Hedonic Pricing of Vehicle Characteristics, Safety and Equipment in the UK Car Market

Filip Mandys [a]
[a] School of Economics, University of Surrey, Guildford, GU2 7XH, UK
f.mandys@surrey.ac.uk

Abstract

UK has one of the largest car markets in the world, with a growing share of alternative fuel vehicles (AFVs), and a great importance to the UK economy. However, extensive insight into the vehicle characteristics that influence prices and sales in the UK car market is lacking. To provide such insight, this paper constructs a novel, extensive and contemporary dataset for the entire UK car market, for the period 2008-2018. This dataset represents over 99% of the market and includes a wide range of information on car characteristics, attributes, equipment, prices and sales. A hedonic pricing model is then applied, using the adaptive Lasso, ordinary least squares, weighted least squares, quantile, and penalized weighted quantile regressions. The main goals of the paper include: finding the key car characteristics that influence market prices, comparing these characteristics between conventional vehicles (CVs) and AFVs, constructing quality constant hedonic price indices, and finding the most important car characteristics for AFV consumers. The results of the research show that the key characteristics influencing car prices are vehicle size/weight, performance, safety ratings, range and lower interior noise. In terms of equipment, these are automatic air conditioning, full electric mirrors, automatic wipers, rear-view camera and infotainment display, but the effect gets weaker as car prices increase. AFV prices are significantly more sensitive to changes in features than CV prices, especially to changes in performance, emissions and extra equipment. Analysis of the hedonic price indices showed that about 70% of the increase in car prices between 2008 and 2018 in the UK car market was caused by improvements in vehicle quality, causing a rise in car prices of about 2.5% per year. Additionally, the AFV features found the most important to consumers are performance, maximum range, environmental friendliness and size/comfort.

1. Introduction

1.1. Vehicle Market Introduction

Automobiles occupy a crucial role in the daily lives of the majority of people, and therefore represent an interesting and beneficial topic for research in many academic fields. Recently, increased attention has been paid to the car market by researchers, governments, and the public, as the market is entering a unique stage of transformation and development. With the focus on environmentally friendly vehicles being greater than ever, we are seeing a boom of alternative fuel vehicles (AFVs) after decades dominated by conventional vehicles (CVs). The number of available AFV models in the UK increased from under 1% in 2008 to just under 8% in 2018, while for purely electric vehicle models, this increase was from 0.4% in 2008 to almost 3% in 2018. Furthermore, the introduction of the first commercially sold hydrogen[1] car, the Hyundai ix35 FCEV in 2013, and the follow up introductions of Toyota Mirai, Honda Clarity and Hyundai Nexo only further stress the ongoing transformation of the global car market. This can be further amplified in the foreseeable future with the potential introduction of semi and fully autonomous vehicles.

Due to its size in terms of revenues, employment and exports, the car industry stands as a crucial sector of the economy for most major countries. It represents around 12% of exports in Germany (with a turnover of over 400 billion euros), 5% in France, 6% in Sweden and 13% in the Czech Republic (OEC, 2017). The UK is no different as it represents one of the country's most important economic pillars. It directly employs almost 170 000 people and over 820 000 across the wider car industry, has a yearly turnover of more than £82 billion and represents about 14.4% of all the exports (SMMT, 2018). UK has one of the largest car markets in the world (6th), and 2nd in Europe.

Due to the importance of the car market for the economy of many countries, there is a wide and growing literature looking into various aspects of the market. Many researchers apply hedonic pricing regressions to examine the effects of various attributes on prices and sales. However, there has been relatively little work done on the UK car market, despite UK being a very large market and of significant importance to the national economy. Consequently, an extensive, contemporary, and direct empirical insight into the UK car market is missing from the literature, which this paper aims to provide, using an extensive, unique and current dataset and applying the hedonic pricing technique.

1.2. Research Significance

This paper encompasses the following four research questions, answered in the remainder of the work:

1. How do vehicle characteristics, features and equipment influence vehicle price in the UK car market?
2. How does this influence differ between conventional vehicles and alternative fuel vehicles?

---

[1] Hydrogen cars react hydrogen with oxygen in a fuel cell, to run the electric motor and produce water as waste.

