Price Determinants and Depreciation of Used Cars Post-COVID-19

Ayaan Sundeep Patel

Professor Michelle Connolly, Ph.D., Faculty Advisor

Professor Andrea Lanteri, Ph.D., Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University
Durham, North Carolina
2023

Acknowledgements

I would like to thank my main advisor, Andrea Lanteri, for helping me brainstorm ideas and for providing me with crucial guidance along the way. I want to thank my secondary advisor, Michelle Connolly, for consistently supporting me through various obstacles that arose throughout this project. I would like to thank Kent Kimbrough, my Econ 495S professor, for your helpful advice and encouragement while I was starting my research. I want to thank my honors seminar classmates for their advice and suggestions during my project. Lastly, thank you to my friends and family for supporting me throughout my entire time at Duke.

Abstract

Throughout the COVID-19 pandemic, the price of used cars has fluctuated greatly due to numerous factors. Inflation and supply chain issues have been at the forefront of the news and have affected not only cars but most consumer goods. While the majority of society has seemingly progressed past COVID-19, its effects still linger in the used car market, as prices rose 4.6% from January 2023 to February 2023. Therefore, in an effort to study this phenomenon, I scraped data from autotrader.co.uk on February 23, 2023. This study aims to understand the effect of various factors, including mileage, age, and engine size, on various classes of used cars. The five classes being studied are compact cars, luxury sports sedans, luxury mid-size sedans, luxury full-size sedans, and luxury SUVs. A log-linear model is used to model the price determinants of the used cars. A linear model is incorporated to model the depreciation rate of the cars in the dataset. Lastly, this model is used to predict the three-year depreciation rate for each car model, which is then compared to the pre-COVID-19 three-year depreciation rate to see the inflated prices in the UK used car market.

Keywords: Used Cars; COVID-19; Inflation; Depreciation; Secondary Markets

1. Introduction

Throughout COVID-19, the price of many new and used goods increased due to supply chain issues. Food prices have risen 65% globally in the last two years as a baseline for inflation of consumer goods (Morgan Stanley, 2022). Additionally, the market of used cars has been greatly affected by general supply chain issues. A chip shortage forced car companies to cancel orders at the beginning of the pandemic, but even as production ramped up near the end of 2020, there was no supply for semiconductors, the chip needed for everything from entertainment systems to power steering (J.P. Morgan, 2023). In cars, semiconductors or chips are needed for everything from entertainment systems to power steering. As a result, many car manufacturers were not able to make as many new cars during COVID-19 due to the chip shortage, increasing the value of both new and used cars. Here, I study the used car market by analyzing current depreciation rates and the various factors that affect the price of used cars in a post-COVID-19 market. To do so, I scrape data from AutoTrader (UK) for 22 car models and split them into five different classes: compact cars, luxury sports sedans, luxury mid-size sedans, luxury full-size sedans, and luxury SUVs. The names of the car models are included in Appendix A, and the car model designations for each class are displayed in Appendix B. Then, a linear and log-linear model is used to model the depreciation rate and price, respectively. This also allows for analysis of the effect of independent variables, such as age, mileage, and engine size, on the response variables.

The UK specifically experienced a notable increase in demand for used cars, as used car prices grew for the thirty-fifth month in a row in February 2023 (Williams, 2023). This streak began in March 2020, approximately when the COVID-19 lockdowns were mandated worldwide. Additionally, AutoTrader (UK) reported that the average retail value of a used car grew by 1.3% year-over-year in February 2023, not indicating much overall growth in price but still signifying that prices are currently inflated (Williams, 2023). The end of supply chain issues and the chip shortage may be near, as experts found that new car volumes rose by 31% from February 2022 to February 2023. However, this new car stock is still 60% below the record levels in early 2021 (Williams, 2023). New car production is ramping up once again, but used car prices in February 2023 still provide for an interesting analysis as they reflect volatile and inflated used car prices that were present during the COVID-19 pandemic.

McKinsey (2019) indicated that the used car market has become more important over time with the introduction of online car marketplaces. They "estimate that the number of used vehicles three years old or less will increase from 51 percent of the total in 2017 to about 60 percent in 2022" (Ellencweig, 2019). They noted that used-car prices have been rising partly due to a shift in the mix of cars that are sold. Car manufacturers have been increasing their selection of full-size trucks and midsize or larger SUVs, and these types of cars tend to be more expensive due to their larger size. As a result, even though the price of used mid-size and compact cars declined by 1.3% and 1.6%, the overall average used car price increased by 2.7% between 2012-2017, indicating that different classes of cars are experiencing varying effects in the current market (Ellencweig, 2019).

This study explores the determinants of price and depreciation rates for used cars in the inflated post-COVID-19 market. Through my analysis, we can also see the effect of a car's class on the various determinants of price and depreciation rates. By using the base MSRP for each year of a model as a proxy measure for the new car price of each model in a certain year, we can understand and estimate the depreciation rates of various cars and put them into context. We hope to see lower depreciation rates for cars in my dataset, compared to cars in 2019 or earlier, which would support the hypothesis that the current used car market is priced above pre-COVID-19 levels.

