signals and overall market efficiency for such a large sector of the economy, this approach could make a useful contribution to the functioning of the national economy.

Problem Statement Two: *Can any systemic bias as a result of developing a linear regression model in Problem Statement One be adequately addressed by comparison to a non-linear regression model?* Through the use of both linear and non-linear analysis, this study aimed to ascertain the true extent to which the individual variables or characteristics of a used car determine the intrinsic value of said car in a South African context.

In their 2006 paper, studying the impact of wine characteristics on wine prices, Priilaid and Van Rensburg argued that linear valuations had caused price anomalies both at the bottom end and at the top end of the wine quality spectrum. By way of redress, Priilaid and Van Rensburg introduced a non-linear or 'dummy-styled' approach to address the asymmetrical (non-linear) price-to-quality relationship.

For the first time, eight characteristics or price cues for used cars offered for sale on the cars.co.za website in 2018 were collated into one database and statistically interrogated. Assuming congruence with Priilaid and Van Rensburg's study in 2006, the findings of this study demonstrate how, when comparing linear to non-linear pricing techniques, the non-linear version is almost always more effective. In other words, this study addressed the possibility that the implicit impact on price of explanatory variables may be asymmetrical or non-ordinal, and that there may in fact be a non-linear relationship between price and value.

This study intended to demonstrate that by introducing a non-linear regression model or dummy-styled price analytics model that can identify, quantify, rank and ultimately price variables on second-hand cars, more accurate and economically sensible valuations are possible, thereby assisting consumers in better identifying cars that offer value for money.

This study also sets out some key implications for driving longer-term investment decisions by car manufacturers, purchasing decisions by customers and strategic pricing implications for used-car dealers.

Finally, and in the instances where these critical, though hitherto unquantifiable, price effects were legitimately identified, this study sought to enable their quantification, valuation and financial reportage.

1.2 Contribution to Literature

In addressing the problems described above, this study contributes to the literature in three key areas:

1. It provides a unique evaluation of linear versus non-linear pricing methodologies with a view to providing a research-based, consumer-facing statistical model for the pricing of second-hand cars in South Africa.
2. This study is interdisciplinary in nature. Drawing on three key research focus areas, as they apply to the second-hand car industry in South Africa, this study provides an empirical analysis incorporating key elements from the fields of strategic marketing and applied management. Perhaps the most ambitious form of synthesis occurs in interdisciplinary work, especially where two or more disciplines are actually integrated, not merely juxtaposed, and where such integration yields understanding and insight that could not have been achieved solely in either of the parent disciplines. For the purposes of this study, economic theory and statistical inference are integrated with the objective of providing numerical values to quantify economic phenomena. In other

words, it turns theoretical statistical models into useful tools for economic policymaking (Baltagi, 2002). This study also explored several streams of research in marketing and economics, particularly those concerned with price setting from an applied viewpoint (Arrondo, Garcia & González, 2018) and consumer behaviour, which is that component of marketing theory that studies the constant decision-making process of consumers (Valaskova, Kramarova & Bartosova, 2015).

3. This study also explored and built on the notion of a 'value frontier' as introduced by Priilaid and Van Rensburg (2006) and confirmed in Priilaid and Van Rensburg (2007).

This study followed on from an earlier study published by Priilaid and Hendry (2019), setting out an predictive valuation methodology that used eight explanatory characteristics in a linear regression framework to define the relationship between the pricing of four entry-level cars (Ford Fiesta, Hyundai i10, Toyota Etios and Volkswagen [VW] Polo Vivo). Data pertaining to eight contending characteristics or hedonic price cues, namely age, mileage, region (province of origin), transmission type, fuel type, colour, engine size and sub-model variant, was recorded manually from the popular cars.co.za website to explain the price of 4 388 cars.

This thesis presents the key findings of a series of explorative studies focused on understanding price formation in the second-hand car market through the development of four linear and four non-linear hedonic pricing models. These models were derived using Ordinary Least Squares (OLS) regression techniques and applied to a dataset drawn exclusively from the year 2018. By limiting the temporal scope in this way, the study purposefully controls for time-variant macroeconomic influences—such as inflation, fluctuations in household income, interest rate adjustments, and shifts in consumer sentiment—as articulated by Mok, Chan and Cho (1995). This methodological choice enables a more precise analysis of how vehicle-specific attributes contribute to market value, independent of external economic volatility.

The thesis unfolds in a structured narrative, beginning with a theoretical and empirical overview of hedonic pricing theory, followed by a detailed account of the dataset and model construction process. Subsequent chapters systematically compare the performance and interpretive strength of linear and non-linear models, offering insights into the significance, directionality, and diminishing marginal utility of key vehicle attributes. The final chapters draw together these findings to explore their implications for consumer value perception, dealership pricing strategies, and policy interventions. Taken together, the thesis advances a more nuanced understanding of how vehicle characteristics are internalized in price—telling a broader story of how value is constructed, perceived, and econometrically measured in dynamic, attribute-driven markets.

By employing a comprehensive statistical analysis of the data sets, this study systematically derived and compared eight key vehicle characteristics to evaluate their relative contributions to the pricing structure of used cars. The investigation focused on understanding how these attributes influence the overall price dynamics across the data sets for the top four entry-level car models in South Africa, offering a detailed examination of their significance in this market segment.

What makes this study so timeous is the fact that the used-car market is in the process of being disrupted by technology. Accelerated by the Covid-19 pandemic and exacerbated by changes in market structure, the continuing growth of online platforms to make purchasing decisions, driven particularly by the importance of technology to millennials, car dealers are being forced to adopt new technology or lose market share. The practical implications are that both buyers and sellers will increasingly move online to engage in technology-driven car sales.

1.3 The Structure of This Thesis

This thesis is organised as follows:

Chapter One introduced the topic, identified and set out the two problem statements, defined the scope of the research and set out the contributions made by the research.

Following this introductory chapter, Chapter Two provides insight into the context and conditions for framing the research undertaken. It does so by looking at the used-car market from a global, regional and a South African perspective, and provides key insights into some of the latest technology disruptions in the industry, particularly those related to price.

Chapter Three presents the literature review conducted and the theoretical framework applied. Key topics covered include a review of some of the most influential authors in the field, a detailed look at where the hedonic pricing model has been used in the literature, a definition and explanation of hedonic pricing and a focus on the risks associated with multicollinearity and how to control for such risks. This chapter also deals with two important sub-topics, namely the implications of hedonic pricing for product efficiency and value frontiers in hedonic pricing.

Chapter Four defines the data set and the data collection process and sets out all the relevant descriptive statistics. Chapter Five specifies each car model, using linear and non-linear approaches. Chapter Six presents the analysis and findings of the research conducted. Finally, Chapter Seven summarises the overall contribution of the study and the inherent weaknesses or shortcomings in the research and identifies useful opportunities for further research.

In conclusion, the concept of hedonic pricing provides a powerful analytical framework for understanding how the inherent utility and value of a product's characteristics influence pricing dynamics. The foundational work of pioneers such as Court, Lancaster and Rosen established the theoretical underpinnings of hedonic price modelling by emphasising the utility derived from individual product attributes. This study builds on this extensive body of literature by employing predictive analytical methodologies to estimate the marginal and relative contributions of extrinsic and intrinsic characteristics to the overall price of differentiated goods, specifically in the context of the South African automotive market.

Through the development and comparative analysis of eight hedonic price regression models, this research offers novel insights into the pricing mechanisms of entry-level used cars. By examining intrinsic and extrinsic attributes across linear and non-linear models, the findings underscore the complexity and multidimensionality of pricing in the automotive sector. Moreover, this study highlights the strategic importance of the South African automotive industry to the nation's economic future, providing a practical contribution to both theory and practice.

Ultimately, the research illustrates that products, such as cars, are best understood not as monolithic entities, but as aggregations of characteristics whose individual values contribute to the overall price. This perspective not only advances the theoretical discourse on hedonic pricing, but also equips stakeholders with actionable tools for more effective pricing strategies, product differentiation and market positioning.

The next chapter provides a comprehensive contextual framework underpinning the research conducted in this study. It systematically navigates the global dynamics of the used-car market, contextualises these trends within the African continent, and narrows its focus to the South African market, offering a multi-scalar perspective essential for understanding the complexities and unique characteristics of the sector.

CHAPTER 2: CONTEXT – THE USED-CAR MARKET

"At the highest level, we assemble great inventory for customers,
give them a great shopping experience via mobile or desktop … and simplify the online buying process so they can pull the trigger without haggling and paperwork.
The car is delivered to your front door."

Paul Hennessy, CEO of Vroom

This chapter explains the broader context within which the research for this study was conducted. Starting with a global overview of the used-car market, it then moves to an Africa-wide context and finally to the South African market.

2.1 Introduction

The size and scope of the used-car market at global, regional and national levels serve to indicate the magnitude of both the economic impact and the relevance for consumers, retailers and manufacturers in the sector. This, in turn, demonstrates the need for a consumer-facing solution to the problem of effectively pricing second-hand cars, not only in South Africa, but also for the continent and the rest of the world.

Importantly, this chapter highlights emerging trends in the global automobile industry, which call for more transparent, accurate and accessible car pricing mechanisms as well as other important factors that should be taken into account when developing a pricing tool. For example, public transport has failed in South Africa (Thomas, 2016), giving rise to a greater demand for cars. Used-car retailers in emerging economies such as South Africa are taking full advantage of smartphone technology and digitisation to lead growth in online sales (Deloitte, 2021). If sales are growing, it means the consumer base for used cars is growing. At the same time, if more and more car sales are occurring online, the implication is that whatever tool is developed should also be based online.

2.2 The Global Used-car Market

The Covid-19 pandemic had a severe impact on the global car market in 2020. Across America, Europe and Asia, many original equipment manufacturers shut down manufacturing plants, as the economic effect of strict government-led social and economic closures began to be felt. As a result of these shutdowns, Moody's Rating Agency, predicted that the global demand for new passenger vehicles in 2020 would shrink by approximately 14% (roughly 13,5 million passenger vehicles). Moody's downgraded a number of global automotive manufacturers, including Toyota (Aa3), Ford (Ba2) and Hyundai (Baa1). The Volkswagen Group was losing €2bn per week following plant closures and a collapse in demand and, along with other leading car firms such as Daimler, Renault and Volvo, VW, the second-largest manufacturer of passenger vehicles in the world, was put under review by Moody's.

Despite the impact of the Covid-19 pandemic, long-term prospects for both the primary (new) and the secondary (used) global car market remain positive (Mordor Intelligence, 2021). With the overall increase of car ownership worldwide, the demand for new cars is expected to grow at a compound annual growth rate of 12,1% for the period 2020–2025. The two main factors driving the long-term growth of car ownership are rising income levels and the impact of technological development on the e-marketing/online retail space. Rising income levels give rise not only to greater demand for luxury cars, but also to a growing preference for current two-wheeler owners to upgrade to entry-level passenger vehicles (Mordor Intelligence, 2021).

Key factors that may have a negative impact on demand in the future are a reduced demand for vehicles cars because of e-hailing services such as Uber and increasing concerns about the environmental impact of cars. The 'Uber effect' referred to here may be mitigated by the fact that many car rental fleets have been decommissioned due to the impact of technology-based platforms such as Uber and Lyft, resulting in many opportunities for people to buy well-maintained fleet cars entering the market at discounted prices.

2.3 The African Used-car Market

On the African continent, the car market is a high-growth opportunity, with 1,4 billion inhabitants currently responsible for purchasing a paltry 1% of new cars globally (Harper, 2021). According to the 2021 Mordor Intelligence report, sub-Saharan Africa alone has the potential for increasing car sales by a factor of 10 and is expected to offer high returns to international car manufacturers that are prepared to invest in the region. With the establishment of the Association of African Automotive Manufacturers in 2015, the industry is expected to benefit from reforms and policies. For example, any car company that sets up a local assembly plant in one of the KING countries (Kenya, Ivory Coast, Nigeria and Ghana) could qualify for significant exemptions in import and export duties as well as tax holidays from these governments for periods of up to 10 years (Mordor Intelligence, 2021).

The African region's automotive market will also be boosted by an increase in production infrastructure, with countries such as Ghana (West Africa) and Morocco (North Africa) driving growth in the sector. The Ghanaian automotive market is anticipated to grow at a rate of 30% per annum until 2026, while Morocco's integration into the global economy was enabled by the signing of numerous free-trade agreements with the European Union (Deloitte, 2021). Morocco saw Renault enter the local car market in 2015 as part of an investment cluster partnership with the French government. Renault was the only global automaker assembling cars in the country until the arrival of Peugeot in 2019 (Mordor Intelligence, 2021).

2.4 The South African Used-car Market

Of the 1% of global new-car sales in Africa, 85% are new cars being sold in South Africa (Harper, 2021). The three key factors driving private vehicle demand and, by implication, growth of the second-hand car market are shorter car ownership periods, a booming import–export market for second-hand cars and the previously referred to failure of the public transport system in South Africa.

Strategically, the South African automotive sector is critical to the industrial future of South Africa (Naamsa, 2021). South Africa is the only country in Africa with seven international automotive manufacturers. The automotive industry is the largest manufacturing sector in the country, accounting for 30,1% of manufacturing output and contributing to 6,9% of South Africa's GDP (4,4% manufacturing and 2,5% retail). South Africa also has one of the most competitive automobile trading environments in the world (Deloitte, 2021) and, with over

2 000 car models and sub-models being produced locally, South African consumers enjoy one of the biggest choices to market size ratios in the world (Naamsa, 2021).

From automobile production and assembly to component manufacturing, the industry employs more than 112 250 people directly and over 457 000 indirectly, generating revenues of R500 billion in 2019 ("WeBuyCars may get new owner", 2021). In addition, the automotive industry in South Africa is the country's fifth-largest exporting sector and accounts for 15% of South Africa's total exports, where as much as 11% of total production is exported to the rest of Africa. In 2019, a record 387 125 cars worth R148 billion, along with a record R53 billion in automotive components, were exported to 151 international markets (Naamsa, 2021).

Like many other industries, South Africa's automotive industry came to a standstill in March 2020. With the implementation of the national lockdown due to the Covid-19 pandemic, most car manufacturers were forced to shut down. South Africa's share of global car production decreased from 0,69% in 2019 to 0,58% in 2020, with the country's ranking remaining at 22nd in the world. South Africa remained the dominant market on the African continent and accounted for 447 218 cars, or 62,1% of the total African car production of 720 156 cars (Naamsa, 2021). The impact of the global pandemic on the automotive industry was unprecedented. By comparison, the sub-prime mortgage crisis in 2008 only led to an 8% decline in automotive demand over a two-year period.

How does all of this specifically relate to the used-car market in South Africa?

Firstly, according to the TransUnion Vehicle Price Index, while new passenger finance deals decreased by 5% in Quarter 1 of 2021, used passenger vehicle deals increased by 7,4% year on year, with the ratio of pre-owned vehicle sales to new vehicle sales increasing from 1,2:1 to 2,4:1. The primary reason for the increase in used-car sales over new-car sales was the increase in new-car prices that led to a spike in the demand for pre-owned cars. This increase in demand for used cars means that more and more consumers are looking for a pricing tool that will help them negotiate the best possible price for their vehicle purchase.

Secondly, although new-car sales reached a record high in 2019, 2020 saw new-car sales decline by 29% year on year to just 380 206 units. The decline in sales was caused by the Covid-

19 pandemic, a weak economy, a consequent drop in consumer confidence and increased financial pressure on consumers. With foreign markets also in crisis, this decline was exacerbated by the fact that South Africa exports almost two-thirds of its locally produced cars. Because the purchase of a car is the second-biggest purchasing decision most middle-class South Africans will make, after the purchase of a property, this combination of low consumer confidence and limited disposable income made entry-level used cars a very attractive option, as many people are forced to find cheaper purchases.

In a proactive plan to boost the ailing automotive sector, the Naamsa has worked with government and local manufacturing stakeholders to develop the "South African Automotive Masterplan". Launched in July 2021, the Masterplan aims to boost industry growth by expanding annual production from an average of 525 227 vehicles per annum to 1,4 million vehicles per annum by 2035. The Masterplan also aims to increase the percentage of automobile components manufactured locally from 39% to 60%.

The Masterplan has received a serious boost from three key stakeholders, namely Ford South Africa, Toyota and Nissan, who have each committed to multibillion-rand local investments. These investments mean the design and development of world-class auto engineering facilities that will boost local production as well as the local supply base. A boost in local production in turn requires local auto engineers to focus on designing and adding only those characteristics that maximise the overall long-term value of their vehicles; for example, should an entry-level car have a petrol or diesel engine – a conundrum solved in Chapter Four of this study.

In 2021, the R15-billion investment by Ford Motor Company South Africa in Gauteng is projected to increase production capacity of the new Ford Ranger from the current level of 168 000 units to over 200 000 units per annum, which will lead to the creation of 1 200 direct new jobs and a further 10 000 jobs across the supply chain ("Ford invests $1 billion (R15.8) to modernize, expand...", 2021).

In addition, Toyota South Africa has pledged over R5,5 billion for the manufacturing of the new Corolla Cross SUV, with production starting in October 2021. This project is anticipated to create a further 1 500 jobs. In 2019, Nissan announced that it is investing R3 billion in its facility in Rosslyn, Pretoria to produce the next-generation Nissan Navara pickup.

In 2021, Naamsa projected a 15% growth in new vehicle sales, aiming for 438 000 units, and anticipated a further 5,7% increase in 2022, targeting 463 000 units. They also forecasted a 16,4% rise in exports, from 271 228 to 315 700 units.

According to Naamsa, new-car sales in October 2024 reached 47 942 units, marking a 5,5% increase compared to the 45 418 vehicles sold in October 2023 (Naamsa, 2024). This growth was primarily driven by a 14,5% rise in passenger car sales, which totalled 34 228 units. Despite the positive trend in domestic sales, vehicle exports experienced a significant downturn. In October 2024, exports fell by 42,6%, decreasing from 40 666 units in October 2023 to 23 342 units. This decline has been attributed to factors such as model changes by major local original equipment manufacturers, stricter emissions regulations in Europe and an influx of lower-cost electric vehicles from China. Overall, these figures indicate a mixed performance in South Africa's automotive sector, with domestic sales showing resilience while export markets face significant challenges.

2.5 The Impact of Technology on Used-car Sales

Used-car retailers in emerging economies are taking advantage of smartphone technology and digitalisation to lead growth in online sales (Deloitte, 2021). If sales are growing, it means the consumer base for used cars is growing. The convergence of the automotive and smartphone industries is being enabled by technological innovations where the key objectives remain the same: to provide a quality experience for customers (McKinsey, 2019).

Traditionally, used-car sales dealerships have been asset-intensive businesses that were heavily laden with debt and entirely dependent on traditional sales channels. Today, used-car dealers can leverage technology to build online platforms to be used for sales and marketing, often without even owning a single vehicle. These platforms provide search functionality, attractive landing pages, access to car specifications and performance reports, as well as instant finance options and doorstep delivery. Several global examples include the following:

- New technology players such as Vroom are combining big data analytics, artificial intelligence (AI) and virtual reality technology to leverage the sheer scale and volume of selling used cars online. In six years, Vroom has become one of the biggest online used-car retailers in the world. This is a clear indication of the direction in which car sales are moving. Their app is powered by their proprietary algorithm that analyses specific data from used cars and meta-data from previous sales to determine price. Sellers upload photos of their car and VIN number onto the app, which then calculates a price quotation. Cars are picked up from the seller and delivered to the buyer, who then has seven days to test-drive the vehicle. In June 2020, Vroom filed to go public at $22 per share, with a valuation of $1,5 billion (CrunchBase News, 2020). The valuation came in above its price range of $18 to $20 per share, but within a month its shares closed more than 117% higher. Subsequently, the share price has slipped back to $19 per share as many competitors have entered the market. Carmax, TrueCar and Ebay Motors are offering competitive service, to name only a few. Venture-backed start-ups such as Carlypso and Beepi are also making inroads into their market share. Vroom generates revenue from the sale of used cars, the sale of spare parts from vehicles that cannot be repaired, the sale of warranties and auto-loan finance. It also runs a huge repair and refurbishing operation, Texas Direct Auto, which it acquired in 2015.
- In 2019, Mercedes Benz acquired PlatON, a Chinese technology start-up specialising in blockchain-based privacy-preserving computation. Founded in 2018, PlatON developed a blockchain platform that utilises advanced cryptographic algorithms, including zero-knowledge proofs and homomorphic encryption, to enable secure and decentralised computing. The platform is built on the Etherium 2.0 network and is capable of monitoring the price of used cars, storing static and dynamic data and then calculating the residual value of vehicles registered on the platform. Initially, the focus was on partnering with the Beijing Mercedes-Benz Sales Service, monitoring the price of used Mercedes-Benz in Beijing, but, in time, the platform was made available to authorised used-car dealerships, vehicle inspection firms and used-car owners across China. In a press release published in May 2019, chief strategy officer for PlatON, Ada Xiao, called for better data collection and monitoring services, stating: "With over 6,5 million used cars in China traded in the first half of 2019 alone, we hope that our collaboration with Beijing Mercedes-Benz will highlight the need for more sophisticated data collection

systems to accurately monitor the value of the vehicles comprising China's substantial used car market" ("Mercedes-Benz leverages on blockchain for secure, tamperproof data tracking powered by PlatON", 2019).

- In 2020, PlatON acquired Ownership Labs, developer of a blockchain-based second-hand car sales platform. The company provides a decentralised data management and off-chain trusted computing middleware system to securely store users' vehicle preference and transaction history data in encrypted personal spaces with distributed storage. This system enables users to explore available car listings more efficiently while maintaining control over their data, aggregating insights from both traditional online car sales platforms (Web2) and emerging blockchain solutions (Web3) in one seamless interface.
- In 2018, Volkswagen, which, as stated earlier, is a major original equipment manufacturer, took the decision to invest in the online used-car sales start-up Heycar in the hope of ramping up global revenues. The global trend where auto manufacturers are increasingly becoming involved in the sale of used cars adds to the overall feasibility of a consumer-facing pricing model.

Given the long-term expected increase of car ownership globally and a compound annual growth rate of 12,1% for the period 2020–2025, Chen and Xie (2017) believe it is essential to develop an accurate, statistically driven used-car price evaluation mechanism for the healthy development of the used-car market.

Technology has also placed more power in the hands of the discerning consumer. In their report titled "Used cars, new platform: Accelerating sales in a digitally disrupted market", McKinsey authors Ellencweig et al. (2019) point out that the expectations of customers are increasingly being driven by the high levels of online service that is provided by the big online retailers. Worldwide, tech-savvy customers in the age group 20–40 are still primarily in search of value for money. The Technavio report (2020) defines 'value' as including the following:

- Turnkey (end-to-end) transactions
- Extensive vehicle data and high-resolution photos
- Good-quality websites with an enhanced search function
- AI-assisted decision making based on driver behaviour and customer lifestyle

- Digitised test drives
- Door-to-door delivery.

Customers purchasing a used car online typically spend 40% more time doing their research than customers buying a new car (Ellencweig et al., 2019). Although they still want to be able to inspect and test-drive the vehicle, they also want convenient access to pertinent information such as photographs of the vehicle, the service history and the accident history on readily available, easy-to-use websites.

In South Africa, the WeBuyCars online platform was founded in 2001 by brothers Faan and Dirk van der Walt (WeBuyCars, 2021). Together, they built the company into the country's largest second-hand car retailer. From humble beginnings and start-up capital of R150 000, the company now employs over 1 200 people, sells an average of 8 000 cars per month and generates revenue of R3,7 billion per annum. Over the last 20 years, the company's revenue has grown an average of 60% per annum and CEO Faan van der Walt still sees plenty of room for growth. This is a sentiment shared by their biggest investor, Transaction Capital, which increased its stake in the business to 74% (*BusinessTech*, 2021).