2. Literature Review

Although car prices have already peaked in December 2021, Manheim reported a 4.3% increase in vehicle prices in February, the largest increase for that month since 2009 (Lee, 2023). Ryan Brinkman from J.P. Morgan mentions that "half of the increase in new vehicle prices relates to the passing along of higher input costs, including raw materials costs" (J.P. Morgan, 2023). This makes sense, but the purpose of my economic analysis will be to understand what else is causing the rise in car prices. One possibility is that there is currently a high level of inflation, so I will be analyzing the change in relative prices adjusted for inflation, as I am looking over the change in prices over a long period of time. This influences my own research, as I plan to take inflation into account when analyzing prices over time by converting all prices to real 2023 British pound sterling. CoPilot, a car shopping app, reveals that consumers in the U.S. today are paying an average of $10,046 more for a used car in August 2022, while also adding that used car prices increased 7.1% year-over-year in June 2022 (J.P. Morgan, 2023).

In "A Quantitative Analysis of the Used-Car Market", Alessandro Gavazza mentions that "the dispersion of used-car prices is approximately five times as large as the dispersion of new-car prices," (Gavazza, 2014). This indicates that the secondary market for vehicles widens the quality of goods that are available to consumers due to the variability in new car prices. Additionally, the used car market will always inherently be more volatile than the new car market since the new car market is regulated by car companies who set the MSRP of each new car. This is promising for my research, as more volatility indicates that it may be easier to see the effect of certain variables on the price of a used car compared to a new car's MSRP. Gavazza primarily examines the difference in secondary car markets across different geographies. For example, the American used-car market is significantly more active than the French market (Gavazza, 2014). I plan to understand how variables within the domestic UK market affect the car market, but this study still provides context on how to perform an analysis of the used car market, specifically which ratios and response variables to examine.

To help with my formation of an econometric model, I analyzed an article from Susanna Esteban and Matthew Shum (2007) that studies the impact of car durability and secondary markets on equilibrium firm behavior in the car market. This article aimed to demonstrate the feasibility of fitting an econometric model for the secondary car market. They used a car quality parameter, alpha, that is estimated for various car models and types of cars. For example, car size is positively related to quality, and Ford and GM cars tend to offer higher quality than Chrysler cars (Esteban, 2007). Alpha also helps understand the depreciation of a used car's quality over time, especially across variables such as brand, size, price, etc. Although I do not plan to use this alpha concept in my own research, this paper taught me how to fit a model for depreciation rate and the alpha car parameter showed which car characteristics are responsible for the decrease in the price of a car.

Bertolotti et al. (2021) authored a paper about the role of new product quality on durable goods, which includes vehicles. He initially describes two very interesting reasons for primarily studying cars within durable goods. First, vehicles are "a large and procyclical component of durable-goods expenditures." (Bertolotti et al., 2021). Second, we are able to acquire detailed information about car markets, which allows us to study price changes. The dataset spans from 2004-2012, allowing Bertolotti to analyze the new car market before, during, and after the Great

Recession. The study follows four steps of analysis. They begin their analysis by studying new car expenditure during the Great Recession, which is then followed by an analysis of the dispersion of new-car expenditures during the recession. Then, a hedonic regression is used to "construct a measure of vertical quality that summarizes the main vehicle characteristics" (Bertolotti et al., 2021). This regression type is especially important because it estimates the value of a product by breaking it down into its constituent characteristics. It is especially useful in this case when comparing cars with different attributes to each other, which is why I use this regression type in my own analysis to value a used car. Using log-linear hedonic regression, Bertolotti et al. (2021) were able to display the "dynamics of the average transaction price and the average of our measure of quality." The measure of quality comes from the variables mentioned earlier (wheelbase, horsepower, engine size, mileage, age). This motivates my research as I plan to use engine size, mileage, and age in my model to measure car quality.

3. Data

3.1 Data Sources

My research project has three main data sources: AutoTrader (UK), Kelley Blue Book, and CarGurus. AutoTrader is the most utilized marketplace for used cars in the UK, so its car listings provide the greatest representation of the entire UK used car market. Kelley Blue Book (KBB) is a popular American automotive valuation company that provides pricing and value information for new and used cars, trucks, and motorcycles. In addition to its pricing guide, Kelley Blue Book provides a variety of other resources and services related to the automotive industry, including expert reviews, ratings, and comparisons of vehicles across different makes and models. Lastly, CarGurus is an online automotive marketplace and search engine that was founded in 2006. It allows users to search for new and used cars from dealerships and private sellers across the United States, Canada, and the United Kingdom.

3.2 Dataset Construction

The dataset was scraped from autotrader.co.uk on February 28, 2023. Although this time period could be viewed as post-COVID-19, the used car market still had inflated prices during this period. There is still a great deal of pressure on the used car market, meaning that my dataset and analysis are still relevant to the effects of COVID-19. I determined which car models to scrape by referencing Kelley Blue Book's "top 10" lists for the following classes: compact cars, luxury sports sedans, luxury mid-size sedans, luxury full-size sedans, and luxury SUVs. KBB.com editors examine vehicles in each class based on fuel economy, best value, safety options, features, technology, level of comfort, driving performance, and more.

The scraped dataset includes data about price, year manufactured, name of the car, transmission, mileage, number of owners, engine size, fuel type, and base horsepower. Although the price variable represents the listing price, not the selling price, I am assuming that these prices are an upper bound for the general transaction price for each of the observations in the dataset. I filtered out observations with a listing price greater than £250,000 to remove exorbitantly unreasonable prices in the market. Approximately 80% of data points had "NA" values for base horsepower, so it is not included in the analysis. A new "age" variable is created to provide a more sensible variable for how old a car is. This is done by subtracting the car's manufactured year from 2023. Additionally, I created a new variable named "class" that displayed whether a car model was a compact car, luxury sports sedan, luxury mid-size sedan, luxury full-size sedan, and luxury mid-size SUV. The total number of observation in the cleaned dataset is 88,743.

Kelly Blue Book is utilized to compile the base MSRP of each model from the years 2016-2023. This led me to filter out data points in the scraped dataset that had a model year before 2016, so all observations in my original datasets are complete and have no NA values for MSRP. MSRP serves as a guidepost for retailers, but it is not necessarily the price at which the product will ultimately be sold. Retailers may choose to sell the product for more or less than the MSRP, depending on factors such as supply and demand, competition, and their own profit margins. Within the context of buying a car, the MSRP is often used as a starting point for negotiations between the buyer and the dealer. For my project, MSRP is used to represent the new car price of each car model in a given model year. Since my scraped data are from one point in time, all of my price values are in real 2023 British pounds. As a result, I converted all of the MSRP values to real 2023 British pounds in order to account for the inflation that has occurred within the British market (Clark, 2023). Then, I create a depreciation variable that is found by calculating the percent change between MSRP and the listing price for each observation.

CarGurus Price Trends tool is the last data source used for my research. CarGurus tracks the prices of millions of used car listings every year. This means that you can track the depreciation rate of each individual model year for a certain car model over time. In the Results section, I will compare the 3-year post-COVID-19 depreciation rate predictions from my linear model to both the pre-COVID-19 and post-COVID-19 depreciation rates from CarGurus for each car model in the dataset which will help to put my findings into context.

3.2 Summary Statistics + Variable Analysis

Table 1: Summary Statistics for Independent Variables by Class

		Compact car	Luxury full-size	Luxury mid-size	Luxury sports	Luxury SUV
Observations	Count	24054	2550	17468	27867	16804
Price	Mean	19504.79	44701.28	27836.23	25052.24	59630.27
	Std. Dev.	6500.98	19715.69	10375.48	9601.75	27438.49
Mileage	Mean	35926.94	40965.47	40821.31	40678.58	36745.29
	Std. Dev.	64840.57	30342.31	28419.22	27226.61	26720.54
MSRP	Mean	22463.80	93897.47	52140.88	40381.99	72855.90
	Std. Dev.	2815.89	8001.43	5225.83	2954.46	13652.97
Age	Mean	4.30	4.59	4.41	4.60	4.11
	Std. Dev.	1.68	1.78	1.73	1.74	1.95
Engine	Mean	1.64	3.23	2.18	2.16	3.23
	Std. Dev.	0.31	0.79	0.43	0.51	0.85

Table 1 provides the mean and standard deviation of several characteristics for five different classes of cars: compact car, luxury full-size, luxury mid-size, luxury sports, and luxury SUVs. The first column indicates the characteristic being measured (Number of Obs, Price, Mileage, MSRP, Age, Engine), and the second column shows whether the data in the subsequent columns represent the count, mean, or standard deviation for that characteristic. When comparing the classes for specific characteristics, some interesting trends emerge. For example, the Luxury SUV class has the highest mean price and MSRP, while the compact car class has the lowest mean price and MSRP. In terms of mileage, the luxury sports, luxury full-size, and luxury full-size sedans all have a mean mileage of approximately 40,000 miles while the compact cars have the lowest mean mileage. For age, the luxury sports class has the highest mean age, while the luxury SUV class has the lowest mean age. Finally, in terms of engine size, the luxury full-size and luxury SUV classes have the largest mean engine sizes (3.23 L), while the compact car class has the smallest mean engine size (1.64 L). This makes sense, as engine size is likely correlated to the size of the car, and compact cars are the smallest class of cars in this dataset. The standard deviation values in the table can also provide insight into the amount of variation within each class for a given characteristic. For example, the luxury SUVs have a high standard deviation for price, suggesting that there is a wide range of prices within that class. In contrast, the compact car class has a relatively low standard deviation for age, suggesting that the ages of the cars in that class are more tightly clustered around the mean.

Figure 1: Stacked Bar Chart of Class by Transmission

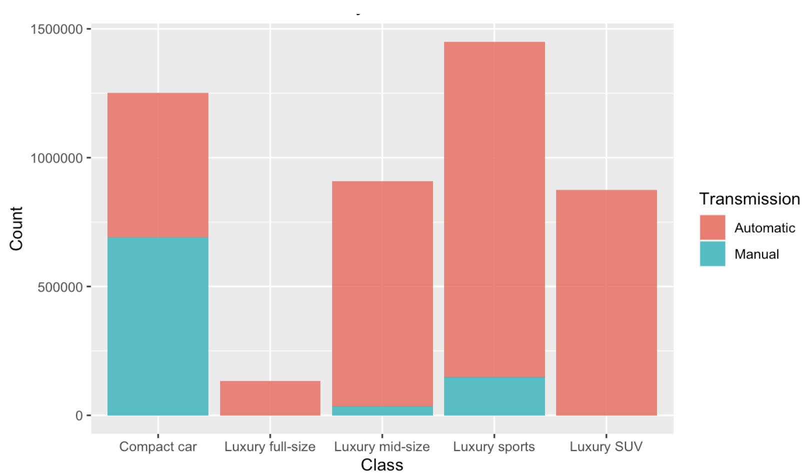

Figure 1 is a stacked bar chart that shows the relationship between the transmission variable and class, while also providing more information on the number of cars in each class.