Despite employing over 120 buyers and more than 200 drivers in 16 offices and warehouses located across the Western Cape, KwaZulu-Natal, the Eastern Cape and Gauteng, until November 2018, the business was still being run off Excel spreadsheets, with only the buying process being automated. With the increasing number of monthly transactions, digitisation became imperative. In order to fund the digital transformation process and expand market share, the founders sold 40% of the business to an investment company that could provide both the funding and the technical expertise required. The infusion of capital funded the development of a bespoke inventory management system, followed by a dynamic database that could provide real-time statistics on the buying and selling of cars, expedite decision making and performance review, and assist managers with the effective pricing of second-hand cars.

2.6 The Future of Car Retailing

According to Dr Martyn Davies, managing director of Emerging Markets & Africa at Deloitte, the future of automotive retailing means evolving away from the traditional sales model towards a digitally led model (Deloitte, 2021). A number of factors support this contention:

- *The impact of the Covid-19 pandemic*. The 2021 report from Deloitte on the state of the South African automotive retail sector highlights just how the transition to a new digital realm has been accelerated by the Covid-19 pandemic, when companies were forced to learn how to do business online (Deloitte, 2021).
- *Global trends*. In order to keep up with new global trends that increasingly focus on consumer demands for convenience and value, South African automotive companies are having to become more agile, more digitised and quicker to adopt new technology. An example of this kind of technology is ACES, or autonomy, connectivity, electric and shared mobility, which is reshaping the automotive landscape (Deloitte, 2021).
- *Rapid advances in technology*. Eighteen years ago, Priilaid and Van Rensburg (2006) had already warned that customers will soon be equipped with more than enough consumer-facing information to make more efficient purchasing decisions. This implied that the days of arbitrarily nominating what set of hedonic price cues could best explain price would soon be over (Thrane, 2004) and that the focus should shift to seeking out what characteristics can really serve customer needs as a meaningful guide to car prices (Melichar et al., 2018).
- *The benefits of digitisation for consumers*. These benefits include increased convenience through real-time engagement, fewer paper-based applications and the ability to make purchases from the comfort of one's own couch. To remain competitive, automotive retailers have started introducing end-to-end solutions that assist customers in finding, selecting and financing a used car, as well as after-sales service and insurance. Effective pricing plays a critical part in these solutions.
- *Alternative vehicle ownership models*. Models such as vehicle subscription and leasing are also gradually disrupting the traditional car market in South Africa and are likely to play a bigger role in the future of automotive retail (Deloitte, 2021). Car subscription

models, for example, are typically designed to appeal to individuals who do not need a vehicle full time and who want to avoid the costs, expenses and risks associated with vehicle ownership.

- *Rapidly changing consumer demands*. These demands are forcing both traditional role players in the motor industry, such as retailers and vehicle financing institutions, and new digital players, such as Vroom and WeBuyCars, to recognise this changing landscape. This has important implications for car pricing models. The models need to be transparent and trustworthy, able to leverage the large and diverse array of applicable technology, such as AI and virtual reality, and well-researched and based on sound predictive analytical principles.

In conclusion, the used-car market, both globally and locally, is undergoing a transformative shift, driven by economic, technological and consumer trends. This chapter demonstrated the immense scale and strategic importance of this market, particularly in emerging economies such as South Africa, where the automotive industry serves as a cornerstone of economic development. The convergence of global automotive trends, including the rise of digital platforms, technological innovations and shifts in consumer behavior, underscores the need for more robust and accessible pricing mechanisms tailored to the needs of modern consumers.

Globally, the Covid-19 pandemic accelerated digital transformation in car retail, reshaping how consumers engage with automotive markets and forcing traditional retailers to adapt to the demands of convenience, transparency and digitisation. Locally, South Africa's automotive sector, despite challenges such as public transport failures and economic downturns, remains resilient and central to the continent's automotive growth. This is evident in the increasing reliance on digital platforms for used-car sales, the strategic investments by global manufacturers and the growth of local initiatives such as the South African Automotive Masterplan.

Technological advancements, including blockchain, AI and virtual reality, are redefining the used-car market by enabling innovative business models and empowering consumers with data-driven decision-making tools. These trends highlight the necessity for pricing models that are not only transparent and trustworthy, but also capable of leveraging advanced market research techniques to address the complexities of vehicle valuation in diverse market contexts.

Looking forward, the future of car retailing will be shaped by further digital integration, evolving ownership models and a deeper understanding of consumer preferences. For the used-car market, the development of accurate, data-driven pricing tools will be critical to navigating this dynamic landscape. Such tools will not only enhance consumer confidence, but also strengthen the competitiveness of market players, fostering sustainable growth in the automotive sector both in South Africa and globally.

The next chapter synthesises the foundational literature on hedonic pricing theory, tracing its development through the seminal works of Court (1939), Griliches (1961), Lancaster (1966), and Rosen (1974) (Sirmans, Macpherson & Zietz, 2005). The chapter also explores the application of hedonic pricing to evaluate the underlying characteristics that define the value of used cars, addressing critical methodological challenges such as multicollinearity. In addition, the next chapter highlights some of the practical applications in the literature of hedonic pricing in understanding product efficiency and value frontiers. It demonstrates that pricing is a multifaceted decision-making process requiring integration across management, marketing and statistical disciplines. Greenstone (2017) emphasises that pricing reflects the equilibrium between utility-maximising buyers and profit-maximising sellers. Hedonic pricing, as defined by Mok, Chan and Cho (1995), quantifies the relationship between price and the multitude of product attributes, aggregating the relative value of each characteristic into a unified framework (Priilaid, 2016). This approach enables the fair valuation of products by accounting for the marginal contributions of their constituent attributes (Sirmans, Macpherson & Zietz, 2005). Ultimately, price is a function of both tangible and intangible characteristics, as well as external market factors, underscoring its complexity. For the purposes of this study, the comparative pricing of vehicles is shown to be contingent upon the inherent value of their respective attributes, as supported by Monson (2009), further validating the relevance of hedonic pricing in evaluating the second-hand car market.

CHAPTER 3: LITERATURE REVIEW AND THEORETICAL FRAMEWORK

"A complex, critical decision in marketing and economics is pricing"

(Arrondo, Garcia & González, 2018)

3.1 Introduction

This chapter provides an overview of the literature relating to hedonic pricing theory. It includes sections on key contributors, namely Court (1939), Griliches (1961), Lancaster (1966) and Rosen (1974), as well as a section on the application of hedonic pricing in the literature, with a focus on the characteristics that define the underlying value of used cars.

The chapter also discusses the issue of multicollinearity, which is one of the biggest risks in hedonic pricing, and finally ends with a brief discussion of two important applications of hedonic pricing theory that have been highlighted in the literature review, namely product efficiency and value frontiers.

Pricing is a complex, critical decision-making process that requires interdisciplinary focus from an applied management, marketing and statistical perspective. Greenstone (2017) shows that price represents the value point where the utility-maximising buyer meets the profit-maximising seller.

Hedonic pricing specifically is defined by Mok, Chan and Cho (1995) as a quantitative method that describes the relationship between price and the many factors relating to the product in question. This means that hedonic pricing is essentially an aggregator of the relative value of each underpinning product characteristic (Priilaid, 2016) and forms the foundation of a quantitative model, which can empower consumers to determine the fair value of a product based on the marginal contribution of its constituent characteristics (Sirmans, Macpherson & Zietz, 2005). In other words, price is the sum of all the marginal, relative, weighted-average values of each characteristic or price cue (Chin & Chau, 2003). By implication, therefore, market price is a function of tangible and intangible attributes and additional external factors.

For the purposes of this study, the price of one car for sale, relative to another, differed based on the inherent value of their respective characteristics, as supported by Monson (2009).

3.2 Literature on Hedonic Pricing

A substantial body of literature, based largely on empirical research, has contributed to the modelling of hedonic price functions and their impact on price and value across many different product categories (Arrondo, Garcia & González, 2018).

3.2.1 Court (1939)

Haas (1922) seems to have been the first to use a price regression method, in his paper titled "Sale prices as a basis for farmland appraisal", where he correlated the sale prices of 160 farms in Blue Earth County, Minnesota, to the value of buildings per acre, type of land, crop yields and distance from market. However, Court (1939) is generally considered to be the pioneer of hedonic pricing (Moresino, 2019).

Most of Court's pioneering work was based on hedonic price analysis of examples from the automotive industry. In his 1939 paper titled "Hedonic price indexes: The dynamics of automobile demand", prices per vehicle were divided by the hedonic index to adjust for changes in vehicle specifications. It was, in fact, Court who coined the term 'hedonic' when he described the weighting of the relative importance of various automobile components such as engine size, weight and braking capacity to construct a 'usefulness and desirability index' (Goodman, 1997).

Andrew Court was an economist for the Automobile Manufacturers' Association in Detroit from 1930 to 1940 and worked for General Motors until his retirement in 1966. His research suggests that from as early as 1935, he was already interested in automobile price indices, statistical modelling of implicit prices and non-linearity. It is clear, therefore, that hedonic pricing theory evolved from the car industry. Court realised that passenger vehicles can serve so many diverse purposes that no single specification could be the sole determinant of price and, having examined several other pricing models to no avail, he decided to combine several specifications to form a single composite measure (Court, 1939).

Following on from Court were three of the biggest contributors to the development of hedonic price modelling theory, namely Griliches (1961), Lancaster (1966) and Rosen (1974). A detailed review of their contributions is set out below. While their models have some fundamental differences, both Lancaster and Rosen theorised that goods are made up of utility-affecting characteristics or attributes, valued by both producer and consumer (Chin & Chau, 2003). Automobiles are a primary example of such goods, which can be qualified as a function of their defining characteristics. These characteristics may include age, mileage and model (Andersson, 2005). These characteristics may be empirically measured and, based on the relationship between the observed prices of differentiated goods, the relative, implicit price of individual characteristics can be estimated.

3.2.2 Griliches (1961)

From 1939 to 1960, there was little follow-up on Court's hedonic work, until the empirical economist Zvi Griliches revisited Court's analysis and applied his model in a 1958 article on the demand for fertiliser (Goodman, 1997). While existing fertiliser price indices used equal weightings in the construction of their prices, Griliches used a hedonic framework to better explain demand for fertiliser by relating the prices of different fertilisers to their specific underlying chemical components of nitrogen, phosphoric acid and potash (Griliches, 1958).

In his 1961 paper titled "Hedonic price indexes for automobiles: An econometric of quality change", Griliches continued developing and extending the hedonic framework; this time in the automobile sector. His underlying hypothesis was that price indices consistently failed to account for the full effect of changes in the quality of the underlying variables that make up that price. Unlike Court, Griliches' work was well received, and hedonic pricing models quickly became an important part of the predictive analytical toolbox (Heckman, 2006).

3.2.3 Lancaster (1966)

Traditional price theory assumed that the utility derived from a good stemmed from the innate value of the good itself. In his seminal study titled "A new approach to consumer theory", Lancaster (1966) broke with this traditional theory by arguing that goods are made up of properties or characteristics and that consumers purchase goods because of the utility derived from these characteristics. In other words, the total benefit of a product for a consumer depends on the aggregate of product characteristics (Ladd & Suvannunt, 1976). Utility or value, therefore, is derived from the bundle of characteristics making up the good and not the good itself. Lancaster envisioned products as bundles of attributes on a quality–price continuum (Arrondo, Garcia & González, 2018).

Lancaster's theory of the demand for characteristics plays a vital role in the development of the hedonic pricing method and hedonic demand analysis (Fried & Tauer, 2015; Maurer, Pitzer & Sebastian, 2004). Importantly, Lancaster also made two critical assumptions:

- He assumed that there is a linear relationship between the price of goods and the characteristics contained in those goods. In other words, implicit prices are constant relative to their underlying characteristics. Price can only change when there is a change in the relative combination of the underlying characteristics of the goods consumed.
- He further assumed that the characteristics of a good have the same value for all consumers. In other words, because the value of the characteristics is empirically determined, they are inherently objective. This assumption is crucial in enabling hedonic price modelling (Goodman, 1997).

3.2.4 Rosen (1974)

In his seminal thesis on hedonic pricing titled "Hedonic prices and implicit markets: Product differentiation in pure competition", Rosen (1974) demonstrated that there are competitive markets where implicit prices for characteristics embodied in products can be defined and evaluated for customers making purchasing decisions (Orrego, Defrancesco & Gennari, 2012). Hedonic pricing is the thesis that the observed price of any product is the sum of the unobserved prices of the attributes associated with it (Chen & Rothschild, 2010). Rosen demonstrated how the market matches buyers and sellers of multidimensional goods, where goods are typically priced as a function of the relative, perceived value of their inherent utility-bearing characteristics (Greenstone, 2017). As most goods are heterogeneous, i.e. made up of multiple characteristics, Rosen's model has proven to be applicable across a broad range of economic sectors. His 1974 paper is the sixth most-cited in the history of the *Journal of Political Economy* (Greenstone, 2017).

According to the hedonic price theory described by Rosen (1974), each good can be described using any number (n) of objectively measurable characteristics oriented in a function of x, where:

$$x = (x1, x2, \ldots, xn).$$

The overall value of a good is price (p), where price in terms of its characteristics can be expressed as:

$$p(x) = p(x1, x2, \ldots, xn).$$

Importantly, and in line with Lancaster (1966), where the customer's valuation of a bundle of characteristics is completely subjective, it is assumed that the characteristics themselves are objectively determined, i.e. customers' perceptions of the embodied amount are identical. In other words, price analysis in this context can be viewed as a micro-statistical analysis (Goodman, 1997), where the objective is to understand not only the relative value of the individual characteristics, but also the impact they have on the overall price.

Building on Rosen's theory, Greenstone (2017) contributed further to hedonic pricing theory in two important ways:

- His predictive analytical approach showed the connection between the market value of a good and its characteristics by formulating a hedonic price function that determined the intrinsic values of a key set of characteristics.
- His analysis proved that the optimum hedonic price function is a result of buyers and sellers aiming to maximise the preferred outcomes for themselves in a competitive market, where the price that a utility-maximising buyer is willing to pay for a good and its characteristics should be perfectly matched by a profit-maximising seller (Greenstone, 2017). For example, the price of a car offered for sale should represent the perfect inflection point where the buyer's willingness to pay for the car meets the seller's willingness to sell that car. By implication, therefore, the price of one car relative to another will differ based on the inherent value of the different attributes of that car (Monson, 2009).

Where Rosen's (1974) model for pure, competition-differentiated products still forms the principal theoretical foundation for all hedonic price studies, the purpose of the current study was to build on that foundation by developing and analysing a comparison between linear and non-linear regression methodologies.

3.3 Application of Hedonic Pricing in the Literature

Hedonic pricing models have been used extensively in research related to the food industry, such as breakfast cereals (Morgan et al., 1979; Shi & Price, 1998), coffee prices (Maietta, 2003) and even frankfurters (Harris, 1997). They have also been used for research into consumer products such as personal computers (Berndt & Rappaport, 2001; Pakes, 2003) and mobile phones (Karato, Movshuk & Shimizu, 2015).

This study relied heavily on two of the most important papers in the literature pertaining to the study of hedonic pricing in the South African context, where Van Rensburg and Priilaid, (2004) and Priilaid and Van Rensburg (2006) studied the application of hedonic pricing in the South African wine industry.

In the original application of hedonic wine analysis, Oczkowski (1994) used a conventional multiple regression framework to show how the price of a particular wine was made up of the summation of each of the marginal, relative and implicit prices attached to certain wine attributes. In so doing, he demonstrated that the price of one bottle of wine relative to another would vary with the additional unit of the different characteristics inherent in one bottle relative to any other. A decade later, following on from Oczkowski (1994), Van Rensburg and Priilaid (2004) used a linear regression model to map out the relationship between wine price and value using sighted and blind 'star-styled' quality ratings as explanatory variables.

Where their 2004 paper assumed a straight-line (linear) relationship between quality rating and value, Priilaid & Van Rensburg (2006) subsequently identified price anomalies at both the bottom end and the top end of the wine quality spectrum. In this second study, they introduced a non-linear or dummy-styled approach to address the asymmetrical (non-linear) price-to-quality relationship.

Hedonic pricing models have probably been most extensively used in the literature on housing (Arrondo, Garcia & González, 2018). There are studies that provide estimates of the pricing of real estate (Edmonds, 1984; Hsueh, 2000; McMillen & Redfearn, 2010; Monson, 2009). This has extended to estimates of consumer valuation of environmental externalities such as air pollution (Palmquist, 1984) as well as public goods, such as school quality and neighbourhood amenities (Bayer, Fernando & McMillan, 2007; Black, 1999). In the pricing of real estate, studies have found, for example, a significant, negative relationship between air pollution and property values (Beron et al., 2004; Greenstone, 2017, Lee, Park & Kim, 2003). The model has also been used in determining the extent to which proximity to the Brussels city centre, urban parks and the tube station were significant determinants of the implicit value of houses in the city (Melichar et al., 2018).

An extensive body of research has focused on hedonic pricing in the context of the hotel industry (Arora & Mathur, 2020), where studies have linked hotel room prices to a number of different attributes. These include paying marine conservation fees (Lorde, Jadon & Weekes, 2019; seasonal attributes (Chen & Rothschild, 2010); site-specific factors such as age, capacity, hotel and room amenities (Abrate, Capriello & Fraquelli, 2011; Hung, Shang & Wang, 2010; Torres, Adler & Behnke, 2014); and locational aspects such as proximity to airport, tourist attractions and city centre (Espinet et al., 2003; Kim et al., 2020; Thrane, 2007; White & Mulligan, 2002).

Hedonic pricing models are also widely found in the literature pertaining to the pricing of second-hand cars, automobile demand and hedonic price indices. Part of the reason for this is that due to the size of the used-car market, none of the players are big enough to affect price, and prices for used cars can be considered an untainted reflection of the value placed on them by the buyer. This, in turn, is a strong indication that the used-car market is very close to pure competition, which makes it ideal for empirical investigations (Betts & Taran, 2006).

Atkinson and Halvorsen (1990) were the first to use hedonic pricing to estimate the value of car safety using data from the American market for automobiles. They were followed by Dreyfus and Viscusi (1995), Mount et al. (2001) and Andersson (2005). Pazarlioglu and Gunes (2000) used the hedonic pricing method to analyse the factors affecting new-car prices in Turkey. According to the results of their study, the determinants of car prices are mainly the automatic gearbox system, engine capacity, engine power and fuel consumption. In their study, Erdem and Şentürk (2009), used the hedonic regression technique to determine factors affecting used-car prices in Turkey.

Some of the most useful examples of empirical studies in the automobile market include the predictive analytics of quality change and automobile price indices (Griliches, 1961), hedonic quality measurement in automobiles (Triplett, 1986), the non-linear effects of age and reliability on price (Betts and Taran, 2006), hedonic prices for cars in the Spanish car market (Matas, Raymond & Roig, 2009), a new hedonic technique for estimating attribute demand (Atkinson & Halvorsen, 1984), survey evidence of the willingness of US consumers to pay for automotive fuel economy (Greene, Evans & Hiestand, 2013), the non-linear effect of reliability on used-car prices (Prieto, Caemmerer & Baltas, 2015) and policy implications for the green car market (Chowdhury, Salam & Tay, 2016; Gonzalez, Arrondo & Carcaba, 2017; Gössling & Metzler, 2017; Moresino, 2019).

From an extensive review of the literature to date, it is clear that no one has yet conducted a comparative evaluation of linear versus non-linear pricing methodologies, with the specific goal of providing a research-based, consumer-facing, market analysis model that can be applied to the pricing of second-hand cars in South Africa.

3.4 Utility-bearing Product Characteristics

Virtually all goods are heterogenous, or as Rosen (1974) puts it, “goods are a vector of their ‘utility-bearing’ attributes or characteristics”. In terms of the consumer decision-making process, the result of the buying decision will be determined by the attributes of the product and the importance the consumer attaches to each of those attributes (Arrondo, Garcia & González, 2018). Haas (1922) refers to these characteristics as “elements of difference or variations”, Priilaid and Van Rensburg (2006) refer to them as “consumer desirables”, while Allenby, Brazell and Howell (2014) and Gonzalez, Arrondo and Carcaba (2017) refer to “product features”. In statistical terms, these characteristics or price cues can also be referred to as ‘independent variables’ or ‘predictor variables’, because they explain changes in the dependent variable relative to price.

Kotler and Armstrong (2004) describe product attributes as those characteristics that complement the basic function of the product. Court (1939) was visionary in this aspect, theorising that where commodity price indices were traditionally made up of raw materials or semi-fabricated products and pricing problems were mainly related to weight and the correct formula to be used, 20th-century technology would lead the evolution of a new group of products. These products would be designed for complex functioning, made from many separate parts and subject to rapid improvement in both design and construction. Court (1939) predicted that automobiles would make up one of the most important groups of new products and that any improvement in measuring changes in their prices would have real benefit to the economics of pricing.

When considering the purchase of a second-hand car, it is assumed that purchasers choose a car with a set of characteristics that will maximise their utility or pleasure, within a given budget (Erdem & Şentürk, 2009). In their study, Betts and Taran (2006) refer to purchasing a car as a "high involvement decision", with consumers actively seeking and using information that could help them make a better decision. A used car is a complex and technologically advanced product, and purchasing one is normally the second-biggest financial decision people must make in their household (Gongqi, Yansong & Qiang, 2011).

To clarify the differences in vehicle characteristics, some researchers distinguish between basic features and additional features (Gongqi, Yansong & Qiang, 2011), physical attributes and performance attributes (Ohta & Griliches, 1976) and internal characteristics and external characteristics (Erdem & Şentürk, 2009). These characteristics may include structural or production-related attributes (e.g. engine size, transmission type, fuel type and colour), attributes related to use (age and mileage) and sub-model variants. Following the perspective of Ohta and Griliches (1976), a vehicle's characteristics can be differentiated based on physical versus performance characteristics, with their research showing that the set of physical characteristics plays a significantly more important role in the evaluation of vehicle value than performance characteristics. In his study, Andersson (2005) chose explanatory variables based on economic grounds, i.e. variables assumed to be "utility-bearing characteristics". Andersson (2005) also reminds us that when deciding which car characteristics to include in the hedonic equation, the researcher has the usual trade-off between including too many variables and risking multicollinearity and not including enough variables and struggling with omitted variable bias.

3.5 The Relationship between Hedonic Pricing and Consumer Behaviour

In today's highly competitive markets, managers are only able to satisfy consumer needs if they understand those needs and the decision-making processes involved in satisfying them (Arrondo, Garcia & González, 2018). To this end, considerable research has been conducted into consumer purchasing decision processes (Kotler & Armstrong., 2004; Shamsher, 2014). Product attributes play an important role in the decision-making process from the viewpoint of both buyers and sellers (Akoyomare et al., 2018). Whereas buyers evaluate product attributes in order to determine the best value to meet their needs, sellers combine product attributes in order to differentiate their products and brands from those of competitors (Shamsher, 2014). At best, buyers are able to estimate the overall value of the product, given its product attributes. This is where the hedonic pricing model is so useful. Estimated implicit prices for characteristics are the most important empirical results from a hedonic function (Triplett, 1969). As the hedonic function is performed in a competitive market, the implicit prices, albeit estimated, are an empirical reflection of the supply and demand for the product characteristics.

One of the key reasons for selecting this research topic was precisely to understand and thereafter address the consumer's need for value by researching and developing a price-based regression model that will allow consumers to make more efficient purchasing decisions at the point of sale (see Van Rensburg & Priilaid, 2004). As most consumers have neither the time nor the money and very rarely the expertise to assess the value of used cars available to them, what then do they use as a meaningful guide to determine whether a vehicle is fairly priced?

Therefore, in the case of purchasing a used car, what is it that consumers can do to mitigate their risk in what is probably the second-biggest purchasing decision of their lives? Consumers (buyers) bear two types of risks in making this decision:

1. The biggest risk arises from the fact that the seller has more knowledge about the quality of the car. As it is almost impossible for a buyer to tell the difference between a good and a bad car, an asymmetry in available information has developed and as a result, bad cars sell at the same price as good cars (Akerlof, 1970). In their paper on the use of hedonic price functions with incomplete information on house prices, Kumbhakar and Parmeter (2010) show that the asymmetry of information between buyer and seller has had a significant impact on house prices in America. They developed a technique that combines the effects of asymmetrical information on both demand and supply side and then estimates the differences between the house price and the buyer's maximum willingness to pay and the seller's minimum willingness to accept.
2. The buyer's second risk is related to the transaction itself: Is the seller able to perform (Akerlof, 1970; Mishra, Heide & Cort, 1998)? The seller's primary risk is the possibility of non-payment. The buyer's risk exceeds that of the seller, with the potential cost of information asymmetry being much higher. It is clear, therefore, that any price-based regression model that will allow a buyer to make a more efficient purchasing decision at the point of sale should also mitigate such buyer's risk (Khumbakar, 2010). An innovative example of a commercial solution for mitigating buyer risk is Lightstone Auto (2021). By scanning and decrypting the barcode on a car license disc and a driver's licence card, the Lightstone app can verify both car and driver instantly, returning accurate, verification information in real time. Features include instant electronic processing of contact information, storing copies of driver's licences for test drives and pre-population of finance applications, a basic affordability test on a potential customer, without performing a credit enquiry, and providing trade and retail values on current year models, including a comprehensive list of cars available in the market.

As consumers turn increasingly towards digital platforms such as Lightstone to assist them with their decision making, the need for a consumer-facing, transparent and trustworthy pricing method can only grow. As Rosen (1974) reminds us, the hedonic pricing process of modelling how buyers and sellers optimise their behaviour drives a data-generating process that delivers the potentially observable equilibrium between characteristics and their relative prices (Greenstone, 2017). In the age of digitisation, which is primarily data-driven, the many practical applications of the hedonic price method could add substantial value to both buyer and seller as they struggle to reach an equilibrium price in a competitive market (Huang & Lin, 2007).

3.6 Multicollinearity

This section provides an overview of the concept of multicollinearity, highlighting its implications for statistical analysis and its treatment in prior research. Chapter Five will delve further into methods for correcting this issue, offering a framework for enhancing the robustness of regression models. By addressing multicollinearity, this study seeks to contribute to more accurate and reliable statistical modelling practices.

3.6.1 Introduction

Given the critical role that independent, explanatory characteristics play in regression models and given that the basic premise behind an independent characteristic is that it must be independent, intercorrelations between explanatory characteristics can lead to statistical difficulties (Triplett, 1969). To mitigate against this type of risk, this section briefly defines the concept of multicollinearity and explains why it is a problem and how it has been dealt with in previous studies. How to correct for this issue is dealt with in Chapter Five.

3.6.2 What is Multicollinearity?

A regression coefficient by definition represents the average change in a dependent variable for every unit of change in an independent variable when all the other independent variables are held constant (Moresino, 2019). It is only possible to obtain statistically accurate results if the model is able to regress one independent variable at a time. If variables are correlated, then changes in one variable are associated with changes in other variables, thereby rendering the model inaccurate. The stronger the correlation, the more difficult it is to change one variable without it having an impact on another. Research also shows that large numbers of economic variables from a single sample are almost certain to be highly intercorrelated (Farrar & Glauber, 1967).

3.6.3 Why is Multicollinearity a Problem?

In most cases, the question is not whether or not there is multicollinearity, but rather the extent to which the problem exists. As interdependence grows among explanatory variables, so the correlation matrix approaches singularity and the data becomes less and less useful, i.e. observed data contain less and less useful information and statistical data becomes less valid in attempting to define the empirical relationship between price and price cue (Priilaid & Van Rensburg, 2006).

Multicollinearity can lead to all sorts of statistical difficulties (Triplett, 1969). For example, in a Swiss study where three vehicle quality labels (independent variables) were regressed against the price (dependent variable) of used cars in Switzerland, it quickly became clear that a high level of correlation existed between two of the three labels (Moresino, 2019). This meant that only one out of the three research questions could be answered. Another risk with multicollinearity is that researchers tend to oversimplify their models in the hope of avoiding high intercorrelations. Oversimplification can lead to overidentification, which in turn means misleading structural estimates (Triplett, 1969). One of the most serious problems is that multicollinearity among so-called independent variables can result in estimates with theoretically incorrect signs (Atkinson and Halvorson, 1984).

Successful model building therefore requires correct specification of the most relevant variables and thereafter the selection of the correct estimating procedure (Farrar & Glauber, 1967), in the case of this study, the OLS model. Multicollinearity can contribute to difficulty in the specification as well as the estimation of the economic relationship between dependent and independent variables.

3.6.4 How to Control for Multicollinearity

The first step in the proper treatment of the multicollinearity problem is detection or diagnosis (Farrar & Glauber, 1967). Detection requires a clear understanding of where and why the variables are correlated as well as the extent to which there is multicollinearity. Economists are becoming more willing to agree that the second step, correction, requires the generation of additional information (Meyer & Glauber, 1964). However, any new information can only be collected and applied effectively once there is a clear understanding of where and how the pattern of interdependence that undermines present data occurs. Moresino (2019) considered the radical approach of discarding one of the correlated variables, but realised quickly that the major drawback would be that he might be excluding variables that affect price. In the end, the author used a principal component regression to control for multicollinearity. This method begins with a principal component analysis to find orthogonal variables capable of modelling variance in an optimal manner. As the variables are orthogonal or independent, there can be no correlation.

3.7 Two Applications of Hedonic Pricing for Future Research

The investigation of product efficiency and value frontiers represents a crucial dimension in understanding pricing dynamics across various industries. This section introduces two significant applications of hedonic pricing theory: product efficiency and the value frontier, both of which have profound implications for manufacturers and consumers alike. By leveraging data envelopment analysis and hedonic pricing models, prior research in industries such as the Spanish automobile and running shoe markets has demonstrated how product attributes and competitive pricing drive efficiency and consumer value. The methodologies employed by Gonzalez, Arrondo and Carcaba (2017) and Arrondo, Garcia and González (2018) provide a robust foundation for assessing the relationship between product characteristics and price, emphasising the role of innovation and competition in optimising market efficiency.

The concept of the value frontier, as explored by Van Rensburg and Priilaid (2004), further enriches this discourse by offering a graphical representation of the quality–price relationship, enabling both consumers and manufacturers to identify optimal pricing strategies. The transition from traditional linear models to advanced non-linear frameworks opens new opportunities for integrating product efficiency theory with value frontier analysis. In doing so, manufacturers can better align product attributes with consumer preferences, optimising pricing strategies for both new and used markets.

This study builds on these foundational theories and methodologies to explore the intrinsic value of individual car characteristics in the South African context. By applying non-linear regression models, it aimed to address systemic biases inherent in linear approaches, thereby advancing the understanding of product efficiency and value determination. Ultimately, the findings have the potential to guide manufacturers and engineers in prioritising attributes that maximise efficiency and consumer satisfaction, while providing a practical framework for developing dynamic and data-driven pricing models.

Application 1: Product Efficiency

In their study on investigating the evolution of product efficiency in the Spanish automobile industry, Gonzalez, Arrondo and Carcaba (2017) collected data on more than 75 technical product features and compared them to prices for a sample of compact cars introduced into the market between 2010 and 2015. By using a data envelopment analysis, where the efficiency rating is derived from the most favourable ratio of the sum of weighted outputs relative to inputs, they derived a 'best-buy frontier'. They then used a Malmquist index to measure the overall increase in product efficiency over the same period and mapped the improvements on their 'efficiency frontier'. In data envelopment analysis, the efficiency frontier represents the lowest possible combination of cost inputs capable of producing the required output (Charnes, Cooper & Rhodes, 1978). Gonzalez, Arrondo and Carcaba (2017) clearly show the extent to which product innovation, improvements in product features and price cutting as a result of competitive behaviour drove overall product efficiency.

In their research on product efficiency in the Spanish running shoe market, Arrondo, Garcia and González (2018) showed empirically that, over time, the competitive market is dynamic enough to reduce price inefficiency. In other words, by using a regression model to value product attributes relative to price, they could prove that the least price-efficient running shoes also showed the biggest reductions in price over the control period. Arrondo, Garcia and González (2018) compiled price data from two different time moments during the year (February and June) on 355 models of running shoes belonging to 31 brands. They used a basic hedonic pricing model to relate the price of the shoe to its bundle of objective attributes, namely weight, cushioning, flexibility, grip and brand. The hedonic price of a running shoe therefore shows the maximum price a consumer would pay for a pair of running shoes and its underlying characteristics.

As product attributes increasingly become more influential in consumer purchasing decisions (Arrondo, Garcia & González, 2018), the valuation of product attributes is a crucial part of the development and commercialisation of new products (Allenby, Brazell & Howell, 2014). The fact that the proposed models used by Gonzalez, Arrondo and Carcaba (2017) and Arrondo, Garcia and González (2018) are useful for manufacturers to dynamically assess product efficiency relative to product attributes and price was of great interest to this study, as it presented a unique opportunity for further research.

If this thesis can solve Problem Statement Two as set out in Chapter One; c*an any systemic bias as a result of developing a linear regression model in Problem One be adequately addressed by comparison to a non-linear regression model,* then the non-linear model developed in this study could have very practical applications for manufacturers of new vehicles. In other words, using non-linear analysis in conjunction with the work of Gonzalez, Arrondo and Carcaba (2017) and Arrondo, Garcia and González (2018), automobile manufacturers committed to greater product efficiency can ascertain the true extent to which the individual characteristics of a used car determine the intrinsic value of said car in a South African context. This means that auto engineers will know where and on which characteristics to focus as they develop new models.

Application 2: The Value Frontier

In an effort to find the best possible quality relative to price (i.e. product efficiency), the informed buyer takes into account all the relevant product attributes before making a purchasing decision (Arrondo, Garcia & González, 2018). The idea of a value frontier was proposed by Van Rensburg & Priilaid (2004) as an empirical construct that reflects these quality–price relationships. Using the results of wine tastings published in two well-known South African wine publications, Van Rensburg & Priilaid assessed the price–quality relationship for three South African-grown cultivars (Cabernet Sauvignon, Merlot and Shiraz) and graphically depicted the empirical results on a value frontier.

The idea of a best-buy frontier was proposed by Gonzalez, Arrondo and Carcaba (2017) in their empirical study reflecting quality–price relationships. By using data envelopment analysis to measure product efficiency for each period, they could compare price and product features across different models of cars. This technique enabled consumers to identify overpriced used cars. Arrondo, Garcia and González (2018) also combined hedonic pricing with frontier analysis in order to estimate product efficiency. Their research went a step further in that they not only identified overpriced products, but were also able to model the extent to which manufacturers would have to discount their prices in order to make their new cars competitive.

Conventional hedonic pricing models demonstrate empirically that, despite a degree of random error (e), price is fundamentally aligned with product attributes. This study aimed to enhance the pricing methodology by evaluating and contrasting linear (ordinal) and non-linear (non-ordinal) models, thereby addressing the second research problem outlined in Chapter One. Resolving this issue provides a clear trajectory for future research. The transition from linear to non-linear pricing methodologies offers an opportunity to integrate product efficiency theory with the conceptual framework of a value frontier. The development of such a value frontier would hold substantial significance for manufacturers and retailers, providing a robust tool for determining optimal pricing strategies for both new and used vehicles in a competitive market landscape.

3.8 Summary and Conclusion

In hedonic price analysis, characteristics are independent, homogeneous, economic variables that are valued by both buyer and seller and, when bundled together, form a heterogeneous good or product (Triplett, 1986). This study applied regression analysis techniques to identify the specific relationship between the dependent variable (price) and each of eight independent variables (price cues) and then to determine their relative impact on the overall price of four entry-level second-hand cars in South Africa.

Where it is impossible to price the eight characteristics separately, the overall price of the used car represents the sum of the marginal, implicit price for each characteristic embodied in the bundle. In other words, once the key characteristics of the car have been identified, the hedonic price function will disaggregate the price of the product into the relative, implicit prices of each of the characteristics. Because these prices can only be estimated as part of an empirical model, rather than directly observed, they are usually termed ‘implicit’ prices (Andersson, 2005).

This literature review has demonstrated how suited the hedonic price regression models are to the pricing of used cars. A used car is a complex and technologically advanced heterogeneous object made up of many homogenous characteristics and often there is a high price differential between cars which from the perspective of a buyer may seem exactly the same (Gongqi, Yansong & Qiang, 2011). In his paper, Stigler (1961) attributed price dispersion to two overarching components: heterogeneity and asymmetrical information between buyer and seller. This study explored a market that is sufficiently 'thick', i.e. made up of many products with many characteristics, thereby eliminating price dispersion for almost identical goods. It is only where markets become 'thin', i.e. made up of products that are increasingly unique, that price dispersion occurs.

This literature review has also demonstrated that consumers are perpetually in search of value for money and often go to great lengths to find 'the perfect deal'. This study therefore aimed to understand and address the consumer's need for value by researching and developing a price-based regression model that will allow a customer to make a more efficient purchasing decision at point of sale. The consumer's need for value coincides with a global shift towards digitisation – a shift that has led to new customer needs evolving, effectively creating a buyers' market (Deloitte, 2021). In other words, in theory, discerning consumers have more power than suppliers. As a consequence, there is a clear understanding in both academia and the popular media that the second-hand car market should be aiming to be more consumer-facing in order to survive the current economic challenges. As most ordinary consumers have neither the time nor the money to assess the plethora of available second-hand cars, options and variables, they are faced with two challenges: (1) identifying which characteristics best reflect value for money and (2) developing the means by which these characteristics can best be evaluated.

In light of the predicament facing consumers, this study presents four predictive analytical valuation models using a regression analysis to map out the linear relationship between car price and value. In the study by Priilaid and Hendry (2019), this approach provided useful insights into the three variables that have the biggest impact on car prices, namely age, mileage and province of origin. There was, however, a risk that valuations, at the top end of the spectrum, tended to overvalue cars relative to the two time-decay factors (mileage and age) (Priilaid & Van Rensburg, 2006). At the bottom end of the price spectrum, linear valuations have also caused problems. These valuation anomalies have led to a view that the linear hedonic pricing model is inadequate (Priilaid & Van Rensburg, 2006), thereby necessitating the development and comparison of four non-linear price regression models. In order to leverage the strengths of both the linear and the non-linear models and to mitigate the weaknesses, in an analysis similar to that of the study by Priilaid and Van Rensburg (2006), this study also compared and presents four statistical valuation models using a regression analysis to map out the non-linear relationship between car price and value. The comparative analysis of four linear and four non-linear regression models represents a novel contribution to the field. This study sought to demonstrate that such a methodological comparison can yield more precise, consumer-centric and economically rational vehicle valuations. By leveraging these insights, the study aimed to enhance the identification of value for money in the used-car market, offering a more robust framework for pricing strategies and consumer decision making.

The next chapter provides a comprehensive overview of the data set employed in this study. By emphasising key factors such as data accuracy, completeness and consistency, the chapter underscores the importance of high-quality data in facilitating precise analyses and sound conclusions. In the next chapter, the data set is systematically detailed and key independent variables and price are carefully outlined. The chapter concludes with a presentation of the descriptive statistics, forming a solid foundation for the analytical work that follows.

CHAPTER 4: DATA COLLECTION AND DESCRIPTION OF DATA SET

"An accurate evaluation of used car prices is a prerequisite for the development of a healthy used car market."

(Chen et al., 2017)

4.1 Introduction

This chapter presents and describes the data set employed in this study. Good-quality data is important in any research study. Factors such as accuracy, completeness and consistency all define good data. Poor-quality data can have negative consequences on research outcomes, including inaccurate analysis and poor conclusions. This chapter starts with a brief description of the data set, followed by a description of the independent variables and price, and culminates with the descriptive statistics.

4.2 Description of the Data Set

The data set for this study was segmented and analysed across each of the four top-selling, second-hand, entry-level vehicles in South Africa: Ford Fiesta (n = 1 008), Hyundai i10 (n = 1 137), Toyota Etios (n = 1 133) and VW Polo Vivo (n = 1 108). A summary of the data set is presented in Table 4.1. To ensure completeness and consistency, data from 4 386 vehicles were manually recorded from the same website, namely cars.co.za. Data collection focused on eight key independent variables, namely age, mileage, (region) province of origin, transmission type, fuel type, colour, engine size and sub-model variants.

Table 4.1: Data Set Summary

Period	2018	
Entry-level vehicles included	Ford Fiesta Hyundai i10 Toyota Etios VW Polo Vivo	(n = 1 008) (n = 1 137) (n = 1 133) (n = 1 108)
Size of final data set		4 386 lines
Treatment of prices	The price of each car is given in South African rand (1US$ = ZAR14,00) and sourced online from the cars.co.za website.	

Note: For the purposes of modelling, car prices do not need to be adjusted for inflation, because the sample is for the year 2018 only (see Priilaid & Hendry, 2019). Limiting this study to one year minimised the effects of such time-variant variables and effectively controlled for them (Mok, Chan & Cho, 1995).

4.3 Description of the Independent Variables (Price Cues)

The eight consumer-facing independent variables (price cues) selected for this study to inform the dependent variable (used car prices) are explained below.

4.3.1 (Model and Year)

Motor vehicle age is calculated in years, from the first year of registration, and is presented in Table 4.2. Categorisation by age is common in durable goods such as cars, where the price that a buyer is willing to pay for a particular used car reflects that car's utility for the buyer. Consumers rightly assume that age affects utility where, for example, older cars require more maintenance and repairs and are often less reliable (Betts & Taran, 2006). This study measured the exact, empirical cost of this decline in utility by reviewing and analysing models from 2008 up to and including 2018. This is in line with other similar studies (Gongqi, Yansong & Qiang, 2011).

Table 4.2: Number of Vehicles Recorded Per Year

Model and year	Number of vehicles recorded per year				TOTAL
	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	
2018	79	9	187	138	413
2017	331	315	199	125	970
2016	303	273	190	125	891
2015	143	137	143	179	602
2014	116	141	104	180	541
2013	63	154	67	124	408

Model and year	Number of vehicles recorded per year				TOTAL
	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	
2012	34	104	42	90	270
2011	39	0	29	90	158
2010	21	0	47	57	125
2009	5	0	0	0	5
2008	3	0	0	0	3
TOTAL	**1 137**	**1 133**	**1 008**	**1 108**	**4 386**

There is a strong tendency for the overall mean (3,00) and median (3,64) age to be approximately three years across all of the data sets, other than the VW Polo Vivo data set, which has a mean (4,50) and median (4,00) of around four years. This may be related to sales promotions that were in place during this time, influenced by length of guarantees and maintenance agreements provided, and so-called balloon payments.

Table 4.3 provides the descriptive statistics of the motor vehicle age for the study sample of 4 386 motor vehicles, as sourced from the website cars.co.za in 2018. The following descriptive statistics are included:

1. Central tendency statistics (mean, median, mode)
2. Distribution statistics in the form of frequency distributions (summary of the frequency of individual variables)
3. Dispersion statistics (standard deviation, variance and range), which refer to the spread of the values around the central tendencies (Trochim, 2006).

Table 4.3: Descriptive Statistics – Age (Years)

Vehicle type	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Observations	1 137	1 133	1 008	1 108	4 386
Minimum	1,00	1,00	1,00	1,00	1,00
Maximum	11,00	7,00	9,00	9,00	11,00
Mean	3,52	3,85	3,57	4,50	3,86
Median	3,00	3,00	3,00	4,00	3,00
Mode	2,00	2,00	2,00	5,00	2,00
Std. deviation	1,93	1,70	2,20	2,32	2,08

There is a strong tendency for the overall mean (3,00) and median (3,64) age to be approximately three years across all of the data sets, other than the VW Polo Vivo data set, which has a mean (4,50) and median (4,00) of around four years. This may be related to sales promotions that were in place during this time, influenced by length of guarantees and maintenance agreements provided, and balloon payments.

4.3.2 Mileage

Ironically, the term 'mileage' is measured in kilometres (km) and is recorded from the odometer reading of each vehicle. The mean of the mileage is comparable among the different data sets. With the overall mean at 55 857 km, the mileage means for the Toyota Etios (61 253 km) and VW Polo Vivo (67 654 km) are higher than that of the Ford Fiesta (49 327 km) and the Hyundai i10 (44 792 km). The overall median of the mileage (45 500 km) is somewhat below the mean for all four data sets, implying that the distribution of the data is skewed to the left, with a small number of outlying data with large mileage.

Table 4.4 provides the descriptive statistics of the mileage for the study sample of 4 386 motor vehicles, as sourced from the website cars.co.za in 2018. The following descriptive statistics are included:

1. Central tendency statistics (mean, median, mode)
2. Distribution statistics in the form of frequency distributions (summary of the frequency of individual variables)
3. Dispersion statistics (standard deviation, variance and range).

Table 4.4: Descriptive Statistics – Mileage (km)

Vehicle type	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Observations	1 137	1 133	1 008	1 108	4 386
Minimum	0	0	1	1	0
Maximum	307 318	227 000	238 000	345 000	345 000
Mean	44 792	61 253	49 327	67 634	55 857
Median	34 000	50 000	44 000	65 000	45 500
Mode	30 000	85 000	62 000	75 000	85 000
Std. deviation	34 819	43 130	39 106	49 472	42 983

A distribution table reflecting mileage across the data set is presented in Table 4.5.

Table 4.5: Mileage Categorisation

Mileage levels	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
0–25 000 km	295	178	317	235	1 025
25 000–35 000 km	282	162	98	69	611
35 000–60 000 km	289	355	214	202	1 060
60 000–95 000 km	156	241	281	364	1 042
>95 000 km	115	197	98	238	648
TOTAL	**1 137**	**1 133**	**1 008**	**1 108**	**4 386**

4.3.3 Fuel Type

In the 2018 sample, there are only two types of fuel available for entry-level vehicles in South Africa – petrol and diesel. The VW Polo Vivo, Toyota Etios and Hyundai i10 only offer petrol-powered options. The Ford Fiesta offers a diesel option; however, 92,9% of Ford Fiesta vehicles sold are petrol-powered. One of the reasons why only the Ford Fiesta supplies a diesel-powered version as an option could be because diesel and petrol should not be seen as perfect substitutes (Erdem & Şentürk, 2009). A study of used cars in Germany (Prado, 2009) tested some key assumptions about diesel-powered cars and found, for example, that because diesel engines initially cost more than petrol engines, used vehicles equipped with diesel engines are expected to be significantly more expensive. The study found that diesel engines, on the one hand, tend to last longer, consume less fuel and are more suitable for vehicles that require low revolutions but high torque, and are therefore generally more suitable for use in commercial vehicles. Petrol engines, on the other hand, are more suited to high-performance vehicles such as many of the top-end private cars on the road.

The Prado (2009) study may explain why so few entry-level vehicles are diesel-powered. It therefore stands to reason that vehicles that are light, small and purchased predominantly for inner-city driving do not require diesel engines. The fuel-type split in South Africa is reflected in Table 4.6.