We see that there are no observations with manual transmission in the luxury full-size and luxury SUV classes, indicating that transmission will not be a significant independent variable for either of these classes. However, more than half of compact cars have manual transmission, so there is some potential for transmission to be worked into the full model. There are very few observations for the luxury full-size class compared to the other classes, but this is likely a result of their lack of popularity among the general UK population.

I hypothesized that age and transmission are likely to be correlated because most newer cars are only sold with automatic transmissions, while manual transmissions are much more common in older cars. To test the correlation between “age” and “transmission”, I used a statistical method called an analysis of variance (ANOVA). ANOVA allows for comparison of the means of a numerical variable across different levels of a categorical variable. As can be seen, the p-value is very small (< 0.05), indicating that there is a significant difference in the means of “age” across different levels of “transmission.” This suggests correlation between the two variables, which is important to note and will be taken into account during model selection.

4. Methods and Empirical Framework

4.1 Depreciation Rate Linear Model

When modeling depreciation rate as the response variable, I use linear regression, as this response variable does not have a rate of exponential decay. Depreciation rate is used as a dependent variable because each observation is scaled to the MSRP of its given production year, which allows for some standardization of the listing price variable. The independent variables of the full model are name, mileage, transmission, age, and engine. Stepwise selection was used for model selection and fuel type was not included as a significant predictor of depreciation rate, so it was removed from the model. Number of owners was also dropped from the model, as the theory behind its interpretation did not make sense. When included in the model, the coefficient for number of owners would be negative, even though one would expect more owners to lead to a higher depreciation rate. Additionally, there are interaction terms between class and engine size, class and mileage, and class and age, while a quadratic term is included for mileage. The interactions allow for the possibility that the effects of engine size, age, and mileage on the depreciation rate varies depending on the class, which is what we ultimately see in the results when running the models for each separate class. The formula of the linear model for depreciation rate is shown below:

$$(1) DepreciationRate_i = \beta_0 + \beta_1 MSRP_i + \beta_2 Name_i + \beta_3 Mileage_i + \beta_4 Mileage_i^2 + \beta_5 Age_i + \beta_6 Transmission_i + \beta_7 Class_i * Engine_i + \beta_8 Class_i * Mileage_i + \beta_9 Class_i * Age_i + \varepsilon_i$$

For Eq. (1), $DepreciationRate_i$ is the depreciation rate of car *i*, and this variable is calculated in the dataset by taking the percentage difference between the car's listing price and its model year's MSRP in real 2023 British pounds. $MSRP_i$ is the MSRP of that car's model year in real 2023 British pounds. $Name_i$ is a fixed effect for each of the model names in the dataset, which is listed in Appendix A. $Mileage_i$ is a continuous variable indicating the number of miles driven by car *i* at the time of the listing, and $Mileage_i^2$ is the quadratic term of the $Mileage_i$ variable. Age_i represents the age of car *i*, and $Transmission_i$ is a binary variable indicating whether car *i* has manual transmission (1 = manual transmission, 0 = automatic transmission). The interaction terms are $Class_i * Engine_i$, $Class_i * Mileage_i$, and $Class_i * Age_i$, and they represent how the effects of $Engine_i$, $Mileage_i$, and Age_i differ across each $Class_i$. Lastly, ε_i represents the variation in the dependent variables that is not explained by the independent variable, $DepreciationRate_i$, in the model.

On the other hand, log-linear models are a common method used for modeling processes that exhibit a decreasing trend over time. This includes a wide range of phenomena, including radioactive decay, population growth, and in the context of car depreciation, the decline in value of a car over time. The basic idea behind a log-linear model is that the rate of decrease in the process being modeled is proportional to the current value of the process. This means that the rate of decrease will be greater at the beginning of the process, when the value is higher, and will gradually slow down as the value approaches zero. A similar model selection process was performed for this log-linear model compared to the linear depreciation model mentioned previously. As a result, fuel type and number of owners were not included as independent variables in the model. MSRP is not included in this model as it is already fully described through the name and age variables. Therefore, when adding it to the model, it did not provide any extra explanatory power and was, therefore, removed from the model.

Similar to Eq. (1) the following is the hedonic regression model that I use to estimate the used car prices:

$$(2)\ log(Price_i) = \beta_0 + \beta_1 Name_i + \beta_2 Mileage_i + \beta_3 Mileage_i^2$$
$$+ \beta_4 Age_i + \beta_5 Transmission_i + \beta_6 Class_i * Engine_i + \beta_7 Class_i * Mileage_i$$
$$+ \beta_8 Class_i * Age_i + \varepsilon_i$$

In Eq. (2), $log(Price_i)$ is the log transformation of the listing price of car *i*. $Name_i$ is a fixed effect for each of the model names in the dataset, which is listed in Appendix A. $Mileage_i$ is a continuous variable indicating the number of miles driven by car *i* at the time of the listing, and $Mileage_i^2$ is the quadratic term of the $Mileage_i$ variable. Age_i represents the age of car *i*, and $Transmission_i$ is a binary variable indicating whether car *i* has manual transmission (1 = manual transmission, 0 = automatic transmission). The interaction terms are $Class_i * Engine_i$, $Class_i * Mileage_i$, and $Class_i * Age_i$, and they represent how the effects of $Engine_i$, $Mileage_i$, and Age_i differ across each $Class_i$. Lastly, ε_i signifies the variation in the dependent variables that is not explained by the independent variable, $log(Price_i)$, in the model.