Table 4.6: Fuel-type Split in South Africa

Vehicle type	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Petrol	1 137	1 133	936	1 108	4 314
Diesel	-	-	72	-	72
TOTAL	**1 137**	**1 133**	**1 008**	**1 108**	**4 386**

4.3.4 Transmission Type

The Ford Fiesta, Hyundai i10 and the VW Polo Vivo are offered to buyers with the option of either manual or automatic transmissions. Only the Toyota Etios is equipped solely with manual transmission. As can be seen in Table 4.7, for all vehicle types with an automatic variant, the proportion of automatic used cars available for sale is very small, with the highest being the Hyundai i10 at 11,2% and the lowest of the other three makes being the VW Polo Vivo at 4,7%. Even though automatic gearboxes are more affordable today, the price of a used car with an automatic gearbox is expected to be higher (Erdem & Şentürk, 2009).

Table 4.7: Transmission Split in South Africa

Vehicle type	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Manual	1 010	1 133	9 22	1 056	4 121
Automatic	127	-	86	52	265
TOTAL	**1 137**	**1 133**	**1 008**	**1 108**	**4 386**

4.3.5 Region (Province of Origin)

Table 4.8 shows the number of vehicles recorded per region. From Table 4.8, the largest occurrence of second-hand vehicles is in Gauteng (2 863), followed by KwaZulu-Natal (504) and Cape Town (475). Studies by Andrews and Benzing (2007), Wu et al. (2015) and Du et al. (2018) show that geographical location is a key variable in the pricing of used cars. This is confirmed by Priilaid and Hendry (2019), where they show that province of origin is one of the top three determinants of hedonic price of used cars in South Africa. They suggest that price arbitrage opportunities might arise because of these interregional price discrepancies.

Table 4.8: Region (Province of Origin) Split in South Africa

Province	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Eastern Cape	28	26	19	27	100
Free State	20	26	6	17	69
Gauteng	609	739	737	778	2 863
KwaZulu-Natal	237	84	85	98	504
Limpopo	38	26	11	11	86
Mpumalanga	50	41	13	21	125
North West	49	41	21	39	150
Northern Cape	0	9	4	1	14
Western Cape	106	141	112	116	475
TOTAL	**1 137**	**1 133**	**1 008**	**1 108**	**4 386**

Figure 4.1 illustrates the percentage split by region of the vehicles recorded.

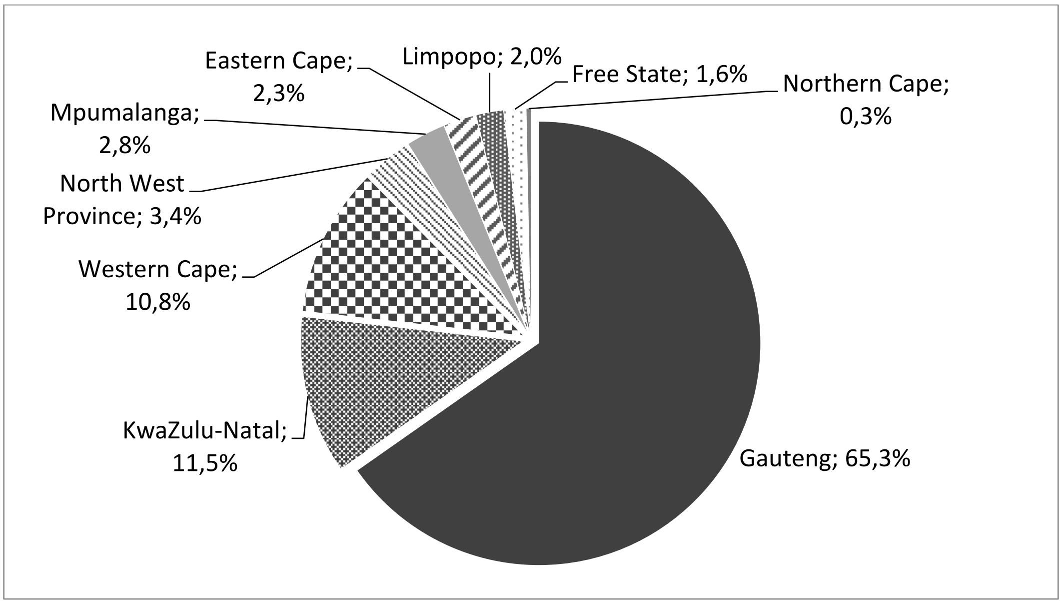

Figure 4.1: Percentage Split of Used Cars by Region (Province of Origin)

4.3.6 Vehicle Colour

The sixth variable in this research study focused on whether statistically significant price differences exist as a result of vehicle colours, where each colour was reviewed relative to the standard white base colour. In Erdem and Şentürk (2009), black and grey colours were found to be key factors driving the hedonic pricing of used cars in Turkey.

4.3.7 Engine Size

Five different engine sizes, measured by volume in litres, were reviewed in this study. The fact that some producers offer a wide range of engine sizes is a strong indicator that this variable is viewed as an important factor that has an impact on car sales. For example, although the Toyota Etios only offers a 1,5-litre engine for the South African market, the Hyundai i10 comes in 1,2-litre, 1,4-litre and 1,5-litre engines and the Ford Fiesta comes with 1,0-, 1,4-, 1,5- and 1,6-litre capacity engines. The VW Polo Vivo is predominantly available with a 1,4-litre engine, but also offers 1,6-litre engines for sale.

Figure 4.2 provides an indication of this engine size split across each of the entry-level vehicles researched, varying from one for the Toyota Etios up to four each for the Hyundai and Ford Fiesta.

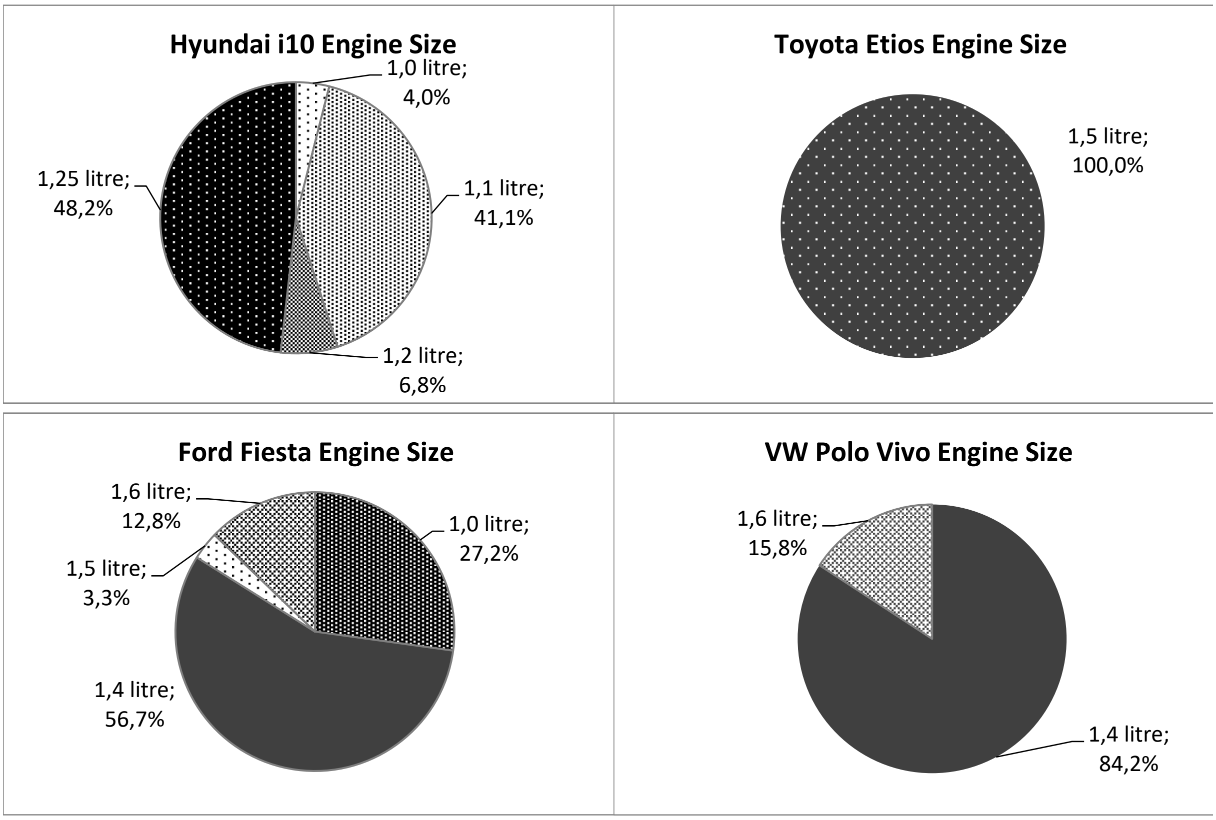

Figure 4.2: Split of Used Cars by Engine Size

4.3.8 Sub-model Variants

The last of the variables selected for this study were the variants on the standard model. In this entry-level segment, only Ford and Toyota carried sub-model versions. In the Ford range, relative to the *Ambient* model, three of the remaining five discernible sub-models carried some statistical significance, namely the *Ecoboost*, *Titanium* and *Sport* models. The remaining sub-models were all found to be statistically insignificant, namely Ford's *Hatch-back* and *Trend*, as well as Toyota's X_s relative to its X_i. Figure 4.3 illustrates the proportionate sales of the two variants offered by Toyota – the X_s and its relative, the X_i.

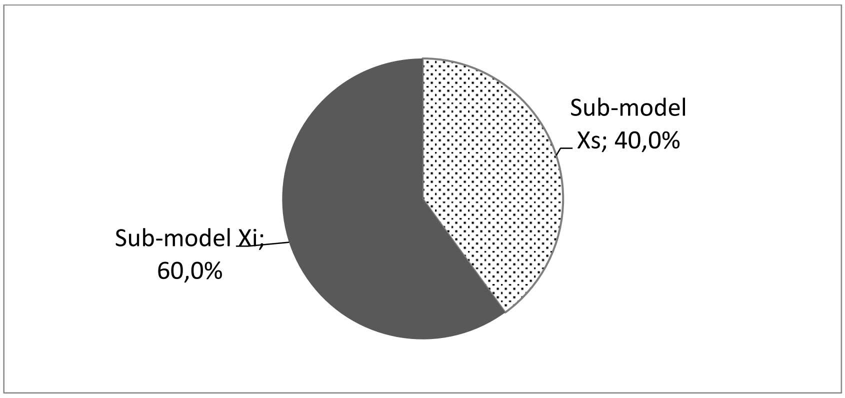

Figure 4.3: Split of Sub-model Variants for Toyota Etios

4.3.8.1 Price

Prices are provided in South African rand (1US$ = ZAR14,00) and were sourced from the website cars.co.za in 2018.

Table 4.9 provides the descriptive statistics for the study sample of 4 386 motor vehicles, as sourced from the website cars.co.za in 2018. The following descriptive statistics are included:

1. Central tendency statistics (mean, median, mode)
2. Distribution statistics in the form of frequency distributions (summary of the frequency of individual variables)
3. Dispersion statistics (standard deviation, variance and range).

Table 4.9: Descriptive Statistics – Price

Vehicle type	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Observations	1 137	1 133	1 008	1 108	4 386
Minimum	54 500	9 990	59 900	65 000	9 990
Maximum	214 900	158 995	424 900	204 900	424 900
Mean	129 844	122 140	160 365	122 651	133 051

Vehicle type	Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo	TOTAL
Median	129 900	129 900	159 900	119 900	129 990
Mode	129 900	149 900	159 900	85 000	149 900
Std. deviation	27 977	25 952	48 679	31 596	37 485

The data comprises four separate data sets across the four vehicle types. The data sets are similar in size, varying from 1 008 to 1 137 observations. The mean and median for the Hyundai i10, Ford Fiesta and VW Polo Vivo data sets are very close in value, with the median differing by less than 2,5% in the case of the VW Polo Vivo data set and by less than 0,5% in the case of the Hyundai i10 and Ford Fiesta data sets. This implies that the observations in these data sets are distributed evenly around the mean.

However, in the case of the Toyota Etios data set, the price median is as much as 6,35% above the mean, implying that the data is somewhat skewed. In Figure 4.4, the distribution of the Toyota Etios prices, grouped in bands of R10 000, is plotted alongside the Hyundai i10 for purposes of comparison. One can see that the Toyota Etios data is skewed to the right compared to the approximately even distribution around the mean in the Hyundai i10 data set.

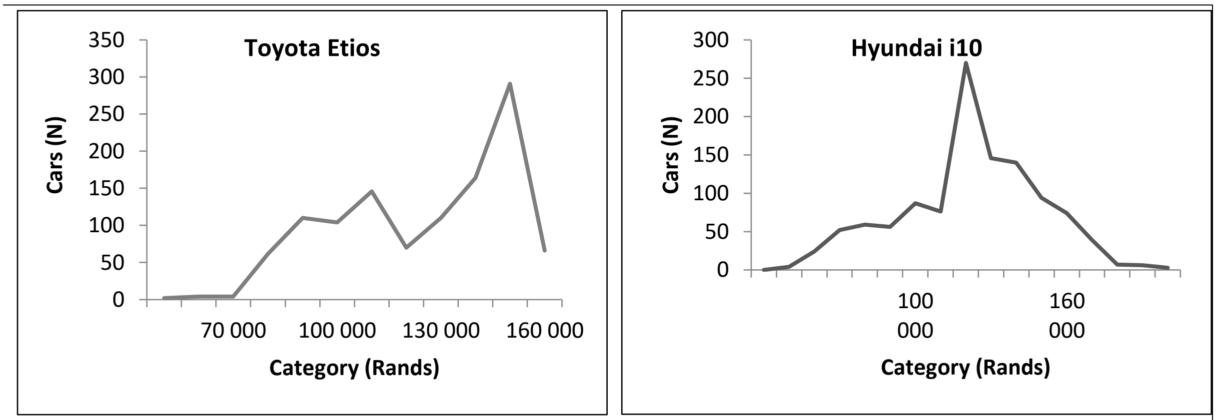

Figure 4.4: Comparison of Toyota Etios and Hyundai i10 Pricing Distribution

From the descriptive statistics set out in Table 4.9, the mean average price for the sample of cars is R133 051. The price maximum, price minimum and standard deviation for the sample are R424 900, R9 990 and R37 485, respectively. From the study, the Ford Fiesta is, on average, the most expensive second-hand entry-level car in South Africa (R160 365), followed by the Hyundai i10 (R129 844), the VW Polo Vivo (R122 651) and the Toyota Etios (R122 083). Of the four vehicles studied, the Hyundai i10 is the most prolific, accounting for 25,9% of all sales in the sample.

This chapter described the data set used in this study. The data set spans the period 2008 to 2018 and was used to identify, compare the linear and non-linear hedonic pricing of four entry-level vehicles in the South African used-car market.

In Chapter Five that follows, the research methodology is set out.

CHAPTER 5: METHODOLOGY AND MODEL CONSTRUCTION

“To most economists, the single equation least-squares regression model, like an old friend, is tried and true.”

(Farrar & Glauber, 1967)

5.1 Introduction

This study leveraged the work done by Van Rensburg and Priilaid (2004) in their development of a predictive statistical valuation methodology for determining the relationship between price and value of South African wine and by Priilaid and Hendry (2019) in their development of a methodology to determine the efficacy of explanatory variables for the pricing of South African motor vehicle brands.

For the first time, this study developed and compared eight price regression models (four linear models and four non-linear models) to explore the effect of eight explanatory variables on the pricing of four second-hand entry-level motor vehicle brands.

For the four linear price regression models, ‘linear’ assumes that there is a straight-line relationship between various explanatory variables and price. For the four ‘non-linear’ models, the researcher used the dummy technique to emphasise individual increments in the top three explanatory variables, namely age, mileage and region (province of origin).

This chapter explains the methodology used and sets out the specification employed for comparing and analysing the four linear and four non-linear (dummy-styled) pricing models that follow. The chapter is organised as follows: Section 5.2 defines linear vs. non-linear valuation techniques. Section 5.3 sets out the rationale behind the choice of explanatory variables or price predictors. Section 5.4 details the methodology and the model specification employed in the empirical analysis and Section 5.5 concludes.

5.2 Linear vs. Non-linear Models

For the linear model, it was assumed that the fair pricing increment associated with each successive increment in economic variable is equal, i.e. there is a straight-line relationship between the implicit value of the economic variable or characteristic and its impact on price. In the main, this linear approach provided workable solutions for Van Rensburg and Priilaid (2004) in their wine valuations; however, they found that at the extreme ends of the quality spectrum, there seemed to be pricing inconsistencies, such as mispricing.

Similarly, in their study of the dynamic relationship between oil price and inflation in South Africa, Balcilar, Uwilingiye and Gupta (2018) found that due to the nature of a linear regression, its accuracy in describing the linear relationship in real scenarios could be imprecise. They found that, with respect to time, a valuable item does not appreciate or depreciate in a consistently linear fashion, and that as the linear regression model moves further from a central mean, so the prediction errors grow larger.

By way of remedy, this study incorporated what Priilaid and Van Rensburg (2006) refer to as a dummy-styled approach to address the possibility of a non-linear relationship between value and price. This study demonstrated that the dummy or non-ordinal approach results in more economically sensible valuations across the value continuum, thereby assisting both buyers and sellers in better identifying cars that offer value for money (Priilaid & Van Rensburg, 2007).

5.3 The Characteristics of Goods and How to Pick Them

Based on the discussion of characteristics in Chapter Two and the discussion of the data and descriptive statistics presented in Chapter Four, a series of specific stepwise regressions was developed, each to explain and contrast the implicit impact of eight functional variables driving the price of cars. The eight functional variables are age, mileage, region (province of origin), transmission type, fuel type, colour, engine size and sub-model variant. The eight pricing models provide a unique opportunity in the literature to compare the marginal, relative impact of a linear regression model with that of a non-linear model.

In the linear model, age was regressed against price. In the non-linear model, age was broken up into calendar year bands by means of dummy variables and then regressed against price. The relative split of the calendar year bands can be seen in Figure 5.1.

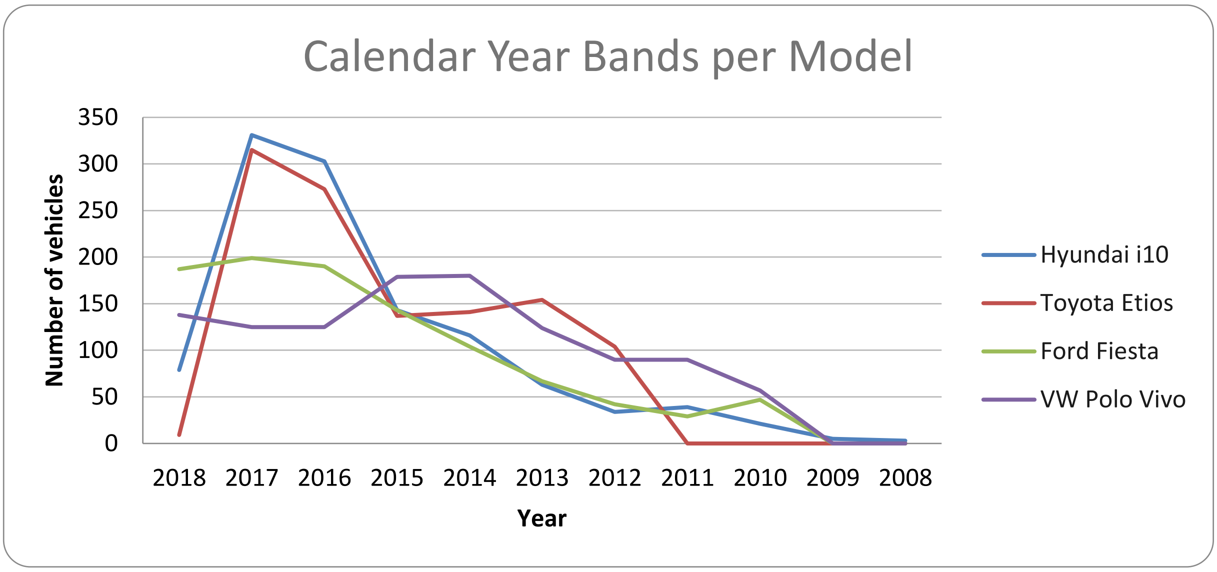

Figure 5.1: Calendar Year Bands per Model

In the linear model, mileage was regressed against price. In the non-linear model, mileage was broken up into bands that roughly split the data evenly by means of dummy variables and then regressed against price. The split of mileage into bands can be seen in Figure 5.2.

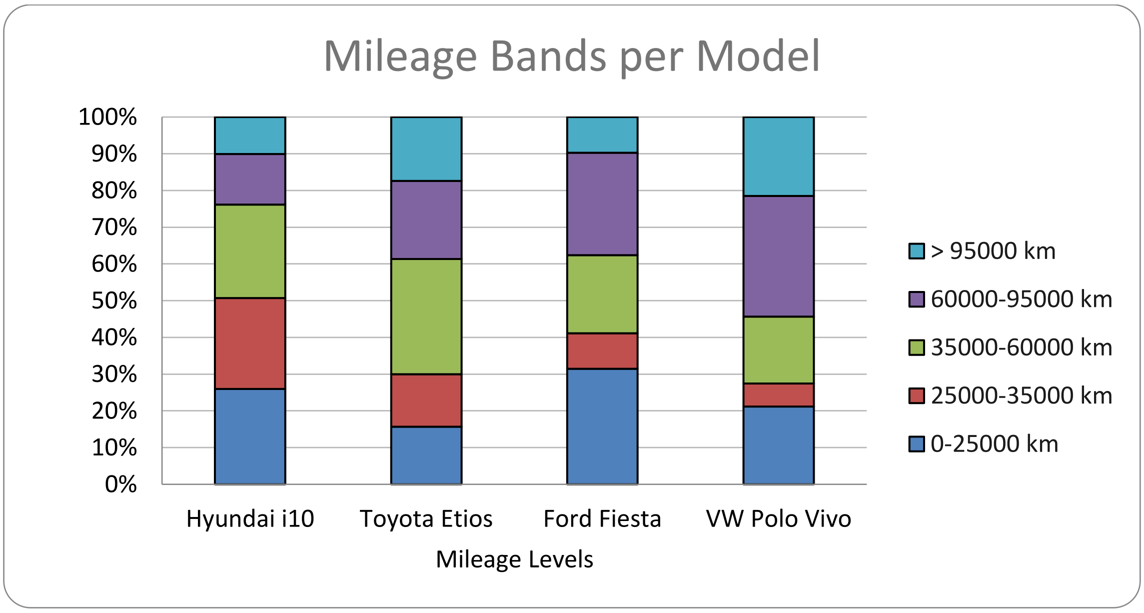

Figure 5.2: Mileage Bands per Model

As already discussed, the specification of the explanatory variables included only the top eight variables that were most likely to influence hedonic quality. A combination of insight, experience and theory (see Farrar and Glauber, 1967) saw the following eight variables chosen as explanatory variables for the purposes of this study; age, mileage, region (province of origin), transmission type, fuel type, colour, engine size and sub-model variant.

The choice of each variable was based on economic grounds, i.e. variables were assumed to be “utility-bearing attributes” (Greenstone, 2017). Based on the work by Priilaid and Hendry (2019) and their use of OLS regression models to reduce the top eight characteristics even further in terms of statistical significance, empirical evidence shows that the three most statistically relevant variables are age, mileage and province of origin. While characteristics or variables were discussed in detail in Chapter Two, what follows here is a rationale for the methodology employed when choosing the correct variables in the first instance.