5. Results

5.1 Depreciation Rate Linear Model

Table 2: Model Output for Linear Model

term	estimate	std.error	statistic	p.value
(Intercept)	0.5891675	0.0115215	51.1365559	0.0000000
MSRP	-0.0000021	0.0000002	-8.5066768	0.0000000
nameAudi A6	-0.0391378	0.0061984	-6.3141611	0.0000000
nameAudi A8	0.0473416	0.0163461	2.8961929	0.0037781
nameBMW 3 Series	-0.0063761	0.0021012	-3.0345477	0.0024097
nameBMW 5 Series	0.0775254	0.0073405	10.5612877	0.0000000
nameBMW 7 Series	0.0105129	0.0168131	0.6252793	0.5317893
nameBMW X5	-0.4226880	0.0081188	-52.0629890	0.0000000
nameHonda Civic	-0.1712307	0.0073085	-23.4290827	0.0000000
nameLand Rover Range Rover	-0.3129126	0.0150726	-20.7602991	0.0000000
nameLand Rover Range Rover Sport	-0.4810297	0.0099684	-48.2553112	0.0000000
nameMazda3	0.1374756	0.0087240	15.7582864	0.0000000
nameMercedes-Benz C Class	-0.0249150	0.0022934	-10.8637441	0.0000000
nameMercedes-Benz E Class	0.0165291	0.0074522	2.2180008	0.0265573
nameMercedes-Benz S Class	-0.0392764	0.0197768	-1.9859867	0.0470379
namePorsche Cayenne	-0.5705840	0.0095564	-59.7071306	0.0000000
nameVolkswagen Golf	-0.0549193	0.0070885	-7.7476249	0.0000000
nameVolvo XC90	-0.5675253	0.0065914	-86.1007935	0.0000000
mileage	0.0000046	0.0000000	115.4868214	0.0000000
I(mileage^2)	0.0000000	0.0000000	-110.3149294	0.0000000
transmissionManual	0.1172815	0.0013018	90.0925241	0.0000000
age	0.0162151	0.0006688	24.2468783	0.0000000
engine	-0.3877589	0.0027208	-142.5141975	0.0000000

engine:classLuxury full-size	0.3140071	0.0041004	76.5788364	0.0000000
engine:classLuxury mid-size	0.2171400	0.0034464	63.0051335	0.0000000
engine:classLuxury sports	0.1832456	0.0030337	60.4041959	0.0000000
engine:classLuxury SUV	0.2945754	0.0029583	99.5756617	0.0000000
mileage:classLuxury full-size	-0.0000032	0.0000001	-30.2019904	0.0000000
mileage:classLuxury mid-size	-0.0000027	0.0000001	-49.1939597	0.0000000
mileage:classLuxury sports	-0.0000020	0.0000001	-38.9546027	0.0000000
mileage:classLuxury SUV	-0.0000018	0.0000001	-28.8549657	0.0000000
age:classLuxury full-size	0.0544620	0.0018647	29.2064508	0.0000000
age:classLuxury mid-size	0.0436005	0.0008991	48.4930540	0.0000000
age:classLuxury sports	0.0322351	0.0007913	40.7344753	0.0000000
age:classLuxury SUV	0.0800303	0.0008818	90.7592661	0.0000000

All coefficients in the model are statistically significant with p-values below 0.05, except for the BMW 7 Series fixed effect. This is likely because the average depreciation rates of Audi A4, the baseline model for the name variable, and BMW 7 Series are very similar to each other, reducing this particular fixed effect's significance in the model. The only other model with an average depreciation rate that is closer to that of the Audi A4 is the BMW 3 Series, and it is still significant likely because of the high number of observations for the BMW 3 Series (11,766) compared to that of the BMW 7 Series (909).

Holding all other variables constant, a car with a manual transmission is expected to increase the depreciation rate by approximately 11.7 percent compared to one with an automatic transmission. For luxury SUVs, each additional year of age suggests an increase of approximately 10 percent, holding all else constant, which is the highest value among all classes.

On the other hand, the depreciation rate of compact cars, the baseline class, is only expected to increase by 1.6 percent for each additional year of age. Mileage is correlated to the highest increase in depreciation rate for compact cars with an increase of 10,000 miles, indicating a 4.6% increase in the depreciation of a vehicle. The effect of mileage on the depreciation rate for luxury full-size sedans is nearly one-third of the effect of mileage on the depreciation rate for compact cars. Lastly, for compact cars, with each Liter increase in engine size, the depreciation rate is expected to decrease by 38%, which is a much larger decrease than the other classes. This is likely because the mean and standard deviation of the engine variable are lowest for compact compared to all other classes. The R-squared value of the linear depreciation rate model is 0.82, indicating that 82% of the variability in the dependent variable, depreciation rate, can be explained by the independent variables in the model. The remaining 18 percent of variation in the response variable is due to factors not accounted for in the model or due to random error.

5.2 Car Price Log-Linear Model (Hedonic Regression)

Table 3: Model Output for Car Price Log-Linear Model

term	estimate	std.error	statistic	p.value
(Intercept)	10.2886999	0.0037173	2767.82240	0
nameAudi A6	0.1226563	0.0063682	19.26076	0
nameAudi A8	0.4098623	0.0128724	31.84035	0
nameBMW 3 Series	0.0354689	0.0025816	13.73886	0
nameBMW 5 Series	0.1767737	0.0067639	26.13496	0
nameBMW 7 Series	0.4513937	0.0124511	36.25333	0
nameBMW X5	0.9050431	0.0065125	138.97019	0
nameHonda Civic	-0.5151380	0.0064576	-79.77200	0
nameLand Rover Range Rover	1.1825206	0.0071450	165.50397	0
nameLand Rover Range Rover Sport	1.0635044	0.0066328	160.34050	0
nameMazda3	-0.8424514	0.0082094	-102.62041	0
nameMercedes-Benz C Class	0.0852141	0.0026239	32.47629	0
nameMercedes-Benz E Class	0.2318147	0.0067546	34.31973	0
nameMercedes-Benz S Class	0.6824625	0.0127815	53.39443	0
namePorsche Cayenne	1.1202623	0.0071906	155.79437	0
nameVolkswagen Golf	-0.4718972	0.0064502	-73.15974	0
nameVolvo XC90	0.8095329	0.0071531	113.17280	0
mileage	-0.0000070	0.0000000	-175.65235	0
I(mileage^2)	0.0000000	0.0000000	156.40408	0
transmissionManual	-0.1398556	0.0014813	-94.41629	0
age	-0.0958403	0.0003521	-272.18165	0
engine	0.4646519	0.0031501	147.50318	0

mileage:classLuxury full-size	0.0000021	0.0000001	22.99696	0
mileage:classLuxury mid-size	0.0000019	0.0000001	34.00108	0
mileage:classLuxury sports	0.0000014	0.0000001	27.28156	0
mileage:classLuxury SUV	0.0000019	0.0000001	31.19959	0
age:classLuxury full-size	-0.0421963	0.0008412	-50.16418	0
age:classLuxury mid-size	-0.0266011	0.0005994	-44.38222	0
age:classLuxury sports	-0.0243974	0.0004842	-50.38330	0
age:classLuxury SUV	-0.0471189	0.0005636	-83.60172	0
engine:classLuxury full-size	-0.2755330	0.0045340	-60.77062	0
engine:classLuxury mid-size	-0.2639702	0.0038445	-68.66089	0
engine:classLuxury sports	-0.2245802	0.0033909	-66.23102	0
engine:classLuxury SUV	-0.3585636	0.0034956	-102.57488	0

Table 3 shows the coefficients and statistical significance of the variables in the model where the dependent variable is the natural logarithm of the price. The intercept is 10.288, which means that the predicted price is $10,815.41 when all other variables are zero. The coefficients of the name variables indicate the estimated differences in log(price) for each category of the name variable, holding all other variables constant. For each Liter increase in engine size, the price of a car is expected to increase by about 46% and 10% for compact cars and luxury SUVs, respectively, holding all else constant, which was the largest margin in coefficients between classes. As mentioned previously, this is likely due to the smaller engine sizes in compact cars. A manual transmission also correlates to a 14 percent lower car price than an otherwise similar car with an automatic transmission. For each one-year increase of the age variable, the price of a car is only expected to decrease by about 5 percent for both luxury full-size and luxury SUVs , indicating that these cars hold their value the most over time compared to the other classes. Moreover, the car price for compact cars is expected to decrease by 7 percent for each 10,000-mile increase in the mileage, holding all else constant. Car price is expected to decrease by approximately 5 percent for all other classes for a 10,000-mile increase. Amongst all car models, the Mazda3 has the lowest coefficient, specifically indicating that the price of a Mazda3 is about 83 percent lower than the price of an Audi A4, the baseline model for the name variable. On the other hand, the Land Rover Range Rover has the highest intercept, showing that the price of a Land Rover Range Rover is 118% greater than an Audi A4. The R-squared value for this log-linear price model (0.9416) is much higher than that of the linear depreciation rate model (0.82). There is little variance in the price variable that is not explained by the independent variable in the model, which is a great sign. The increase in R-squared values between the two models is likely because the log-transformation is able to model the price variable better than a linear trend models the depreciation rate variable.