One of the principal difficulties in hedonic regression is the choice of the variables to be retained in the model (Moresino, 2019). Data limitations, rather than theoretical limitations, are normally responsible for underspecified or oversimplified marketing models (Liu, 1960). However real the dependency relationship between Y and each member of a set of independent variables (X) may be, the growth of interdependence within X as its size increases can only decrease the marginal, relative significance of the contribution of each independent variable (Farrar & Glauber, 1967). Even in second-hand cars, the researcher is faced with a plethora of possible variables, some with strong significance and many less so. When deciding what characteristics to include in the hedonic equation, the researcher faces the usual trade-off between omitted variable bias from excluding too many variables and problems related to multicollinearity from including too many variables (Andersson, 2005). This predicament of knowing which predictors to include and which to exclude is exactly why a stepwise regression can be so useful. Per the literature review in Chapter Two, stepwise regression, despite being somewhat controversial, remains a popular tool and is a widely used method for modelling the impact of variables in a hedonic regression. Here, variables are introduced step by step in the model and their relative impact is measured by means of a t-statistic (Moresino, 2019). The t-statistic is a ratio showing the difference between an estimated value and its standard error. It determines the extent to which the researcher should support or reject the null hypothesis. The higher the t-value, the higher the confidence in the coefficient as a predictor. In general, any t-value greater than +2 or less than -2 is acceptable.

The stepwise regression method involves automatically selecting variables for a model based on statistical criteria, and while it has significant advantages, there are several drawbacks to consider. Stepwise regression may lead to overfitting, especially in the presence of noise or irrelevant variables. Overfitting occurs when the model fits the training data too closely, capturing random fluctuations as if they were meaningful patterns. The risk is that the overfitted models tend to perform poorly on new, unseen data because they have essentially memorised the noise in the training data, rather than capturing true underlying patterns. To prevent overfitting, this model relies on careful feature selection and a large data set. By focusing only on the eight most informative variables, chosen using domain knowledge, the model becomes less prone to overfitting. A large data set also helps the model to generalise better, as data augmentation introduces variability without adding new information.

Multicollinearity is another issue that can arise from stepwise regression. This occurs when two or more independent variables in a regression model are highly correlated, making it challenging to distinguish the individual effects of each variable on the dependent variable. Overcoming multicollinearity issues is crucial for obtaining reliable and interpretable results. Once again, the combined approach of using a large data set and careful variable selection, which limited predictors to only the most essential, reduced the chances of collinearity.

5.4 Methodology and Model Construction

In essence, this study sought to explain how a competitive market solves the problem of matching buyers and sellers of differentiated or multidimensional goods, as termed by Greenstone (2017). These multidimensional or heterogeneous goods derive their value from the sum of the individual values pertaining to the homogeneous, independent economic variables or characteristics (Triplett, 1986). The economic variables or characteristics are valued by both buyers and sellers, which is what supplies their perceived economic value. In this study, the competitive market was the South African used-car market, the multidimensional or heterogenous goods were entry-level second-hand cars and the homogeneous, independent economic variables or characteristics from which the cars derive their value were the eight key attributes discussed in Chapter Four, namely age, mileage, region (province of origin), transmission type, fuel type, colour, engine size and sub-model variant. Furthermore, for the purposes of this study, the advantage of drawing all the relevant data from one website was that the researcher did not need to control for unknowns such as brand effects across different online used-car retailers. Similarly, by conducting this study over one year (2018), the effects of time-dependent variables such as inflation, changes in consumer confidence and growth in household income were minimised and therefore effectively controlled for, as supported by Mok, Chan and Cho (1995) and Priilaid and Hendry (2019).

As it is not possible to allocate individual prices to each independent characteristic of a vehicle, the purchasing decision is based largely on the explicit price for the vehicle. In that sense, the price represents the sum of the marginal, relative, implicit prices of all the characteristics that make up the car. In other words, the price of the vehicle (dependent variable) depends on the sum of the marginal, relative, implicit prices of all the attributes (independent variables) of the car in question.

In this pricing analysis, the OLS regression methodology was employed. In so doing, the cross section of prices (price) was modelled for cars (i = 1, …, n) as a function of consumer desirable characteristics (K),
where (CDk = 1, …, K) and:

α = the intercept term as estimated by OLS
b = the K slope coefficients as estimated by OLS
CD = the K consumer desirables or explanatory variables as detailed in Chapter Four
ε = a random residual error term conforming to classic assumptions.

By using a regression model, where the coefficient represents the relationship between each independent variable and the dependent variable, one is able to understand how changes in the predictor values (CD) have an impact on the overall mean of the response variable (i). For example: Suppose you have the following regression equation: Y = 2X + C. In this equation, +2 is the coefficient or slope, X is the independent variable or set of independent variables and C is the constant. The coefficient represents the change in mean (arithmetic average) in the dependent variable given a one-unit change in the independent variable. For example, if the coefficient is +2, the mean value of the dependent variable increases by two units for every one-unit change in the independent variable. The sign of the coefficient shows the direction of the relationship between the variables:

1. A positive sign shows that as the independent variable increases, the dependent variable also increases.
2. A negative sign shows an inverse relationship between the independent variable and the dependent variable, i.e. as the predictor variable increases, the response variable decreases.

As the independent, price explanatory variables (product cues) explain changes in the dependent variable (price), what is interesting is the following:

1. The extent to which a marginal, relative change in one or more independent or predictor variable changes the dependent variable (price).
2. The extent to which one can compare the marginal, relative impact on price of one or more independent variables relative to the impact on price of other independent variables.
3. The role of the hedonic price function as a 'disaggregator'.

Hedonic modelling allows one to disaggregate the price of any heterogenous good into the relative, implicit prices and quantities of key characteristics. Because the prices are derived from an predictive analytical model, rather than directly observed, they are usually termed 'implicit' prices. The hedonic pricing function provides a credible, well-established model that has been tested over a broad range of products. As virtually all goods are made up of multiple characteristics, it is a versatile model with many applications.

5.4.1 Linear Specification

In constructing the eight stepwise regression models, two key steps are required:

Step 1: Value each car hedonically by employing linear analysis (addressing Problem Statement One as set out in Chapter One).

As the 'intrinsic value' of each car, *i*, is then estimated, the fitted estimates yield the measure of intrinsic value; with the extent of mispricing (ε_i) approximating the difference between value and price. This done, each car's price and value may be graphically depicted, allowing for the identification of that space on the graph where maximum value relative to price may be identified. In wine and whisky valuation studies, this region has been termed the 'value frontier' (Priilaid, 2016; Priilaid & Van Rensburg, 2006).

In order to avoid the 'dummy trap', the following categorical comparators are specified: fuel type: "petrol"; transmission type: "manual"; engine size: variable as specified above; region (province of origin): "Gauteng"; colour: "white". Therefore, the coefficients derived in each model should be reflected against these comparators, which in aggregate represent the derived constant intercept term α.

Step 2: Value each car using a non-linear regression analysis in order to compare the statistical feasibility of the linear vs. non-linear models (addressing Problem Statement Two, as set out in Chapter One).

If one expands on the equation in section 5.4 above, the final linear model for each of the four vehicles in this study reads as follows:

$$\begin{aligned}
\text{Price}_i &= \alpha + b_1{*}\text{Age} + b_2{*}\text{Mileage} + b_3{*}D_{\text{Diesel}} + b_4{*}D_{\text{Automatic}} + b_5{*}D_{\text{Engine size 1.1}} + \\
& b_6{*}D_{\text{Engine size 1.2}} + b_7{*}D_{\text{Engine size 1.25}} + b_8{*}D_{\text{Engine size 1.4}} + b_9{*}D_{\text{Engine size 1.5}} + b_{10}{*}D_{\text{Engine size 1.6}} \\
& + b_{11}{*}D_{\text{WC}} + b_{12}{*}D_{\text{EC}} + b_{13}{*}D_{\text{FS}} + b_{14}{*}D_{\text{NW}} + b_{15}{*}D_{\text{KZN}} + b_{16}{*}D_{\text{NC}} + b_{17}{*}D_{\text{LIM}} + b_{18}{*}D_{\text{MPH}} + \\
& b_{19}{*}D_{\text{Blue}} + b_{20}{*}D_{\text{Red}} + b_{21}{*}D_{\text{Brown}} + b_{22}{*}D_{\text{Green}} + b_{23}{*}D_{\text{Black}} + b_{24}{*}D_{\text{Silver}} + b_{25}{*}D_{\text{Grey}} + \\
& b_{26}{*}D_{\text{Metallic}} + b_{27}{*}D_{\text{Orange}} + b_{28}{*}D_{\text{Gold}} + b_{29}{*}D_{\text{Beige}} + b_{30}{*}D_{\text{Yellow}} + b_{31}{*}D_{\text{Other}} + \\
& \text{Sub-model variant} + \varepsilon_i
\end{aligned}$$

Employing this methodology, a valuation model was estimated for pricing effects stemming from the age, mileage, region (province of origin), transmission type, fuel type, colour, engine size and sub-model variant of each make of vehicle,
where:

α = the estimated intercept term of the regression

D_{Diesel} = 1 if the engine runs on diesel, 0 if otherwise

$D_{Automatic}$ = 1 if the engine is an automatic transmission, 0 if otherwise

$D_{Engine\ size\ 1.1}$ = 1 if the engine size is 1.1cc, 0 if otherwise

$D_{Engine\ size\ 1.2}$ = 1 if the engine size is 1.2cc, 0 if otherwise

$D_{Engine\ size\ 1.25}$ = 1 if the engine size is 1.25cc, 0 if otherwise

$D_{Engine\ size\ 1.4}$ = 1 if the engine size is 1.4cc, 0 if otherwise

$D_{Engine\ size\ 1.5}$ = 1 if the engine size is 1.5cc, 0 if otherwise

$D_{Engine\ size\ 1.6}$ = 1 if the engine size is 1.6cc, 0 if otherwise

D_{WC} = 1 if the car is for sale in the Western Cape, 0 if otherwise

D_{EC} = 1 if the car is for sale in the Eastern Cape, 0 if otherwise

D_{FS} = 1 if the car is for sale in the Free State, 0 if otherwise

D_{NW} = 1 if the car is for sale in the North West, 0 if otherwise

D_{KZN} = 1 if the car is for sale in KwaZulu-Natal, 0 if otherwise

D_{NC} = 1 if the car is for sale in the Northern Cape, 0 if otherwise

D_{LIM} = 1 if the car is for sale in Limpopo, 0 if otherwise

D_{MPH} = 1 if the car is for sale in Mpumalanga, 0 if otherwise

D_{Blue} = 1 if the car colour is blue, 0 if otherwise

D_{Red} = 1 if the car colour is red, 0 if otherwise

D_{Brown} = 1 if the car colour is brown, 0 if otherwise

D_{Green} = 1 if the car colour is green, 0 if otherwise

D_{Black} = 1 if the car colour is black, 0 if otherwise

D_{Silver} = 1 if the car colour is silver, 0 if otherwise

D_{Grey} = 1 if the car colour is grey, 0 if otherwise

$D_{Metallic}$ = 1 if the car colour is metallic, 0 if otherwise

D_{Orange} = 1 if the car colour is orange, 0 if otherwise

D_{Gold} = 1 if the car colour is gold, 0 if otherwise

D_{Beige} = 1 if the car colour is beige, 0 if otherwise

D_{Yellow} = 1 if the car colour is yellow, 0 if otherwise

D_{Other} = 1 if the colour or make of car is not previously cited, 0 if otherwise

b_1 = the marginal effect of vehicle vintage, where 2018 = year 1, 2017 = year 2, etc.

b_2 = the marginal effect of the vehicle mileage

b_3 = the marginal effect of a diesel engine

b_4 = the marginal effect of automatics

b_5 *to* b_{10} = the marginal effect of 1.1, 1.2, 1.25, 1.4, 1.5 and 1.6cc engines, relative to the designated base case per vehicle make, these being 1.0 for the Ford Fiesta and Hyundai i10, and 1.4 for the VW Polo Vivo; the Toyota Etios comes with a uniform 1.5 engine size

b_{11} *to* b_{18} = the marginal effect of the following provinces of origin relative to the Gauteng base case: Western Cape, Eastern Cape, Free State, North West, KwaZulu-Natal, Northern Cape, Limpopo and Mpumalanga

b_{19} *to* b_{31} = the marginal effect of blue, red, brown, green, black, silver, grey, metallic, orange, gold, beige, yellow and colours designated as 'other' relative to the base-case colour of white.

5.1.1 Non-linear Specification

If one introduces dummy variables into the equation in section 5.4 above, one can demonstrate the non-linear effect of age and mileage on the perceived value of the cars. For the purposes of the non-linear modelling, the following dummy specifications were used to update age and mileage:

A_1 = 1 if the car was registered in 2018, 0 if otherwise

A_2 = 1 if the car was registered in 2017, 0 if otherwise

A_3 = 1 if the car was registered in 2016, 0 if otherwise

A_4 = 1 if the car was registered in 2015, 0 if otherwise

A_5 = 1 if the car was registered in 2014, 0 if otherwise

A_6 = 1 if the car was registered in 2013, 0 if otherwise

A_7 = 1 if the car was registered in 2012, 0 if otherwise

A_8 = 1 if the car was registered in 2011, 0 if otherwise

A_9 = 1 if the car was registered in 2010, 0 if otherwise

A_{10} = 1 if the car was registered in 2009, 0 if otherwise

A_{10} = 1 if the car was registered in 2008, 0 if otherwise

M_1 = 1 if the car's odometer reads 0–25 000 km, 0 if otherwise

M_2 = 1 if the car's odometer reads 25 000–35 000 km, 0 if otherwise

M_3 = 1 if the car's odometer reads 35 000–60 000 km, 0 if otherwise

M_4 = 1 if the car's odometer reads 60 000–95 000 km, 0 if otherwise

M_5 = 1 if the car's odometer reads >95 000 km, 0 if otherwise.

5.5 Conclusion

For the first time, eight price regression models (four linear and four non-linear) were developed to explore the effect of explanatory variables on the pricing and valuation of entry-level second-hand motor vehicles in South Africa. This chapter meticulously set out the methodology and model construction employed. The distinction between linear and non-linear models was elucidated, highlighting the traditional assumptions of a straight-line relationship in linear models and the incorporation of a dummy-styled approach in non-linear models. In the pursuit of understanding the influence of various characteristics on vehicle pricing, the chapter outlined the rationale behind the selection of eight key explanatory variables, with a specific focus on age, mileage and province of origin as the most statistically relevant factors. The stepwise regression approach was justified, providing a systematic means to address the trade-offs associated with variable inclusion and exclusion in the hedonic equation.

The application of OLS regression was thoroughly explained. The hedonic pricing function employed in this study has proven to be a versatile and credible model, providing a means to disaggregate the price of heterogeneous goods into the implicit prices of their individual characteristics. In this case, the marginal, relative value of each characteristic was statistically proven. What follows in the next chapter is the analysis, conclusions and findings, which, together with the rigorous methodology and model construction employed in this chapter, contribute to the overall advancement of predictive valuation techniques in the context of the South African used-car market.

CHAPTER 6: FINDINGS AND DISCUSSION

"Price is what you pay, value is what you get."

(Van Rensburg & Priilaid, 2004)

6.1 Introduction

This chapter sets out findings and discussions as the researcher leveraged both advanced linear and non-linear price regression analyses into a comprehensive examination of pricing determinants of four prominent entry-level vehicles: the Ford Fiesta, Hyundai i10, Toyota Etios and VW Polo Vivo.

The chapter begins by scrutinising the constants derived from each of the eight regression models. These serve as a baseline representation of vehicle pricing under controlled conditions. This introductory analysis establishes a fundamental understanding of the factors influencing the intercept, illuminating the base-case scenarios for these vehicles; specifically, brand-new, manual, petrol-driven white vehicles, purchased in Gauteng, with zero mileage.

The analysis showed that age was by far the most statistically relevant variable. The empirical evidence clearly showed the impact of calendar years on vehicle pricing. The half-life calculations provided valuable insights into the duration it takes for each vehicle's value to diminish by half, serving as a useful metric for both prospective buyers and sellers.

Mileage is another critical variable but showed significantly less statistical relevance compared to age. The differential impact of mileage across the four vehicles shed light on its significance or lack thereof in shaping pricing dynamics.

Furthermore, this chapter undertakes an examination of adjusted R-squared values, gauging the goodness of fit for both the linear and the non-linear regression models. These values served as a yardstick for the accuracy and explanatory power of the models, paving the way for a nuanced comparison between the two methodologies.

Through a detailed discussion of findings, this chapter elucidates the intricate interplay between age, mileage, region (province of origin) and other categorical variables, offering an enriched comprehension of the complex pricing determinants in the entry-level vehicle market. This knowledge not only contributes to the academic discourse on automotive economics, but also holds practical implications for industry practitioners, policymakers and consumers navigating the ever-evolving landscape of the second-hand vehicle market.

6.2 Preliminary Findings and Discussion

Section 6.2 sets out the eight tables showing the four linear and four non-linear regression analyses. Each table is followed by a discussion. Section 6.3 summarises the findings and discussion and Section 6.4 concludes.

6.2.1 Linear Findings and Discussion

The following four tables present the results of the linear regression analysis. Each table is followed by a discussion.

Table 6.1: Model 1: Hyundai i10 – Linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	171 109,84	1 192,31	143,511	<0,0001
Age	-7 326,10	339,81	-21,560	<0,0001
Mileage	-0,15	0,02	-8,692	<0,0001
Western Cape	2 079,34	1 450,04	1,434	0,152
KwaZulu-Natal	3 307,16	1 058,89	3,123	0,002
Eastern Cape	6 744,71	2 646,98	2,548	0,011
Limpopo	4 233,95	2 284,72	1,853	0,064
Mpumalanga	7 144,95	2 029,68	3,520	0,000
North West	3 326,61	2 035,03	1,635	0,102
Free State	5 295,94	3 100,75	1,708	0,088
Automatic	14 732,12	1 319,95	11,161	<0,0001
Blue	-745,51	1 386,09	-0,538	0,591
Red	-2 154,23	1 539,76	-1,399	0,162
Grey	-7 177,39	1 271,03	-5,647	<0,0001
Silver	-3 576,49	1 333,82	-2,681	0,007

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Orange	6 935,73	4 193,18	1,654	0,098
Metallic	5 388,97	1 279,24	4,213	<0,0001
Engine 1.2	-20 605,01	1 794,55	-11,482	<0,0001
Engine 1.1	-22 379,21	893,42	-25,049	<0,0001
Engine 1.0	-14 997,31	2 168,51	-6,916	<0,0001

The intercept for the Hyundai i10 (171 110) had a t-score of 143,51, indicating that it was highly significant in the model. Age had a t-score of -21,560, suggesting that it was highly significant. As the t-score was negative, it indicates a negative relationship with vehicle price. With a t-score of -8,692, mileage was also significant and negatively related to vehicle price.

From Model 1 set out above, it is clear that relative to the base case of Gauteng, the Western Cape with a value of 2 079 (t-score: 1,434) did not show a significant price premium. The opposite is true of Mpumalanga with a value of 7 145 (t-score: 3,520) and the Eastern Cape with a value of 6 745 (t-score: 2,548), where a general price-to-distance relationship was evident, with large and statistically significant price premia identified.

It is interesting to note that the 1.1-litre engine had a value of -22 379 (t-score: -25,049), which means this engine size had a highly significant and negative impact on price. By the same token, the 1.0-litre engine with a value of -14 997 (-6,960) seemed to have a significantly lower impact on the price of the Hyundai i10. The 1.0-litre engine seems to be a more popular engine.

This analysis attempted to discern price differences between colours per vehicle relative to the standard white base-case offering. Overall, the significance of colour on vehicle price was very low. This was particularly true for the Hyundai i10. In the linear regression, only grey, with a value of -7 177 (t-score: -5,647), indicating an oversupply, and metallic, with a value of 5 389 (4,213), indicating a scarcity of supply, were statistically relevant. In other words, only two colours out of six played a role in determining the value of the vehicles.

Table 6.2: Model 2: Toyota Etios – Linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	159 972,04	4 783,79	33,440	<0,0001
Age	-11 278,65	356,28	-31,656	<0,0001
Mileage	-0,04	0,01	-2,889	0,004
Mpumalanga	12 267,31	5 296,64	2,316	0,021
North West	10 583,16	5 304,41	1,995	0,046
KwaZulu-Natal	11 306,34	5 077,15	2,227	0,026
Western Cape	9 854,08	4 951,69	1,990	0,047
Eastern Cape	11 087,71	5 547,81	1,999	0,046
Free State	5 438,41	5 605,00	0,970	0,332
Limpopo	13 243,13	5 593,25	2,368	0,018
Northern Cape	13 530,34	4 873,48	0,647	0,518
Black	-5 363,05	2 753,30	-1,948	0,052
Blue	9 180,37	2 089,32	4,394	<0,0001
Gold	5 078,29	1 238,34	4,101	<0,0001
Orange	9 501,56	14 553,72	0,653	0,514
Red	3 265,13	1 758,03	1,857	0,064
Silver	3 709,02	1 071,65	3,461	0,001
Engine 1.5	-168,56	868,10	-0,194	0,846

The intercept for the Toyota Etios (159 972) was significant with a t-score of 33,440. As expected, age was significantly negative with a t-score of -31,656, while mileage was also negatively related to vehicle price with a t-score of -2,889.

In Model 2 above, a general price-to-distance relationship is clear, with a large and statistically significant price premium identified in Limpopo with a value of 13 243 (t-score: 2,368), compared to the Free State with a value of 5 438 (t-score: 0,970), suggesting that this province is too close to Gauteng to warrant a distance premium. This information could play an important role in helping customers to determine which provinces provide better value for money. This in turn has the effect of forcing vehicle prices down in the more affluent areas.

The Toyota Etios came with a standard engine size of 1.5-litre engine and a value of -168,56 (t-score: -0,194), indicating that the engine size is not a significant price differentiator.

Overall, the significance of colour on price for the Toyota Etios was very low, with only three colours featuring as statistically significant, namely blue with a value of 9 180 (t-score: 4,394), gold with a value of 5 078 (t-score: 4,101) and silver with a value 3 709 (t-score: 3,461). The Toyota Etios carried the greatest number of significant colour premia (blue, gold and silver – all positive), followed by the Hyundai i10 (grey and metallic). In other words, colour only played a role in determining the value of the car in two out of the four entry-level vehicles studied.

Table 6.3: Model 3: Ford Fiesta – Linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	220 111,07	2 307,33	95,396	<0,0001
Age	-13 362,94	660,18	-20,241	<0,0001
Mileage	-0,09	0,04	-2,557	0,011
North West	13 138,27	5 864,27	2,240	0,025
Eastern Cape	16 599,57	6 112,76	2,716	0,007
Mpumalanga	8 069,32	7 350,17	1,098	0,273
KwaZulu-Natal	12 978,86	3 054,82	4,249	<0,0001
Free State	16 370,20	10 740,56	1,524	0,128
Northern Cape	18 628,54	13 206,21	1,411	0,159
Western Cape	13 273,95	2 705,49	4,906	<0,0001
Limpopo	24 577,93	8 052,40	3,052	0,002
Blue	-5 687,34	2 541,32	-2,238	0,025
Black	6 196,01	6 106,47	1,015	0,311
Red	11 189,43	2 722,63	4,110	<0,0001
Silver	0,00	0,00		
Brown	-7 130,22	5 128,88	-1,390	0,165
Green	7 539,11	8 068,90	0,934	0,350
Automatic	29 331,13	3 335,94	8,792	<0,0001
Engine 1.4	-27 213,69	2 265,60	-12,012	<0,0001
Engine 1.5	35 228,50	6 762,16	5,210	<0,0001
Engine 1.6	10 521,34	3 680,86	2,858	0,004
Diesel	-19 563,21	4 793,04	-4,082	<0,0001

The intercept for the Ford Fiesta (220 111) was highly significant with a t-score of 95,396. Age was highly significant and negatively related to price with a t-score of -20,241. Mileage was negatively related to car price with a t-score of -2,557. Some provinces had statistically more significant predictors, e.g. the Western Cape with a value of 13 274 (t-score: 4,906) and KwaZulu-Natal with a value of 12 979 (t-score: 4,249). As anticipated, Limpopo, with a value of 24 578 (t-score: 3,052), the Northern Cape, with a value of 18 629 (t-score: 1,411), and the Eastern Cape, with a value of 16 599 (t-score: 2,716), were identified with large and statistically significant price premia.

In their study, Haliloglu and Berument (2021) found that the impact on price of diesel fuel is greater than that of petrol. They surmised this because the price of diesel is cheaper than that of petrol and fuel economy for diesel cars is bigger than that of petrol cars. The researcher's analysis of entry-level vehicles indicated that diesel was used in only one of the four vehicles, namely the Ford Fiesta. With a statistically significant coefficient of -19 563 (t-score: -4,082), there is a price penalty attached to the use of diesel engines versus the more popular petrol versions. This could explain why only 72 of the 1 008 Fiestas sampled were fitted with diesel engines and why Ford Fiesta is the only brand to make a diesel engine at entry level. With fewer than 1,6% of the cars in entry-level category being diesel, managers should recognise that diesel has become a significant price differentiator.

Automatic variants were found in all cars but the Toyota Etios. Across the remaining vehicles, automatic transmissions carried significant (at the 99% cut-off) price premia, namely Ford Fiesta at R29 331, Hyundai i10 at R14 732 and VW Polo Vivo at R7 805. What is noteworthy here is the variance across these three figures, with Ford Fiesta's automatic feature valued as twice as expensive as that of the Hyundai i10 and four times as expensive as that of the VW Polo Vivo. This may be because the Ford Fiesta is front-wheel-driven and equipped with a six-speed automatic transmission that is built around a torque converter, as opposed to a dual-clutch automatic gearbox.

With the Ford Fiesta, relative to the 1.0-litre engine control, the respective values of the 1.4-litre, 1.5-litre and 1.6-litre engines were -27 213, 35 229 and 10 521, respectively. The 1.4-litre engine had a significant t-score of -12,012, while the 1.6-litre engine model had a t-score that was much less significant at 2,858.

Red, with a value of 11 189 (t-score: 4,110) was the only colour that was statistically relevant in predicting the price of a Ford Fiesta.

Table 6.4: Model 4: VW Polo Vivo – Linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	163 719,715	1 392,853	117,543	<0,0001
Age	-10 556,533	330,510	-31,940	<0,0001
Mileage	0,003	0,015	0,180	0,857
Mpumalanga	15 454,336	4 041,188	3,824	0,000
North West	19 003,710	2 895,772	6,563	<0,0001
KwaZulu-Natal	15 421,383	1 919,855	8,033	<0,0001
Western Cape	17 796,618	1 771,520	10,046	<0,0001
Eastern Cape	16 791,658	3 462,785	4,849	<0,0001
Free State	16 384,538	4 331,494	3,783	0,000
Limpopo	22 582,252	5 351,350	4,220	<0,0001
Northern Cape	31 496,427	17 735,778	1,776	0,076
Silver	-1 549,032	1 258,097	-1,231	0,218
Red	-1 191,291	2 161,186	-0,551	0,582
Beige	9 462,894	3 570,644	2,650	0,008
Gold	-7 234,717	2 130,435	-3,396	0,001
Blue	3 760,653	3 031,905	1,240	0,215
Grey	5 253,510	5 644,302	0,931	0,352
Black	5 676,760	3 638,031	1,560	0,119
Yellow	607,390	5 593,753	0,109	0,914
Other	-3 688,453	6 711,468	-0,550	0,583
Orange	16 203,902	7 946,205	2,039	0,042
Silver gold	-15 613,924	8 842,152	-1,766	0,078
Cream	-24 625,364	17 608,725	-1,398	0,162
Navy blue	-12 149,093	10 176,003	-1,194	0,233
Champagne	-8 685,302	8 844,467	-0,982	0,326
Automatic	7 805,499	2 546,299	3,065	0,002
Engine 1.6	9 056,410	1 515,711	5,975	<0,0001

The intercept for the VW Polo Vivo (163 720) was highly significant with a t-score of 117,543. Age was highly significant and negatively related to price with a t-score of -31,940. It is interesting that with the VW Polo Vivo, mileage did not seem to have a significant impact on price, with a t-score of 0,180.

Some provinces, such as the Western Cape with a value of 17 997 (t-score: 10,046) and KwaZulu-Natal with a value of 15 421 (t-score: 8,0330), were significant price premia with significant price predictors compared to the Northern Cape, with a significant price premium of 31 496, but statistically irrelevant with a t-score of 1,776.

Across the entire sample of cars, engine size varies from between 1,0 litre and 1,6 litres. The VW Polo Vivo came with a standard engine size of 1,6 litres and a value of 9 056 (t-score: 5,975), which was significant at the 99% confidence level.

Of the 13 VW Polo Vivo colours analysed, only three colours showed some statistical relevance, namely gold with a value of -7 234 (t-score: -3,396), orange with a value of 16 203 (t-score: 2,059) and beige with a value of 9 463 (t-score: 2,650).

6.2.2 Non-linear Findings and Discussion

The following four tables present the results of the non-linear regression analysis. Each table is followed by a discussion.

Table 6.5: Model 5: Hyundai i10 – Non-linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	170 643,60	2 012,45	84,794	<0,0001
Yr_2017	-16 168,17	1 986,73	-8,138	<0,0001
Yr_2016	-22 773,04	2 129,58	-10,694	<0,0001
Yr_2015	-33 788,00	2 343,37	-14,419	<0,0001
Yr_2014	-43 872,92	2 438,85	-17,989	<0,0001
Yr_2013	-53 768,19	2 804,90	-19,169	<0,0001
Yr_2012	-60 527,24	3 241,30	-18,674	<0,0001
Yr_2011	-55 704,45	3 308,20	-16,838	<0,0001
Yr_2010	-57 225,65	3 850,03	-14,864	<0,0001
Yr_2009	-58 657,20	6 657,12	-8,811	<0,0001
Yr_2008	-53 655,59	8 053,29	-6,663	<0,0001
25 000	-211,67	1 209,00	-0,175	0,861
35 000	-3 520,48	1 284,80	-2,740	0,006
60 000	-8 707,99	1 659,81	-5,246	<0,0001
95 000	-12 953,11	2 057,07	-6,297	<0,0001
Western Cape	2 743,54	1 425,48	1,925	0,055
KwaZulu-Natal	4 076,32	1 039,75	3,920	<0,0001
Eastern Cape	4 108,96	2 588,62	1,587	0,113

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Limpopo	2 930,67	2 202,01	1,331	0,183
Mpumalanga	6 951,47	1 978,51	3,513	0,000
North West	2 919,85	1 981,33	1,474	0,141
Free State	4 378,05	2 999,15	1,460	0,145
Automatic	14 925,51	1 290,74	11,564	<0,0001
Blue	-776,84	1 338,72	-0,580	0,562
Red	-2 524,12	1 485,88	-1,699	0,090
Grey	-5 932,14	1 237,00	-4,796	<0,0001
Silver	-3 744,19	1 288,81	-2,905	0,004
Orange	6 623,02	4 070,46	1,627	0,104
Metallic	4 048,44	1 250,26	3,238	0,001
Engine 1.2	-22 519,14	1 798,37	-12,522	<0,0001
Engine 1.1	-24 082,87	895,39	-26,896	<0,0001
Engine 1.0	-19 970,53	2 412,45	-8,278	<0,0001

The intercept for the non-linear regression model of the Hyundai i10 (170 644) was highly significant with a t-score of 84,794. In the non-linear price regression model, age was broken up into calendar years, and it is clear from the t-scores ranging from -19,169 in 2013 and -16,838 in 2011 to -6,683 in 2008 that age was highly significant.

As with the region (province of origin) predictors in the linear regression model for the Hyundai i10, the predictors in the non-linear model had relatively low values, with the lowest being the Western Cape with a value of 2 744 (t-score: 1,925), North West with a value of 2 919 (t-score: 1,474) and Limpopo with a value of 2 930 (t-score: 1,331). The two provinces with the highest price premia and the most statistically relevant were KwaZulu-Natal with a value of 4 076 (t-score: 3,920) and Mpumalanga with a value of 6 951 (t-score: 3,513).

As with the linear regression model for the Hyundai I10, the 1.1-litre engine had a significant negative effect on the price with a value of -24 083 (t-score: -26,896), compared to the 1.0-litre engine with a value of -19 973 (t-score = -8,278).

This non-linear regression analysis attempted to discern price differences between colours per vehicle relative to the standard white base-case offering. Overall, the significance of colour on Hyundai i10 car prices was very low. In the non-linear regression, only grey, with a value of -

5 932 (t-score: -4,796), indicating an oversupply, and metallic, with a value of 4 048 (t-score: 3,238), indicating a scarcity of supply, were statistically relevant. In other words, only two colours out of six played a role in determining the value of the Hyundai i10.

Table 6.6: Model 6: Toyota Etios – Non-linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	147 609,56	4 715,58	31,303	<0,0001
Yr_2017	-1 037,93	5 035,62	-0,206	0,837
Yr_2016	-12 916,21	5 127,66	-2,519	0,012
Yr_2015	-26 940,86	5 265,82	-5,116	<0,0001
Yr_2014	-43 471,62	5 328,39	-8,158	<0,0001
Yr_2013	-51 629,43	5 302,74	-9,736	<0,0001
Yr_2012	-55 765,33	5 376,29	-10,372	<0,0001
25 000	1 826,61	1 551,74	1,177	0,239
35 000	366,85	1 510,77	0,243	0,808
60 000	2 036,48	1 819,17	1,119	0,263
95 000	238,55	1 960,68	0,122	0,903
Mpumalanga	1 743,27	2 702,77	0,645	0,519
North West	-498,43	2 710,75	-0,184	0,854
KwaZulu-Natal	0,00	0,00	0,00	0,00
Western Cape	-1 148,91	1 957,73	-0,587	0,557
Eastern Cape	-95,19	3 201,65	-0,030	0,976
Free State	-4 000,77	3 227,77	-1,239	0,215
Limpopo	2 642,96	3 237,15	0,816	0,414
Northern Cape	-7 281,77	1 640,34	-4,439	<0,0001
Black	-4 735,58	2 720,76	-1,741	0,082
Blue	11 133,28	2 079,17	5,355	<0,0001
Gold	5 606,01	1 230,34	4,556	<0,0001
Orange	9 147,56	14 405,86	0,635	0,526
Red	4 970,18	1 741,08	2,855	0,004
Silver	4 174,46	1 061,68	3,932	<0,0001
Engine 1.5	-208,34	871,18	-0,239	0,811

The intercept for the non-linear regression model for the Toyota Etios (147 620) was highly significant with a t-score of 31,303. With the introduction of calendar years as dummy-style variables in the non-linear model, it was clear that some years were statistically more significant than others. For example, calendar year 2012 with a t-score of -10,372 was statistically more significant than calendar year 2017 with a t-score of 0,206. Interestingly, mileage seemed to have a significantly lower impact on the price of a Toyota Etios, with very low t-scores ranging from 1,117 for cars with mileage ranging from 0 to 25 000 km to a t-score of 0,243 for cars with mileage ranging from 25 000 km to 35 000 km and a low t-score of 0,122 for cars with mileage of 95 000 km or more. In terms of statistically relevant provinces of origin, only the Northern Cape with a value of -7 281 (t-score: -4,439) seemed to be significant.

Overall, in the non-linear model for the Toyota Etios, the significance of colour on price for the Toyota Etios was very low. Only three colours featured as statistically significant, namely blue with a value of 11 133 (t-score: 5,355), gold with a value of 5 605 (t-score: 4,556) and silver with a value 4 174 (t-score: 3,932). The Toyota Etios carried the most significant colour premia (blue, gold, and silver – all positive).

Table 6.7: Model 7: Ford Fiesta – Non-linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	229 721,00	2 091,64	109,828	<0,0001
Yr_2017	-71 587,95	2 831,20	-25,285	<0,0001
Yr_2016	-77 779,07	3 501,10	-22,216	<0,0001
Yr_2015	-90 019,72	3 710,06	-24,264	<0,0001
Yr_2014	-98 648,94	4 200,77	-23,484	<0,0001
Yr_2013	-101 334,38	4 462,82	-22,706	<0,0001
Yr_2012	-135 258,22	5 159,08	-26,218	<0,0001
Yr_2011	-141 963,37	5 751,32	-24,684	<0,0001
Yr_2010	-141 150,78	5 051,85	-27,940	<0,0001
25 000	5 914,21	2 942,90	2,010	0,045
35 000	7 141,35	2 890,56	2,471	0,014
60 000	595,71	3 188,31	0,187	0,852
95 000	3 761,30	3 888,41	0,967	0,334
North West	8 935,25	4 729,59	1,889	0,059
Eastern Cape	15 986,93	4 951,44	3,229	0,001
Mpumalanga	11 283,77	6 018,80	1,875	0,061
KwaZulu-Natal	8 483,73	2 539,15	3,341	0,001
Free State	20 163,13	8 684,49	2,322	0,020
Northern Cape	24 578,36	10 757,28	2,285	0,023
Western Cape	12 675,69	2 201,21	5,759	<0,0001
Limpopo	18 293,34	6 521,83	2,805	0,005

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Blue	-3 180,87	2 130,20	-1,493	0,136
Black	3 119,40	4 997,82	0,624	0,533
Red	8 008,31	2 343,62	3,417	0,001
Silver	1 943,76	1 797,78	1,081	0,280
Brown	1 790,64	4 206,78	0,426	0,670
Green	11 763,26	6 599,22	1,783	0,075
Automatic	23 413,00	2 726,59	8,587	<0,0001
Engine 1.4	-9 992,04	2 021,14	-4,944	<0,0001
Engine 1.5	21 845,68	5 571,86	3,921	<0,0001
Engine 1.6	23 637,36	3 081,99	7,670	<0,0001
Diesel	-22 612,28	3 914,67	-5,776	<0,0001

The intercept for the Ford Fiesta (229 721) was highly significant with a t-score of 109,828. Age was clearly highly significant, indicating a strong, negative influence on price, with large t-scores ranging from -27,940 (2010) and -26,128 (2012) to -25,285 (2017). It is interesting to note that when the very high t-scores related to age for the Ford Fiesta are compared to the much lower t-scores related to age for the Toyota Etios, the non-linear regression models show empirically that age had a much greater negative effect on the Ford Fiesta than on the Toyota Etios. As with the Toyota Etios, mileage seemed to have a low statistical significance with low t-scores ranging from 2,010 (25 000 km) to 0,967 (95 000 km).

Region (province of origin) was statistically low in significance, with only the Western Cape with a value of 12 676 showing some significance with a t-score of 5,759.

In the non-linear model, the 1.4-litre engine had a much lower t-score of -4,944, while the 1.6-litre model had a higher t-score of 7,670, indicating higher statistical relevance. Red, with a value of 8 008 (t-score: 3,417) was the only statistically useful price predictor.

Table 6.8: Model 8: VW Polo Vivo – Non-linear Regression

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Intercept	162 699,68	1 460,72	111,383	<0,0001
Yr_2017	-8 603,36	2 195,92	-3,918	<0,0001
Yr_2016	-33 371,84	2 499,69	-13,350	<0,0001

Source	Value	Standard error	t-score	p-value (Pr > |t|)
Yr_2015	-53 520,29	2 707,40	-19,768	<0,0001
Yr_2014	-63 515,83	2 824,73	-22,486	<0,0001
Yr_2013	-74 280,26	2 993,96	-24,810	<0,0001
Yr_2012	-74 992,97	3 162,98	-23,710	<0,0001
Yr_2011	-82 272,83	3 166,01	-25,986	<0,0001
Yr_2010	-87 508,49	3 471,86	-25,205	<0,0001
25 000	-6 383,87	2 472,45	-2,582	0,010
35 000	2 925,37	2 180,79	1,341	0,180
60 000	3 143,70	2 332,56	1,348	0,178
95 000	12 906,19	2 560,88	5,040	<0,0001
Mpumalanga	12 674,46	3 549,69	3,571	0,000
North West	15 172,82	2 549,78	5,951	<0,0001
KwaZulu-Natal	12 767,64	1 703,80	7,494	<0,0001
Western Cape	13 811,56	1 568,68	8,805	<0,0001
Eastern Cape	11 645,54	3 095,99	3,761	0,000
Free State	12 642,61	3 809,34	3,319	0,001
Limpopo	20 960,10	4 714,29	4,446	<0,0001
Northern Cape	27 082,77	15 554,83	1,741	0,082
Silver	-1 532,03	1 106,53	-1,385	0,166
Red	-793,05	1 899,80	-0,417	0,676
Beige	5 617,67	3 148,63	1,784	0,075
Gold	-3 206,38	1 897,16	-1,690	0,091
Blue	-2 374,56	2 697,98	-0,880	0,379
Grey	4 665,01	4 951,56	0,942	0,346
Black	4 141,17	3 237,31	1,279	0,201
Yellow	-6 865,49	4 940,99	-1,389	0,165
Other	1 474,58	5 974,27	0,247	0,805
Orange	9 118,78	6 972,83	1,308	0,191
Silver gold	-7 222,86	7 792,30	-0,927	0,354
Cream	-14 904,57	1 5428,30	-0,966	0,334
Navy blue	-5 366,94	8 917,63	-0,602	0,547
Champagne	309,41	7 763,81	0,040	0,968
Automatic	12 096,06	2 255,06	5,364	<0,0001
Engine 1.6	8 341,45	1 332,79	6,259	<0,0001

The intercept for the VW Polo Vivo (162 700) was highly significant with a t-score of 111,383. Age was also highly significant, indicating a strong influence on price, with t-scores ranging from -25,986 (2011) to -3,918 (2017). It is clear from the non-linear analysis that where age initially had a low impact on price, the VW Polo Vivo tended to age more rapidly than its counterparts. Mileage seemed to have a low statistical significance initially with low t-scores ranging from 2,582 (25 000 km) to 1,341 (35 000 km), but with a significant jump in t-score to 5,040 (95 000 km).

Geographically, the province with the biggest statistical impact on price was the Western Cape with a value of 13 812 (t-score: 8,805). Interestingly, compared to the other two brands, the Ford Fiesta and VW Polo Vivo had large and statistically significant price premia in provinces more distant to Gauteng. This may be an indicator that distribution and holding costs were more expensive for these firms compared to their more central competitors.

In the non-linear model, the 1.6-litre engine had a value of 8 341 (t-score: 6,259), indicating higher statistical relevance.

While Petrich (2017) argues that colour plays a significant role in determining the price of a second-hand vehicle, this analysis showed empirically that overall, the statistical significance of colour was low. For example, in discerning the impact on the price of a VW Polo Vivo, out of 13 different colour options, the most statistically relevant colours were beige with a value of 5 618 (t-score: 1,785), gold with a value of -3 206 (t-score: 1,690) and orange with a value of 9 118 (t-score: 1,308).

Across the various colour permutations of the four non-linear models, colour premia appeared to be both positive, indicating scarcity, and negative, indicating an oversupply. Examples of positive price premia are the orange Hyundai i10 with a value of 6 623 (t-score: 1,627); the orange 9 147 (t-score: 0,635), silver 4 174 (t-score: 3,932) and red 4 970 (t-score: 2,855) Toyota Etios; the red Ford Fiesta with a value of 8 008 (t-score: 3,417); and the beige VW Polo Vivo with a value of 5 617 (t-score: 1,784). Examples of negative price premia are the grey Hyundai i10 with a value of -5 932 (t-score: -4,796), the black Toyota Etios with a value of -4 735 (t-score: -1,741), the blue Ford Fiesta with a value of -3 180 (t-score: -1,493) and the gold VW Polo Vivo with the value of -3 206 (t-score: -1,690).

6.3 Summary of Key Statistical Indicators (Age, Mileage and Adjusted R^2)

For presentation, comparison and discussion purposes, Table 6.9 sets out a summary of the key statistical indicators.

Table 6.9: Key Statistical Indicators (Age, Mileage and Adjusted R^2)

Regression type	Variable	Model			
		Hyundai i10	Toyota Etios	Ford Fiesta	VW Polo Vivo
Linear	Constant	171 110	159 97	220 111	163 720
	Age	-7 326	-11 279	-13 363	-10 557
	Mileage	-0,15	-0,04	-0,09	0,003
	R-squared	0,77	0,70	0,72	0,70
	Adjusted R-squared	0,77	0,70	0,71	0,69
Non-linear	Constant	170 644	147 610	229 721	162 700
	Age 2017	-16 168	-1 038	-71 588	-8 603
	Age 2016	-22 773	-12 916	-77 779	-33 372
	Age 2015	-33 788	-26 941	-90 020	-53 520
	Age 2014	-43 873	-43 472	-98 649	-63 516
	Age 2013	-53 768	-51 629	-101 334	-74 280
	Mileage (25 000 km)	-212	1 552	-5 914	-6 384
	Mileage (35 000 km)	-3 520	1 511	-7 141	2 925
	Mileage (60 000 km)	-8 708	1 819	-596	3 144
	Mileage (95 000 km)	-12 953	1 961	-3 761	12 906
	R-squared	0,79	0,71	0,82	0,77
	Adjusted R-squared	0,78	0,71	0,81	0,77

In the regression analysis, the constant (also known as the 'intercept') of each model is noteworthy, as it represents the price (dependent variable) of the respective vehicle with base-case controls (independent variables) applied, in other words a brand-new manual, petrol-driven white vehicle, purchased in Gauteng, with zero miles on the clock. The primary purpose of the constant is to ensure that the models account for the intercept or baseline value.

From the table above it is clear from the four linear constants that the Ford Fiesta was empirically the most expensive at R220 111, compared to the Hyundai i10 (R171 110), the VW Polo Vivo (R163 720) and the Toyota Etios (R159 972). With the four non-linear constants, the Ford Fiesta was again empirically the most expensive at R229 721, followed by the Hyundai i10 (R170 644), the VW Polo Vivo (R162 700) and the Toyota Etios (R147 610). It is important to note that interpreting these constants may not always yield a meaningful real-world interpretation, especially when the independent variables are not realistically interpretable at zero.

For the Hyundai i10, Toyota Etios, Ford Fiesta and VW Polo Vivo, the respective age coefficients were -7 326 (t-score: -21,560), -11 279 (t-score: -31,656), -13 363 (t-score: -20,241) and -10 557 (t-score: -31,940). To compute the value of an aging vehicle, one needs merely to subtract the age depreciation component of each vehicle from its baseline value. Therefore, for a one-year-old (2017) vehicle, the computed values were R163 784 (Hyundai i10), R148 693 (Toyota Etios), R206 748 (Ford Fiesta) and R153 163 (VW Polo Vivo). Based on the attributed t-scores (in parenthesis above), the most significant factor driving price was age. No other independent variable came close to these levels of statistical significance. From the analysis, one can also deduce that some vehicles lost their value faster than others. In terms of value, the half-life for each vehicle was 11,6 years (Hyundai i10), 7,09 years (Toyota Etios), 8,2 years (Ford Fiesta) and 7,8 (VW Polo Vivo) years. As a basis of value retention, and therefore future resale, these figures indicate that the Hyundai i10 will retain its value the longest (11,2 years) and the VW Polo Vivo the shortest (7,08 years). These insights would not be evident without this analysis.

For the Hyundai i10, Toyota Etios, Ford Fiesta and VW Polo Vivo, the mileage coefficients were -0,15 (t-score: -8,692), -0,04 (t-score: -2,889), -0,09 (t-score: -2,557) and 0,003 (t-score: 0,180), respectively. Although the Hyundai i10 coefficient was highly significant, the VW Polo Vivo coefficient was close to zero and insignificant. After 100 000 km, depreciation on the Hyundai i10, Toyota Etios and Ford Fiesta would be R15 000, R4 000 and R9 000, respectively. The VW Polo Vivo would increase by R3. This regression analysis showed that compared to age, mileage had a less significant role in determining the price of the vehicle. One of the reasons for this could be that these entry-level vehicles are not purchased for driving long distances, but rather as short distance ‘run-arounds’.