6. Discussion

Table 4: Comparison of Depreciation Rates Across Time-Periods

Time Period	2016-2019	2020-2023	
Source	CarGurus	CarGurus	Linear Model Predictions
Audi A4	42.3%	18.0%	31.2%
Audi A6	29.2%	26.4%	33.3%
Audi A8	48.6%	40.3%	47.3%
BMW 3 Series	40.2%	36.0%	25.7%
BMW 5 Series	45.4%	34.2%	38.9%
BMW 7 Series	47.2%	32.6%	42.8%
BMW X5	33.5%	21.6%	13.2%
Honda Civic	22.6%	11.1%	10.3%
Land Rover Range Rover	51.2%	33.5%	15.1%
Land Rover Range Rover Sport	48.6%	27.1%	-1.6%
Mazda3	24.4%	15.9%	11.5%
Mercedes-Benz C Class	43.8%	14.3%	25.2%
Mercedes-Benz E Class	45.9%	18.8%	33.7%
Mercedes-Benz S Class	60.0%	23.0%	37.17%
Porsche Cayenne	30.6%	28.2%	-8.1%
Volkswagen Golf	39.4%	12.2%	6.6%
Volvo XC90	48.6%	22.2%	9.3%

Table 4 compares data gathered from CarGurus with depreciation rate predictions from my linear model. The first column tracks the depreciation rate of the 2016 model year of each car model from its release date until February 2019 to provide a pre-COVID-19 reference point for my depreciation rate predictions. The second column tracks the depreciation rate of the 2020 model year of each car from its release date until February 2023, which encompasses the effect of COVID-19 on the used car market. The third column was calculated by predicting depreciation rate values for each car model using the linear depreciation model mentioned previously. These depreciation rate predictions for each car model are based on the following inputs for the linear model (Eq. (1)). Because this regression provides predictions for each car model, the name variable input is equal to that respective car model being studied. Age is equal to 3 for all models because we are tracking the depreciation rate of 2020-year models until February 2023. The model inputs for MSRP, engine size, and mileage for each car model are equal to the average of these values for all 2020-year vehicles of each car model. Lastly, the transmission input for the model is equal to the proportion of vehicles in a certain car class with a manual transmission. Next, with all of these model inputs, we run the linear regression for the depreciation rate for each car model, which results in the third column of Table 4.

The depreciation rates in Table 4 suggest that there has been an overall decrease in depreciation rate between 2016-19 and 2020-23 for the car models listed, with one exception being the Audi A6. This supports the hypothesis that the used car market during COVID-19 is inflated, compared to the pre-COVID-19 used car market. There does seem to be a noticeable variance between the 2020-23 CarGurus and 2020-2023 AutoTrader depreciation rates. This likely occurs because AutoTrader only has listings from the UK market and CarGurus has listings from the US and UK markets. The used car market varies for these geographies, as the US market generally has more variety due to its larger population and higher car ownership rates. Moreover, CarGurus has access to significantly more data than I do, so their model output is likely much less influenced by potential outliers that exist in my dataset. It is also important to note that we model depreciation rate differently from each other. CarGurus calculates the depreciation rate of each model year of a car model by aggregating all of the used car price data over time for each model year, whereas I calculate the depreciation rate variable by taking the percent change between the car's listing price and that model year's base MSRP in real 2023 British pounds. Then, I use depreciation rate as my response variable and determine inputs for the linear model by aggregating data for certain variables. However, it is still notable that both sources' depreciation rates for 2020-2023 are lower than CarGurus' depreciation rates from 2016-19.

The compact cars (Volkswagen Golf, Mazda3, and Honda Civic) all have three-year depreciation rates between 6 and 12 percent, which is the lowest among all the classes. This is likely due to several factors, including their lower purchase price and lower cost of maintenance compared to luxury vehicles (Barry, 2021). The three luxury full-size sedans (Audi A8, BMW 7 Series, Mercedes Benz S Class) have three of the four highest three-year depreciation rate predictions. There are a few potential reasons for their abnormally high-depreciation rates. These three car models had the lowest number of observations in the dataset. In the UK, smaller cars are also generally preferred due to narrower roads and tighter parking spaces, as well as higher fuel prices and taxes on larger vehicles. Luxury SUVs may be an exception to this reasoning, as there is currently a very high demand for SUVs around the world. SUVs accounted for about 45% of all global car sales in 2021, and the number of SUVs in use from 2021 to 2022 increased by 35 million (Rokke, 2022). This also supports why the depreciation rates for luxury SUVs are considerably lower than other classes, ranging from -8% to 9%. With reference to specific vehicles, it is notable that the 2020 Land Rover Range Rover and 2020 Porsche Cayenne are both estimated to have a positive depreciation rate as of February 2023. It seems abnormal for a 3-year-old vehicle to be worth more than when it was a new vehicle, but that may not be very surprising in the current used car market. Additionally, Land Rover is a British brand, potentially leading to more demand among the UK population for this vehicle and thus increasing the prices of the Range Rover model on the used car market.

There are a couple of limitations to my research project that may be responsible for some skew within the results. Because the new-car price of each observation in my dataset was not available, I used the base MSRP for each model year of each car model as the new car price for a specific car model's year. It is an imperfect proxy, but I am able to control for inflation by putting all MSRP values in real 2023 British pounds. I also account for increases in production costs over time for these vehicles by gathering MSRP data for every consecutive year from 2016-2023. In theory, the changes in MSRP after accounting for inflation should represent the change in production costs. Additionally, I only include the top three to five models within each car class for this study, so my results may not represent the population of cars with lower quality within each class.