The adjusted R-squared values measure the goodness of fit of the linear and non-linear regression models. They indicate the proportion of the variance in the dependent variable (price) that is predictable from the independent variables (CD). Higher values indicate a better fit. For the four linear regression models, the adjusted R-squared values ranged from 0,69 to 0,77, generally indicating a good to strong fit of the models. For the non-linear regression, the adjusted R-squared values ranged from 0,71 to 0,81, generally indicating a good to strong fit of the models.

In the case of the Hyundai i10, the adjusted R-squared value of the linear approach was 0,77 compared to the 0,78 of the non-linear approach. In the case of the Toyota Etios, the adjusted R-squared value of the linear approach was 0,70 compared to the 0,71 of the non-linear approach. In the case of the Ford Fiesta, the adjusted R-squared value of the linear approach was 0,71 compared to the 0,81 of the non-linear approach. In the case of the VW Polo Vivo, the adjusted R-squared value of the linear approach is 0,69 compared to the 0,77 of the non-linear approach. Having compared the two sets of models, the non-linear models consistently show higher adjusted R-squared values, indicating a better fit compared to the linear models. Statistically, the inclusion of the dummy variables made the non-linear models more accurate and in so doing provides customers with a more useful tool for identifying value for money in a second-hand vehicle.

Where the impact of age and mileage on price varied across models and categories, as expected, age and mileage had negative coefficients, suggesting a negative impact on vehicle price as they increased.

6.4 Conclusion

In regression analysis, the constant (intercept) represents the baseline value of the vehicle with base-case controls. From the empirical results, the Ford Fiesta appeared to be the most expensive among the four linear models, both in terms of linear (R220 111) and non-linear (R229 721) constants. The age coefficients for the Hyundai i10, Toyota Etios, Ford Fiesta and

VW Polo Vivo indicated that age has a significant impact on price. The half-life analysis revealed differing rates of depreciation, with the Hyundai i10 retaining value the longest (11,6 years) and the VW Polo Vivo the shortest (7,8 years). This information is useful for assessing future resale values.

While mileage coefficients were significant for the Hyundai i10, Toyota Etios and Ford Fiesta, the statistical impact was significantly less compared to that of age. Mileage had a smaller role in determining price, possibly because these entry-level vehicles are often used for short distances. Adjusted R-squared values indicated the goodness of fit for the linear and non-linear models. The non-linear models consistently exhibited higher values, suggesting better accuracy. The inclusion of dummy variables enhances non-linear models, providing a more useful tool for identifying value for money in second-hand cars.

The Hyundai i10 showed significant regional price differences, with Western Cape and KwaZulu-Natal showing significant price premia. Engine size had an impact on price, with 1.1-litre engines having a significant negative effect. In the case of the Toyota Etios, region (province of origin) influenced price, with provinces more distant to Gauteng showing significant price premia. Diesel engines in entry-level vehicles acted as a significant price differentiator. For the Ford Fiesta, region (province of origin), engine size and fuel type significantly affected prices. For the VW Polo Vivo, region (province of origin) influenced price, with the Western Cape and KwaZulu-Natal showing significant price premia. Engine size (1.6 litres) was a significant price predictor.

Regional predictors for the Hyundai i10 had relatively low values, with KwaZulu-Natal and Mpumalanga showing the most statistically relevant price premia. The impact of colour on price was low, with only grey and metallic colours being statistically relevant. For the Toyota Etios, calendar years remained statistically the most significant, while motor vehicle colours had varying levels of statistical significance. Blue, gold, silver and red were statistically significant colour premia. Age had a greater negative effect on the Ford Fiesta compared to the Toyota Etios in non-linear models. For the Ford Fiesta, red was the only statistically relevant colour predictor. For the VW Polo Vivo, age and mileage had an impact on price, with the Western Cape having the most significant geographical influence. Colour premia varied, with beige, gold and orange showing statistical relevance.

The findings provide a comprehensive understanding of the factors influencing the pricing of entry-level vehicles. Age remained a dominant factor, with varying depreciation rates. Non-linear models consistently outperformed linear models, offering improved accuracy. Region (province of origin), engine size, fuel type and transmission type had a statistical impact on price, while colour generally had a low impact across the models. These insights contribute to a more nuanced understanding of the second-hand car market, aiding both buyers and sellers in making more informed decisions.

Chapter Seven will set out conclusions and recommendations.

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS

"Virtually all goods are heterogeneous"

(Greenstone, 2017)

7.1 Introduction

The exploration of pricing dynamics in the automotive market necessitates a sophisticated analytical approach that captures both conventional and nuanced relationships between vehicle attributes and prices. This study compared linear and non-linear modelling frameworks to provide a comprehensive examination of the determinants of car pricing, offering insights that extend beyond traditional methodologies. By incorporating both established linear models and more complex non-linear models enriched with variables such as dummy indicators for registration years and mileage ranges, this analysis delved into the temporal and segmented impacts of key factors such as age and mileage on vehicle prices. This dual approach not only enhanced the explanatory power of the models, as evidenced by higher adjusted R-squared values in non-linear frameworks, but also highlights the practical implications for market stakeholders. This chapter sets the stage for understanding the critical role of model selection in uncovering pricing patterns, underscoring its relevance to strategic decision making for managers, policymakers and market analysts in the automotive industry.

The inclusion of linear and non-linear models, coupled with adjusted R-squared values, reflects a comprehensive approach. The linear models encapsulate the conventional understanding of how age and mileage influence car prices. The introduction of dummy variables for different years of registration and mileage ranges in the non-linear models enhanced the complexity of the analysis by reflecting an acknowledgment of the potential non-linearity in the relationship between certain variables and car prices. For example, the non-linear model captures the temporal dynamics by considering the specific year of registration, recognising that the impact of age on vehicle prices may vary across different years. Similarly, the segmentation of mileage into ranges acknowledges the non-uniform effect of mileage on pricing. Where the use of adjusted R-squared values serves as a diagnostic measure for model fit, the higher adjusted R-squared values in non-linear models suggest that the inclusion of additional variables provides a more comprehensive explanation of the variability in car prices. Understanding the non-linear nuances in pricing could inform targeted interventions or marketing strategies tailored to specific vehicle characteristics. These findings hold implications for managers, policymakers, market analysts and other industry stakeholders. This final chapter is organised as follows: Section 7.2 presents a summary of the findings. Section 7.3 sets out managerial implications, section 7.4 discusses the limitations of the study, section 7.5 presents potential avenues of future research, and section 7.6 concludes.

7.2 Summary of Findings

The valuation of the characteristics of a product is a fundamental issue in the development and commercialisation of products and services (Allenby, Brazell & Howell, 2014). The regression results from the models above provide a robust statistical framework for understanding key determinants of second-hand vehicle prices, particularly how age and mileage affect value. These findings hold implications not only for consumers, but also for many stakeholders in the automobile industry. With major global brands such as Toyota South Africa Motors, VWSA, Ford Motor Company of Southern Africa, Mercedes-Benz South Africa and Nissan South Africa all driving growth, competitiveness and innovation in the South African automotive sector, understanding consumer decision-making processes is crucial for marketing and sales managers of these automobile manufacturers. Other key stakeholders, including component manufacturers and suppliers (e.g. Bosch, Sumitomo and Metair), government regulatory bodies such as the Department of Trade, Industry and Competition, industry associations such as Naamsa and the Retail Motor Industry Organisation, as well as leading automotive insurers such as Old Mutual, Santam and Budget Insurance, should maintain a vested interest in the development and application of hedonic pricing models. These models provide a more precise and reliable framework for determining motor vehicle pricing, offering critical insights that can enhance decision making and operational strategies across the automotive value chain.

In this study, an examination of constants provided baseline values, notably identifying the Ford Fiesta as empirically the most expensive among the linear models. Non-linear models reinforced this observation, underlining the importance of establishing baseline values for precise price predictions. It has become clear that understanding the non-linear nuances in pricing could inform targeted production interventions or marketing strategies tailored to specific vehicle characteristics. For example, of the 13 VW Polo Vivo colours analysed, only three showed some statistical relevance, namely gold with a value of -7 234 (t-score: -3,39), orange with a value of 16 203 (t-score: 2,05) and beige with a value of 9 463 (t-score: 2,65). It is clear that production managers should reduce the colour range significantly and in so doing save both economic and environmental costs.

Age and mileage emerged as dominant determinants in second-hand vehicle pricing, with varying depreciation rates across the four models. The negative coefficients for the two price explanatory variables aligned theoretical expectations with conventional understanding and proved empirically that depreciation in price is logical as age and mileage increase. What is more interesting is the extent to which age and mileage affect the prices of the different car models. For example, if one uses the linear model to compare the impact of a one-year increase in the age of a Hyundai, the depreciation in price is only R-7 326, as opposed to the Toyota (R-11 280), the Ford Fiesta (R-13 362) and the VW Polo Vivo (-R10 556). The half-life analysis revealed that the Hyundai i10 demonstrated the longest retention of value (11,6 years), with the VW Polo Vivo exhibiting the shortest half-life (7,8 years).

While mileage played a secondary role compared to age in determining prices, it still held significance. A third factor, namely region (province of origin), revealed pricing disparities in provinces of origin, offering insights for both consumers and sellers into value for money in different areas. The importance of the region (province of origin) variable suggests opportunities for regional arbitrage, indicating that consumers may find it cost-effective to purchase a car in a different region and transport it to their location.

The inclusion of linear and non-linear models, coupled with adjusted R-squared values, reflects a comprehensive approach, and a comparison of adjusted R-squared values highlighted the superiority of non-linear models in accuracy. It was clear that the inclusion of dummy variables for different years of registration and mileage ranges in the non-linear models enhanced the complexity of the analysis and model performance, providing a more useful tool for identifying value for money in the second-hand vehicle market. However, if one were to consider the audience or target market, and if the goal is to communicate findings to a non-technical audience, the linear model might be more straightforward and therefore more useful.

7.3 Managerial Implications

The dynamic nature of the second-hand car market presents unique challenges and opportunities for managers across various sectors, including marketing, insurance and online retail. This section explores the practical implications, for managers, of the research findings, providing a foundation for data-driven decision making that aligns with contemporary market demands. By addressing these managerial implications, this chapter bridges the gap between theoretical insights and practical applications, empowering industry stakeholders to make informed, strategic decisions in an increasingly competitive landscape.

7.3.1 Marketing Strategies

For marketing managers in the second-hand car market, understanding the key drivers of vehicle pricing is essential to designing effective marketing campaigns and maximising market competitiveness. By incorporating advanced tools such as hedonic pricing models and non-linear regression analysis, marketing managers can enhance their ability to accurately segment and price vehicles, ultimately aligning with consumer valuation patterns. These data-driven approaches not only improve pricing precision, but also enable the development of campaigns that resonate with target audiences, emphasising transparency and value for money. The following section equips marketing managers with actionable strategies to capitalise on market dynamics, refine promotional messaging and drive customer engagement in the highly competitive second-hand car industry.

7.3.1.1 Age as a primary determinant in pricing strategy

For marketing managers, vehicle age emerges as a pivotal determinant in pricing strategies, particularly in contexts characterised by high price elasticity. It is imperative that promotional campaigns emphasise age as a key variable, ensuring that price points reflect its significance.

By doing so, marketing efforts can effectively align with consumer valuation models in competitive markets.

7.3.1.2 Mileage: A secondary but significant factor

While mileage remains a salient consideration, its impact is secondary to that of age. Marketing strategies should incorporate mileage as a value-enhancing attribute, particularly by assigning premiums to vehicles with low mileage. For consumers, this underscores the importance of investing time and effort to identify vehicles with favourable mileage metrics. In cases where a vehicle's age is relatively low but its mileage is high, the vehicle's age should be the predominant factor influencing purchasing decisions.

7.3.1.3 Regional (province of origin) preferences

If regional preferences or pricing variations exist, as the analysis clearly shows, marketing managers can create localised campaigns. For instance, certain regions may prefer vehicles with higher durability due to rougher terrain or specific climate conditions, influencing how age, mileage and region (province of origin) are marketed.

7.3.2 Insurance premiums

A critical insight for managers of insurance companies lies in the empirical identification of three statistically significant variables that affect the valuation of second-hand vehicles. Integrating these variables into actuarial models offers a strategic opportunity to refine the precision of insurance premium calculations. By aligning premiums more closely with market-relevant risk factors, insurers can enhance the accuracy and fairness of their pricing structures, ultimately improving competitiveness and consumer trust in their offerings.

7.3.3 Online retailers

Managers overseeing major online platforms for second-hand vehicle sales (e.g. WeBuyCars, AutoTrader and cars.co.za) should prioritise the development of innovative, customer-focused applications that seamlessly incorporate hedonic pricing models. These applications should leverage the advanced capabilities of big data analytics and AI in a robust platform-driven business framework. By doing so, they can revolutionise the second-hand car market, enabling more accurate price valuations, enhancing customer engagement and establishing a competitive edge in a rapidly evolving digital marketplace.

7.3.4 Production managers

It is clear from Chapter Two that the automotive sector has a big role to play in the economic and manufacturing capacities of many emerging market countries in Africa. The non-linear model developed in this study could assist new entrants into the African automotive market by guiding production managers focused on product efficiencies as to exactly which car attributes they should be focusing on for their new market. The relatively low statistical significance of four out of eight variables, including colour, fuel type, transmission type and engine size, should highlight to managers that they have a marginal influence on entry-level car pricing. This insensitivity may be attributable to the perception of entry-level vehicles as being optimised for urban commuting. Specifically, these cars are often viewed as suitable for short-distance travel, with features tailored to convenience and fuel efficiency rather than performance or aesthetic preferences. This insight suggests that not only marketing efforts but also production for entry-level vehicles should deprioritise these attributes in favour of more impactful factors such as age and mileage. Given that less than 1,6% of vehicles in the entry-level category are powered by diesel engines, marketing managers should acknowledge that fuel-type premia hold negligible relevance in this segment.

7.4 Research Limitations

Customers have been noted to struggle in their assessments of what constitutes fair value for money when that value is derived from characteristics that cannot be independently priced or even tested before consumption (Balcilar, Uwilingiye & Gupta, 2018). Although this study offers a robust analysis and useful insights into the statistical significance of key extrinsic variables, certain limitations persist. Foremost among these is the non-consumer-facing orientation of the predictive analytical models utilised, which may fail to account for additional determinants of purchasing behaviour, such as demographic characteristics (e.g. socioeconomic status, purchasing power and disposable income) and brand loyalty. To address these shortcomings, future studies should adopt a more holistic approach to examining the second-hand car market, integrating qualitative methodologies, such as consumer interviews, to elucidate the extrinsic factors that buyers perceive as most influential in shaping their purchasing decisions.

Currently there appears to be no literature that explores the impact of brand loyalty or perceived brand value across second-hand car markets. An additional avenue of potential research might be an investigation into why certain customers remain brand loyal, even at the expense of being price inelastic.

Brand loyalty is a multidimensional construct shaped by the dynamic interplay of various interrelated factors. These factors can be classified into psychological, behavioural, emotional and contextual domains, which collectively drive and sustain a customer's allegiance to a brand. The psychological mechanisms underpinning consumer brand loyalty warrant a deeper examination, particularly in terms of the conditions under which such loyalty may diminish or dissolve. To elucidate these dynamics, one could analyse the fundamental psychological attributes that initially foster brand loyalty, or alternatively, conduct a comparative investigation into the hedonic pricing mechanisms across multiple brand effects.

The conceptualisation of a robust brand analytics framework that can identify, quantify and rank non-conscious brand effects holds significant promise. Such a model would empower production and marketing managers in the automotive sector to gain a more nuanced understanding of the factors that constitute brand loyalty in vehicles. Moreover, it would facilitate the identification of these factors across varying price segments and market conditions, thereby enhancing strategic decision making in the context of brand management.

7.5 Opportunities for Further Research

Future research could extend the scope of this study by broadening its analytical framework to include a wider range of market segments and/or vehicle categories. While the current focus on linear and non-linear hedonic pricing for four entry-level vehicles in the South African market provides valuable insights into the pricing dynamics of a specific segment, several avenues exist for future exploration. Future studies could incorporate a more diverse set of consumer demographics and economic strata, considering not only entry-level vehicles but also mid-range and luxury vehicles. This would allow for a comparative analysis of how hedonic pricing influences brand loyalty and consumer preferences across different income groups and purchasing power profiles.

A comparative analysis of hedonic pricing could be extended beyond the South African market to include other emerging markets or developed economies. By incorporating cross-national perspectives, the research would help identify whether the psychological drivers of pricing and loyalty are universal or culturally contingent, thereby contributing to a more global understanding of hedonic pricing. Future research could also explore the impact of hedonic pricing across various vehicle categories, such as electric vehicles, hybrid models and SUVs. These categories are likely to exhibit different consumer valuations based on their unique attributes (e.g. environmental impact, fuel efficiency and size), which may require distinct pricing models.

By examining changes in consumer preferences and pricing behaviour over time, a longitudinal study could provide insights into how hedonic pricing evolves in response to market trends, technological innovations and shifts in consumer values. Such a longitudinal approach would also shed light on the temporal aspects of brand loyalty and its susceptibility to external factors, such as economic downturns or shifts in regulatory policies. Discussing the role of consumer search behaviour, signaling mechanisms, and bounded rationality in price determination would enrich the theoretical contributions. There is an opportunity to explore behavioural economics insights—such as loss aversion or anchoring effects—that may influence pricing in the used-car market.

Finally, further exploration of non-linear pricing could incorporate more advanced econometric or machine learning models to better capture the complexities of consumer behaviour and market dynamics. An econometric study, which adheres to the strict norms and expectations of econometrics, which demand a high degree of mathematical rigor, explicit treatment of endogeneity, omitted variable bias, and robust model specification tests, could help identify non-obvious relationships between vehicle attributes and consumer willingness to pay, providing a more granular understanding of hedonic pricing effects.

The confluence of linear and non-linear hedonic pricing, combined with advancements in big data, AI and the platform-based business model, holds significant potential for transforming the automotive industry, particularly in the context of second-hand cars. By integrating linear and non-linear hedonic pricing with big data analytics and AI, researchers can achieve a far more precise understanding of the factors influencing second-hand car prices. AI-driven algorithms can process large data sets to uncover complex, non-linear relationships between vehicle attributes and market demand, providing more accurate price forecasts. This could lead to dynamic pricing models that better reflect consumer behaviour and regional market variations, enabling car dealers and buyers to make more informed decisions.

Platform-based business models, such as those employed by online marketplaces, could benefit greatly from AI-driven insights into consumer preferences and purchasing patterns and how these relate to hedonic pricing. By leveraging big data, companies can offer hyper-personalised recommendations to consumers based on their search histories, demographics and previous interactions with the platform. This personalisation could extend to pricing, where dynamic, AI-powered pricing models can adjust based on individual consumer profiles and perceived willingness to pay, enhancing both the user experience and the conversion rates for sellers.

For second-hand car dealerships operating in a platform-based ecosystem, combining hedonic pricing with big data and AI could significantly improve inventory management. By analysing real-time data on consumer preferences, market trends and competitor prices, dealerships could optimise their inventory to ensure that they are stocking vehicles that are in high demand. In addition, AI could predict which cars are likely to appreciate or depreciate in value, enabling dealerships to make better investment decisions and enhance profitability.

The combination of big data and AI can help establish greater transparency in the second-hand car market by providing objective, data-driven pricing insights. With platforms offering real-time access to pricing trends, vehicle histories and condition assessments, consumers could gain more confidence in their purchase decisions. Furthermore, AI can aid in detecting price manipulation or discrepancies, helping to regulate and ensure fairness in the pricing of used cars, which is particularly important in regions where trust in used vehicle markets is often low.

By utilising big data and AI to understand how different consumer segments value various vehicle attributes, researchers could identify previously unrecognised market segments. This could lead to the development of new platform-based business models, such as subscription services or pay-per-use options for second-hand cars, where pricing models are tailored based on the specific needs and preferences of individual segments. Such models could cater to shifting consumer trends, such as the growing preference for mobility-as-a-service, offering new opportunities for businesses to innovate in the second-hand car market.

A PhD-level analysis prompts considerations for future research. For example, a PhD study could investigate the underlying factors contributing to non-linear relationships in hedonic pricing, particularly in the context of second-hand cars. This would involve a comprehensive examination of how various vehicle attributes influence consumer perceptions and willingness to pay in ways that cannot be captured by traditional linear models. Identifying the psychological, economic and contextual drivers of these non-linearities could contribute new theoretical insights into pricing theory.

Future research could focus on exploring the interactions between various explanatory variables, such as vehicle characteristics, market conditions and consumer behaviour. Such a study could utilise advanced statistical or machine learning techniques to model complex interactions that may influence pricing dynamics in non-obvious ways. For instance, how might the combined effect of brand and age differ across market segments? Understanding these interactions could lead to more nuanced pricing models that account for multidimensional factors.

A further opportunity for a PhD study could be to examine the applicability of hedonic pricing models across different market segments, such as entry-level versus luxury cars, or in varying geographic and demographic contexts. This would involve testing whether the same pricing dynamics hold in different market conditions, such as in emerging markets versus developed economies, or among different income groups. Insights from such a study could help refine pricing strategies for firms operating in diverse markets such as in South Africa.

A key aspect of PhD-level research could be a detailed analysis of the methodological challenges inherent in hedonic pricing studies, particularly in addressing issues of endogeneity (where explanatory variables are correlated with the error term), multicollinearity (where explanatory variables are highly correlated with each other) and model selection criteria. A PhD study could explore advanced econometric techniques, such as instrumental variable methods, robust regression models or machine learning-based approaches, to resolve these issues and ensure the validity of the findings.

Finally, a holistic approach to a PhD study could involve contextualising the research findings in the broader literature on pricing theory, consumer behaviour and brand loyalty. This would require an in-depth review of existing theoretical frameworks, identifying gaps in the literature and positioning the research as a contribution to advancing the understanding of hedonic pricing in second-hand markets. In addition, exploring the implications of these findings for managerial practice, policy and future research would help bridge the gap between theory and practice.

7.6 Conclusion

In conclusion, this study highlights the multi-faceted factors influencing second-hand vehicle pricing, combining statistical rigor and practical insights for stakeholders across the automotive industry. The findings emphasise the importance of non-linear models in capturing complex market dynamics, providing a nuanced understanding of how variables such as age, mileage and region (province of origin) shape vehicle valuation. These results offer actionable guidance to manufacturers, marketers, insurers and policymakers, paving the way for more efficient production strategies, tailored marketing efforts and informed decision making in both pricing and purchasing practices. By bridging technical depth with practical applications, the study contributes to the evolving landscape of second-hand vehicle markets and offers tools for enhanced stakeholder engagement. The insights derived from this analysis underscore the critical role that nuanced and data-driven strategies play in managerial decision making across multiple facets of the automotive industry. For marketing managers, the primacy of vehicle age and the secondary yet significant role of mileage necessitate refined promotional and pricing strategies that align with consumer valuation frameworks.

The findings further advocate for insurance companies to integrate statistically significant variables into actuarial models, enhancing the alignment between insurance premiums and market dynamics. In the realm of online retail, the adoption of advanced analytics and platform-driven frameworks offers transformative potential, enabling customer-centric solutions that can redefine the second-hand vehicle market. For production managers, the implications extend beyond mere pricing considerations, emphasising the need to prioritise impactful attributes such as age and mileage, while de-emphasising factors with limited relevance, particularly for entry-level vehicles optimised for urban use. Collectively, these managerial implications offer a robust foundation for optimising resource allocation, improving operational efficiencies and fostering competitive advantages in a rapidly evolving automotive market. By leveraging these insights, stakeholders can align their strategies with empirical evidence, ensuring sustained relevance and effectiveness in addressing consumer needs and market realities.

Finally, the proposed opportunities for future research underscore the vast potential for advancing theoretical and practical knowledge in the domain of hedonic pricing, particularly in the context of second-hand vehicles. The exploration of non-linear relationships in pricing dynamics offers an opportunity to deepen our understanding of consumer valuation processes by integrating psychological, economic and contextual dimensions that remain underexplored. By delving into the interactions between explanatory variables, such as vehicle attributes and market conditions, future research can contribute to the development of more sophisticated pricing models that account for multidimensional and interdependent factors.

Expanding the applicability of hedonic pricing models to diverse market segments, including entry-level versus luxury vehicles and varying geographic and demographic contexts, could generate valuable insights for firms operating across heterogeneous markets. Such investigations would also inform the adaptation of pricing strategies in emerging versus developed economies, enhancing the relevance and utility of pricing models across different economic landscapes. Addressing methodological challenges, such as endogeneity, multicollinearity and model selection, is essential for ensuring the robustness and validity of hedonic pricing studies. Employing advanced econometric techniques and machine learning approaches could not only resolve these issues, but also push the boundaries of methodological innovation in pricing research. Finally, integrating findings into the broader theoretical frameworks of pricing theory, consumer behaviour and brand loyalty would position future studies as critical contributions to the academic discourse. By bridging theoretical insights with managerial and policy implications, PhD-level research in this field can offer actionable guidance for stakeholders while simultaneously advancing the frontiers of pricing theory. This holistic approach would ensure that the research is not only methodologically rigorous, but also impactful in terms of shaping industry practices and informing policy decisions.

REFERENCE LIST.

Abrate, G., Capriello, A. & Fraquelli, G. 2011. When quality signals talk: evidence from the Turin hotel industry. *Tourism Management.* 32(4):912-921.

Akerlof, G.A. 1970. The market for “lemons”: quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*. 84(3):488-500.

Akoyomare, G., Balcilar, M., Uwilingiye, J. & Gupta, R. 2018. Dynamic relationship between oil price and inflation in South Africa. *The Journal of Developing Areas*. 52(2):73-93.

Allenby, G.M., Brazell, J.D. & Howell, J.R. 2014. Economic valuation of product features. *Quantitative Marketing and Economics*. 12:421-456. DOI: https://doi.org/10.1007/s11129-014-9150-x.

Andersson, H. 2005. The value of safety as revealed in the Swedish car market: an application of the hedonic pricing approach. *The Journal of Risk and Uncertainty.* 30(3):211-239.

Andrews, T. & Benzing, C. 2007. The determinants of price in internet auctions of used cars. *Atlantic Economic Journal*. 35(1):43-57. DOI: https://doi.org/10.1007/s11293-006-9045-7.

Arora, S.D. & Mathur, S. 2020. Hotel pricing at tourist destinations: a comparison across emerging and developed markets. *Journal Tourism Management Prospective*. 35(3):Article 100724. DOI: https://doi.org/10.1016/j.tmp.2020.100724.

Arrondo, R., Garcia, N. & González, E. 2018. Estimating product efficiency through a hedonic pricing best practice frontier. *Business Research Quarterly*. 21(4):215-224. DOI: https://doi.org/10.1016/j.brq.2018.08.005.

Atkinson, S.E. & Halvorsen, R. 1984. A new hedonic technique for estimating attribute demand: an application to the demand for automobile fuel efficiency. *The Review of Economics and Statistics.* 66(3):417-426. DOI: https://doi.org/10.2307/1924995.

Atkinson, S.E. & Halvorsen, R. 1990. The valuation of risks to life: evidence from the market for automobiles. *The Review of Economics and Statistics.* 72(1):133-136. DOI: https://doi.org/10.2307/2109769.

Balcilar, M., Uwilingiye, J. & Gupta, R. 2018. Dynamic relationship between oil price and inflation in South Africa. *Journal of Developing Areas.* 52(2):73-93. DOI: https://doi.org/10.1353/jda.2018.0023.

Baltagi, B.H. 2002. *Econometrics*. 3rd ed. Berlin: Springer-Verlag.

Bauer, I., Zavolokina, L. & Schwabe, G. 2019. Is there a market for trusted car data? *Electronic Markets*. 30:211-225. DOI: https://doi.org/10.1007/s12525-019-00368-5.

Bayer, P., Fernando, F. & McMillan, R. 2007. A unified framework for measuring preferences for schools and neighbourhoods. *Journal of Political Economy*. 115(4):588-638. DOI: https://doi.org/10.1086/522381.

Berndt, E.R. & Rappaport, N. 2001. Price and quality of desktop and mobile personal computers: a quarter-century historical overview. *American Economic Review. 91*(2):268-273.

Beron, K.J., Hanson, Y., Murdoch, J.C. & Thayer, M.A. 2004. Hedonic price functions and spatial dependence: implications for the demand for urban air quality. In *Advances in spatial econometrics.* L. Anselin, R.J.G.M. Florax & S.J. Rey, Eds. Berlin: Springer. 267-281. DOI: https://doi.org/10.1007/978-3-662-05617-2_12.

Betts, S.C. & Taran, Z. 2006. A test of prospect theory in the used car market: the non-linear effects of age and reliability on price. *Academy of Marketing Studies Journal*. 10:57-75.

Black, S.E. 1999. Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics*. 114(2):577-599. DOI: https://doi.org/10.1162/003355399556070.

Bond, E. 1984. Test of the lemons model: reply. *The American Economic Review*. 74(4):801-804.

Brand South Africa. 2021. *Automobile industry growing the South African economy*. Available: https://www.brandsouthafrica.com/investmentsimmigration/economynews/automobile-industry-growing-the-south-african-economy [March 22, 2021].

Charnes, A., Cooper, W. & Rhodes, E. 1978. Measuring the efficiency of decision-making units. *European Journal of Operational Research*. 2(6):429-444.

Chen, C. & Rothschild, R. 2010. An application of hedonic pricing analysis to the case of hotel rooms in Taipei. *Tourism Economics*. 16(3):685-694. DOI: https://doi.org/10.5367/000000010792278310.

Chen, Y. & Xie, K.L. 2017. Consumer valuation of Airbnb listings: a hedonic pricing approach. *International Journal of Contemporary Hospitality Management*. 29(9):2405-2424. DOI: https://doi.org/10.1108/IJCHM-10-2016-0606.

Chin, T.L. & Chau, K.W. 2003. A critical review of literature on the hedonic price model. *International Journal for Housing and its Applications*. 27(2):145-165.

Chowdhury, M., Salam, K. & Tay, R. 2016. Consumer preferences and policy implications for the green car market. *Marketing Intelligence & Planning*. 34:810-827. DOI: https://doi.org/10.1108/MIP-08-2015-0167.

Court, A.T. 1939. Hedonic price indexes with automotive examples. In *The dynamics of automobile demand: based upon papers presented at a joint meeting of the American*

Statistical Association and the Econometric Society, in Detroit, Michigan, on December 27, 1938, Wydania 1-5. New York: General Motors: 98-117.

Deloitte. 2021. *2021 global automotive consumer study: Global focus countries*. Available: https://www2.deloitte.com/global/en/pages/consumer-business/articles/global-automotive-trends-millennials-consumer-study.html [November 27, 2021].

Dreyfus, M.K. & Viscusi, W.K. 1995. Rates of time preference and consumer valuations of automobile safety and fuel efficiency. *Journal of Law and Economics*. 38(1):79-105.

Du, Q., Wu, C., Ye, X., Ren, F., & Lin, Y. 2018. Evaluating the effects of landscape on housing prices in urban China. *Tijdschrift voor Economische en Sociale Geografie /Journal of Economic and Human Geography*. 109(4):525-541 DOI: https://doi.org/10.1111/tesg.12308

Edmonds, R.G. 1984. A theoretical basis for hedonic regression: a research primer. *Real Estate Economics. 12*(1):1-21. DOI: https://doi.org/10.1111/1540-6229.00311.

Ellencweig, B., Ezratty, S., Fleming, D. & Miller, I. 2019. *Used cars, new platforms: Accelerating sales in a digitally disrupted market*. McKinsey. Available: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/used-cars-new-platforms-accelerating-sales-in-a-digitally-disrupted-market [2021, November 27].

Erdem, C. & Şentürk, I. 2009. Hedonic analysis of used car prices in Turkey. *International Journal of Economic Perspectives*. 3(2):141-149.

Espinet, J., Saez, M., Coenders, G. & Fluvià, M. 2003. Effect on prices of the attributes of holiday hotels: a hedonic prices approach. *Journal of Tourism Economics*. 9(2):165-177.

Farrar, D. & Glauber, R. 1967. Multicollinearity in regression analysis: the problem revisited. *The Review of Economics and Statistics*. 49(1):92-107. DOI: https://doi.org/10.2307/1937887.

Ford invests $1 billion (R15.8) to modernize, expand South African manufacturing for all-new Ranger: adds 1,200 jobs. 2021. *Ford South Africa*. 2 February. Available: https://www.ford.co.za/about-ford/newsroom/2021/ford-sa/71883/ [September 19, 2021].

Fried, H.O. & Tauer, L.W. 2015. An entrepreneur performance index. *Journal of Productivity Analysis. 44*(1):29-38. DOI: https://doi.org/10.1007/s11123-015-0436-0.

Gongqi, S., Yansong, W. & Qiang, Z. 2011. New model for residual value prediction of the used car based on BP neural network and nonlinear curve fit. In *Third International Conference on Measuring Technology and Mechatronics Automation International Conferenc*e, 6-7 January, 2011, Shangshai, China. Vol. 2, 682-685. DOI: https://doi.org/10.1109/ICMTMA.2011.455.

Gonzalez, E., Arrondo, R. & Carcaba, A. 2017. Product innovation in the Spanish auto market: frontier shift and catching-up effects. *Transportation Research Part D: Transport and Environment*. 50:170-181. DOI: https://doi.org/10.1016/J.TRD.2016.10.024.

Goodman, R. 1997. The strengths and difficulties questionnaire: a research note. *Journal of Child Psychology and Psychiatry*. 38:581-586. DOI: https://dx.doi.org/10.1111/j.1469-7610.1997.tb01545.x.

Gössling, S. & Metzler, D. 2017. Germany's climate policy: facing an automobile dilemma. *Energy Policy*. 105:418-428. DOI: https://doi.org/10.1016/j.enpol.2017.03.019.

Greene, D.L., Evans, D.H. & Hiestand, J. 2013. Survey evidence on the willingness of U.S. consumers to pay for automotive fuel economy. *Energy Policy*. 61:1539-1550. DOI: https://doi.org/10.1016/j.enpol.2013.05.050.

Greenstone, M. 2017. The continuing impact of Sherwin Rosen's hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*. 125(6):1891-1902. DOI: https://doi.org/10.1086/694645.

Griliches, Z. 1958. The demand for fertilizer: an econometric reinterpretation of a technical change model. *Journal of Farm Economics*. 40(2):591-606.

Griliches, Z. 1961. Hedonic price indexes for automobiles: an econometric analysis of quality change. In *The Price Statistics of the Federal Government* (Report of the Price Statistics Review Committee). Cambridge, MA: National Bureau of Economic Research: 173-196. Available: http://www.nber.org/chapters/c6492 [2025, June 20].

Haas, G.C. 1922. *Sale prices as a basis for farmland appraisal* (Technical Bulletin 9). University Farm, St. Paul, MN: Agricultural Experiment Station, University of Minnesota.

Haliloglu, E.Y., & Berument, M.H. 2021. The asymmetric effects of crude oil prices and exchange rates on diesel prices for 27 European countries. *Global Business Review*, 25(4):888-912. DOI: https://doi.org/10.1177/0972150921999035.

Harper, J. 2021. Africa begins to emerge as car industry hub. *Deutsche Welle*. 15 October. Available: https://www.dw.com/en/africa-begins-to-emerge-as-car-industry-hub/a-59500532 [December 5, 2024].

Harris, R. 1997. Evaluating internet research sources. *VirtualSalt*. 13 October. Available: https://www.virtualsalt.com/evaluating-internet-research-sources/ [November 22, 2021].

Heckman, J.J. 2006. Skill formation and the economics of investing in disadvantaged children. *Science*. 312(5782):1900-1902. DOI: https://doi.org/10.1126/science.1128898.

Hsueh, L. 2000. The relationship between housing price, tenure choice and saving behavior in Taiwan. *International Real Estate Review*. 3(1):11-33.

Huang, C. & Lin, B. 2007. A hedonic analysis of fresh tomato prices among regional markets. *Journal of Applied Economic Perspectives and Policy*. 29(4):783-800.

Hung, W., Shang, J. & Wang, F. 2010. Pricing determinants in the hotel industry: quantile regression analysis. *International Journal of Hospitality Management*. 29:378-384.

Karato, K., Movshuk, O. & Shimizu, C. 2015. Semiparametric model of hedonic housing prices in Japan. *Asian Economic Journal*. 29(4):325-340.

Kotler, P. & Armstrong, G. 2004. *Principles of marketing*. 10th ed. Upper Saddle River, NJ: Pearson Education.

Kumbhakar, S.C. & Parmeter, C.F. 2010. Estimation of hedonic price functions with incomplete information. *Empirical Economist*. 39:1-25. DOI: https://doi.org/10.1007/s00181-009-0292-8.

Ladd, G.W. & Suvannunt, V. 1976. A model of consumer goods characteristics. *American Journal of Agricultural Economics*. 58(3):504-510. DOI: https://doi.org/10.2307/1239267.

Lancaster, K.J. 1966. A new approach to consumer theory. *Journal of Political Economy*. 74(2):132-157.

Lee, H., Park, K. & Kim, S. 2003. Estimation of information value on the Internet: application of hedonic price model. *Electronic Commerce Research and Applications*. 2(1):73-80. DOI: https://doi.org/10.1016/S1567-4223(03)00002-4.

Lightstone Auto. 2021. *Live brochure*. Available: https://www.lightstoneauto.co.za/downloads/LiveBrochure.pdf [November 23, 2021].

Liu, T.-C. (1960). Underidentification, structural estimation, and forecasting. *Econometrica*, 28(4), 855-865. Available: https://www.jstor.org/stable/1907567 [June 19, 2025]

Lorde, T., Jadon, J. & Weekes, Q. 2019. Price-setting behavior in a tourism sharing economy accommodation market: a hedonic price analysis of Airbnb hosts in the Caribbean. *Tourism Management Perspectives*. 30:251-261. DOI: 10.1016/j.tmp.2019.03.006.

Maietta, O.W. 2003. The hedonic price of fair-trade coffee for the Italian consumer. Paper presented at the *International Conference: Agricultural Policy Reform and the WTO: Where are We Heading?* 23-26 June, Capri, Italy.

Mannes, J. 2016. Online used car dealer Vroom leverages virtual reality to bring the showroom to you. *TechCrunch*. 2 August. Available: https://techcrunch.com/2016/08/02/online-used-car-dealer-vroom-leverages-virtual-reality-to-bring-the-showroom-to-you/ [March 10, 2022].

Matas, A., Raymond, J. & Roig, J. 2009. Car ownership and access to jobs in Spain. *Transportation Research Part A: Policy and Practice*. 43(6):607-617. DOI: https://doi.org/10.1016/j.tra.2009.04.003.

Maurer, R., Pitzer, M. & Sebastian, S. 2004. Hedonic price indices for the Paris housing market. *Allgemeines Statistisches Archiv*. 88:303-326. DOI: https://doi.org/10.1007/s101820400173.

McKinsey & Company. 2019. *The future of mobility is at the intersection of the automotive and smartphone industries*. Available: https://www.mckinsey.com [October 10, 2021].

McMillen, D. & Redfearn, C. 2010. Estimation and hypothesis testing for nonparametric hedonic house price functions. *Journal of Regional Science*. 50(3):712-733.

Mercedes-Benz leverages on blockchain for secure, tamperproof data tracking powered by PlatON. 2019. *International Business Times*. 30 August. Available: https://www.ibtimes.sg/mercedes-benz-leverages-blockchain-secure-tamperproof-data-tracking-powered-by-platon-32282 [October 10, 2024].

Meyer, J.R. & Glauber, R.R. 1964. *Investment decisions, economic forecasting, and public policy*. Boston, MA: Division of Research, Graduate School of Business Administration, Harvard University.

Mishra, D.P., Heide, J.B. & Cort, S.G. 1998. Information asymmetry and levels of agency relationships. *Journal of Marketing Research*. 35(3):277-295.

Mok, H.M.K., Chan, P.P.K. & Cho, Y.S. 1995. A hedonic price model for private properties in Hong Kong. *The Journal of Real Estate Finance and Economics*. 10(1):37-48.

Monson, M. 2009. Valuation using hedonic pricing models. *Cornell Real Estate Review*. 7:62-73.

Mordor Intelligence. 2021. *Automotive market in Africa: Growth, trends, COVID-19 impact, and forecast (2021-2026)*. Available: https://www.mordorintelligence.com/industry-reports/africa-automotive-industry-outlook [February 25, 2023].

Moresino, F. 2019. A hedonic approach to estimate the price of reliability, energy efficiency and safety for new cars in Switzerland. *American Journal of Industrial and Business Management*. 9(3):679-694.

Morgan, A., Daniels, P. & Murphy, J. 1979. Hedonic pricing and the demand for breakfast cereals. *Review of Economics and Statistics*. 61(4):564-572.

Naamsa. 2021. *Quarterly review of business conditions: New motor vehicle manufacturing industry automotive sector*. Pretoria: National Association of Automobile Manufacturers of South Africa.

Naamsa. 2024. Naamsa releases October 2024 new vehicle sales statistics [Media release, 1 November]. Available: https://naamsa.net/wp-content/uploads/2024/11/20241101-naamsa-October-2024-new-vehicle-sales-media-release.pdf [December 12, 2023].

Oczkowski, E. 1994. A hedonic price function for Australian premium table wine. *Australian Journal of Agricultural Economics*. 38(1):93-110.

Ohta, M. & Griliches, Z. 1976. Automobile prices revisited: extensions of the hedonic hypothesis. In *Household production and consumption.* N.E. Terleckyj, Ed. New York, NY: National Bureau of Economic Research, 325-390.

Orrego, E., Defrancesco, M. & Gennari, A. 2012. The wine hedonic price models in the "Old and New World": State of the art. *Revista de la Facultad de Ciencias Agrarias*. 44(1):205-220.

Pakes, A. 2003. A reconsideration of hedonic price indexes with an application to PCs. *American Economic Review*. 93(5):1578-1596.

Palmquist, R.B. 1984. Estimating the demand for the characteristics of housing. *The Review of Economics and Statistics*. 66(3):394-404.

Pazarlioglu, M.V. & Gunes, M. 2000. The hedonic price model for fusion car market. *Proceedings of the 10th Mediterranean Electrotechnical Conference (MELECON 2000)*, Vol. 3. IEEE. 1066-1069.

Prieto, M., Caemmerer, B. & Baltas, G. 2015. Using a hedonic price model to test prospect theory assertions: the asymmetrical and nonlinear effect of reliability on used car prices. *Journal of Retailing and Consumer Services*. 22:206-212.

Priilaid, D. 2016. Positioning fast moving luxury goods in a discount environment: how cue effects explain South African whisky prices. *International Journal of Wine Business Research*. 28(1):22-42. DOI: https://doi.org/10.1108/IJWBR-05-2015-0016.

Priilaid, D. & Hendry, S. 2019. Hedonic pricing analysis of entry-level, second-hand cars in South Africa. *Proceedings of the 31st annual conference of the Southern African Institute for Management Scientists (SAIMS) on "Responsible Management in Disruptive Times"*, 8-11 September 2019. S.M. Farrington, Ed. 576-596.

Priilaid, D.A. & Van Rensburg, P. 2006. Non-linearity in the hedonic pricing of South African red wines. *International Journal of Wine Marketing*. 18(3):166-182.

Priilaid, D. & Van Rensburg, P. 2007. Non-linearity in the hedonic pricing of South African red wines. *16th conference of the Southern African Finance Association*, Cape Town, 17 January 2007.

Rosen, S. 1974. Hedonic prices and implicit markets: product differentiation in pure competition. *The Journal of Political Economy*. 82(1):34-55.

Shamsher, R. 2014. Estimating product efficiency through a hedonic pricing best practice frontier. *Business Research Quarterly*. 21(4):215-224.

Shi, W. & Price, C. 1998. Applications of the hedonic pricing method in forest economics: an overview. *Forestry*. 71(3):213-232.

Sirmans, G.S., Macpherson, D.A. & Zietz, E.N. 2005. The composition of hedonic pricing models. *Journal of Real Estate Literature*. 13(1):3-43.

Stapleford, T.A. 2011. Reconceiving quality: political economy and the rise of hedonic price indexes. *History of Political Economy*. Annual Suppl. 43(5):309-328.

Stigler, G.J. 1961. The economics of information. *Journal of Political Economy*. 69(3):213-225.

Technavio. 2020. *Artificial intelligence (AI) market by end-user and geography: forecast and analysis 2021-2025*. Available: https://www.technavio.com/report/artificial-intelligence-ai-market-industry-analysis [November 25, 2021].

Thomas, D.P. 2016. Public transportation in South Africa: challenges and opportunities. *World Journal of Social Science Research*. 3(3):352-366.

Thrane, C. 2004. In defense of the price hedonic model in wine research. *Journal of Wine Research*. 15(2):123-134.

Thrane, C. 2007. Examining the determinants of room rates for hotels in capital cities: the Oslo experience. *Journal of Revenue and Pricing Management*. 5(4):315-323.

Torres, E.N., Adler, H. & Behnke, C. 2014. Stars, diamonds, and other shiny things: the use of expert and consumer feedback in the hotel industry. *Journal of Hospitality and Tourism Management*. 21:34-43.

Triplett, J.E. 1969. Automobiles and hedonic quality measurement. *Journal of Political Economy.* 77(3):408-417.

Triplett, J.E. 1986. The economic interpretation of hedonic methods. *Survey of Current Business*. 66(1):36-40.

Trochim, W.M.K. 2006. *The Research Methods Knowledge Base*. 3rd ed. Cincinnati, OH: Atomic Dog Publishing.

Turrión, J. & Velázquez, F.J. 2016. Hotels' attitudes towards innovations: an application to hospitality. *Journal of Destination Marketing & Management*. 5(2):97-106.

Valaskova, K., Kramarova, K. & Bartosova, V. 2015. Multi-dimensional analysis of consumer preferences in the food market. *Procedia Economics and Finance*. 34:622-629.

Van Rensburg, P. & Priilaid, D.A. 2004. An econometric model for identifying value in South African red wine. *International Journal of Wine Marketing*. 16(1):53-75.

Wafula, D. 2019. Mercedes joins PlatOn for used car data storage blockchain. *CoinGeek*. 30 August. Available: https://coingeek.com/mercedes-joins-platon-for-used-car-data-storage-blockchain/ [February 25, 2023].

WeBuyCars. 2021. *Our story.* Available: https://www.webuycars.co.za/our-story [November 21, 2021].

WeBuyCars may get new owner. 2021. *Fin24.* 12 May. Available: https://www.news24.com/fin24/companies/industrial/webuycars-may-get-new-owner-20210512 [2021, November 29].

White, P.J. & Mulligan, G.F. 2002. Hedonic estimates of lodging rates in the Four Corners region. *The Professional Geographer*. 54(4):533-543.

Wu, J., Wang, M., Li, W., Peng, J. & Huang, L. 2015. Evaluating the effects of landscape on housing prices in urban China: case study in Shenzhen. *Journal of Urban Planning and Development.* 141(4):Article 05014023. DOI: https://doi.org/10.1061/(ASCE)UP.1943-5444.0000241.