7. Conclusion

Overall, the findings presented in this paper are consistent with existing literature on the current inflation within the used car market. The main contribution of my research is to provide a comprehensive study of the depreciation rate and car price post-COVID-19 in the UK used car market in order to compare results across different models and classes. Based on findings in this paper, luxury full-size sedans are predicted to currently have the highest 3-year depreciation rate, while luxury SUVs are predicted to have the lowest 3-year depreciation rate. Compared to that of other classes, the expected depreciation rate of compact cars is most influenced by an increase in mileage and least influenced by an increase in age. The high R-squared value (0.94) for the car price model indicates that a log transformation of the car price variable allows for the independent variables to explain a high degree of this response variable's variation. Lastly, the comparison between 3-year depreciation rate predictions from 2020-23 and the CarGurus depreciation rates from 2016-19 confirms my hypothesis that the depreciation rate for vehicles is currently lower than it was before COVID-19.

References

Barry, K. (2021, February 3). *Deal breakers: The cars and suvs with highest repair costs for common fixes*. Consumer Reports. Retrieved April 13, 2023, https://www.consumerreports.org/car-repair/cars-suvs-with-highest-repair-costs-for-common-fixes-a1155971972/

Bertolotti, F., Gavazza, A., & Lanteri, A. (2023). Dynamics of expenditures on Durable Goods:The role of new-product quality. *The Economic Journal*, *133*(652), 1641–1656. https://doi.org/10.1093/ej/uead006.

Clark, D. (2023, March 22). *UK CPI 2023*. Statista. Retrieved April 14, 2023, from https://www.statista.com/statistics/306631/consumer-price-index-cpi-united-kingdom/

Ellencweig, B., Ezratty, S., Fleming, D., & Miller, I. (2019). *Used cars, new platforms: Accelerating sales in a digitally disrupted market. McKinsey & Company.* Retrieved December 19, 2022, https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/used-cars-new-platforms-accelerating-sales-in-a-digitally-disrupted-market

Esteban, S., & Shum, M. (2007). Durable-Goods Oligopoly with Secondary Markets: The Case of Automobiles. *The RAND Journal of Economics*, *38*(2), 332–354. http://www.jstor.org/stable/25046309

Gavazza, A., Lizzeri, A., & Roketskiy, N. (2014). A Quantitative Analysis of the Used-Car Market. *The American Economic Review*, *104*(11), 3668–3700. http://www.jstor.org/stable/43495350

J.P. Morgan Research. (2023) *When Will Car Prices Drop?* J.P. Morgan, Retrieved April 13, 2023, www.jpmorgan.com/insights/research/when-will-car-prices-drop.

J.P. Morgan Research. (2022, August 11). *How long will the chip shortage last?* J.P. Morgan.

Retrieved April 13, 2023,
https://www.jpmorgan.com/insights/research/supply-chain-chip-shortage

Morgan Stanley. (2022, July 6). *Seeing a peak to rising food prices*. Morgan Stanley. Retrieved April 13, 2023, from https://www.morganstanley.com/ideas/rising-food-prices-2022-outlook

New and Used Cars For Sale. Auto Trader UK. (n.d.). Retrieved April 13, 2023, from https://www.autotrader.co.uk/?refresh=true

New Car & Used Car Prices. Kelley Blue Book. (n.d.). Retrieved April 14, 2023, from https://www.kbb.com/car-prices/

Rokke, N. (2022, February 16). *SUV sales record highlights Climate challenge ahead*. Forbes. Retrieved April 13, 2023, from https://www.forbes.com/sites/nilsrokke/2022/02/14/suv-sales-record-highlights-climate-challenge-ahead/?sh=72e2e3907012

Sraders, A., & Lambert, L. (2021, November 1). *What to expect in the 2022 used car market*. Fortune. Retrieved April 13, 2023, from https://fortune.com/2021/11/01/used-car-prices-high-carmax-2021/

Wayland, M. (2023, March 8). *Used vehicle prices rising at an unseasonably strong rate*. CNBC. Retrieved April 13, 2023, from https://www.cnbc.com/2023/03/07/used-vehicle-prices-rising-at-an-unseasonably-strong-rate.html

Williams, J. (2023, March 6). *Used car prices grow for 35th month in a row as Auto Trader data also reveals huge leap in new car costs*. Car Dealer Magazine. Retrieved April 1, 2023, https://cardealermagazine.co.uk/publish/used-car-prices-grow-for-35th-month-in-a-row-as-auto-trader-analysis-also-reveals-huge-leap-in-new-car-values/280631

Appendix

Exhibit A: List of car models in dataset

- Audi A4
- Audi A6
- Audi A8
- BMW 3 Series
- BMW 5 Series
- BMW 7 Series
- BMW X5
- Honda Civic
- Land Rover Range Rover
- Land Rover Range Rover Sport
- Mazda3
- Mercedes-Benz C Class
- Mercedes-Benz E Class
- Mercedes-Benz S Class
- Porsche Cayenne
- Volkswagen Golf
- Volvo XC90

Exhibit B: List of car models by class

- Compact cars
  - Honda Civic
  - Volkswagen Golf
  - Mazda3
- Luxury sports sedan
  - Audi A4
  - BMW 3 Series
  - Mercedes-Benz C Class
- Luxury mid-size sedan
  - Audi A6
  - BMW 5 Series
  - Mercedes-Benz E Class
- Luxury full-size sedan
  - Audi A8
  - BMW 7 Series
  - Mercedes-Benz S Class
- Luxury SUV
  - Porsche Cayenne

- Volvo XC90
- BMW X5
- Land Rover Range Rover
- Land Rover Range Rover Sport