3. Constructing the quality adjusted hedonic price index for UK car market, 2008-2018.

4. What are the key vehicle characteristics, features and equipment that influence the sales of AFVs?

The main methods used to answer these questions include: adaptive lasso, ordinary least squares (OLS), weighted least squares (WLS), quantile regressions, weighted quantile regressions and penalized weighted quantile regressions. With the exceptions of Murray and Sarantis (1999), and Crawford and Neary (2019), similar research has not been done for the UK car market. However, Murray and Sarantis (1999) used data on 123 car models starting in 1977, and Crawford and Neary (2019) used data on only 36 models starting in 1988, so both the data sets used are quite old. My contribution is unique, as I constructed and use a completely novel, current and very extensive manually collected database. I have sales, prices and a wide range of characteristics and equipment data for almost every car model in the UK market (the data represents over 99% of the market) in the years 2008-2018. The only missing models are very expensive, high end cars (e.g. Lamborghini or Ferrari) that would skew the data, and unique cars, such as the ones made by Morgan[2]. To the best of my knowledge, this is the first time that anyone has done such an analysis for the UK car market with this extensive and recent dataset. A large and wide dataset brings the benefits of more control variables, more options for inference on independent variables and the possibilities of stronger or completely new conclusions. Further contribution comes from the application of quantile regression to the data, as to my knowledge, this method has not been used before to examine the car market. Different car characteristics may be priced differently for cars in the upper price range, compared to cars in the lower price range. Using just an OLS or WLS regression would not show this phenomenon, as these are only based on the mean of the whole price distribution. Quantile regression that I perform on the other hand allows for different pricing of car characteristics along the overall price spectrum, and thus provides a deeper insight into the UK car market.

The research questions stated above are of interest to many car dealers attempting to set competitive prices, as well as to the government and policy makers. The quality adjusted hedonic price indices are at the forefront of the political debate and on the agenda of many researchers, as pointed out by Reis and Santos Silva (2006). This is because if the effects of quality change are not measured correctly, the ordinary consumer price index can significantly overestimate inflation, which leads to an underestimation of productivity and living standards growth (Boskin et al., 1996). This could also be one of the reasons for productivity growth slowing down in advanced economies (Feldstein, 2017). Consequently, hedonic pricing technique and quality adjusted hedonic indices are of great general interest.

[2] Morgan is a British car manufacturer which specializes in producing old-style, historic vehicles. All vehicles are made using wood, and entirely assembled by hand. The company only produces around 850 cars every year.

The rest of the paper is organized as follows. Section 2 reviews the relevant literature for hedonic modelling in general and hedonic pricing regressions in car markets in particular. Section 3 explains the unique dataset, its construction procedure and what information it contains. Section 4 gives an overview of the UK car market through the use of descriptive statistics. Section 5 delves into the methodology used to answer the four main research questions. Section 6 provides the results of the research and discussion. The final section 7 then concludes the paper.

2. Literature Review

The relevant literature of hedonic pricing regressions can be roughly divided into three main strands. First, one with similar goals as this paper, that applies hedonic regression techniques to different vehicle markets. Second, one that also applies hedonic techniques, but focuses on a different composite good, (e.g. houses). These works don't investigate car markets, but reviewing them is useful as their methodology can be valuable for researchers looking into the vehicle markets as well. Third, one that looks into the more theoretical aspects of hedonic regression techniques, such as consistency or identification. These strands are now going to be reviewed in more detail.

2.1. Hedonic Regressions on Vehicle Markets

The first use of the hedonic modelling procedure, and the coining of the term "hedonic" was done by Court in 1939, however, the paper that popularized the use of hedonic regressions was written by Griliches (1961). Griliches (1961) wanted to explore whether it would be feasible to implement, and whether it would provide worthwhile results. He took note of the fact that general price indices do not consider changes in overall quality of the products and he wanted to alleviate this important problem. Therefore, he investigated the US car market in the years 1930, 1950 and 1954-1960 by regressing the vehicle characteristics on vehicle price, and discovered the marginal percentage effects on price to construct quality adjusted or "hedonic" price indices. The main finding is that the increase in prices of the US car market in 1950s can be explained purely by improvements in vehicle quality.

Calculation of hedonic price indices for a car market was more recently done by e.g. Reis and Santos Silva (2006). The paper aimed to examine how quality change affects the price indices for new cars in the Portuguese car market in 1997-2001. The interest in quality adjusted price indices was amplified with a report by Boskin et al. (1996), and more recently by Feldstein (2017), which stressed that inaccurate price indices could negatively affect the perception of the economy and policy decisions. This bias of the consumer price indices was further explored by Wayne and Rodriguez-Palenzuela (2004). Applying a weighted least squares regression, the authors find that the price index bias may be about 0.15 percentage points per year.

This same question was explored by Murray and Sarantis (1999), using a panel dataset of the UK car market (1977-1991). Similarly to Reis and Santos Silva (2006), weighted least squares method is applied, where each car model is weighted by the market share of that model, as in Hogarty (1975). A recent example of the same approach for the UK car market is in Crawford and Neary (2019), with data for 36 car models over the 1988-1995 period. Murray and Sarantis (1999) conclude that car price is most affected by performance, luxury items, economy and maneuverability, while durability and safety were either insignificant, or had the wrong sign.

A popular area of the car market hedonic modelling literature focuses on cost, value and characteristics of safety in vehicles. Mount et al. (2000) examine automobile safety and aim at obtaining the value of statistical life for children and the elderly, similarly as Blomquist et al. (1996). This is done by inspecting the amount that families with children and elder member spend on car safety, in comparison to families without children or the elderly. Using US survey data from 1995, the authors calculate that a child's value of statistical life is slightly more than adult's, which is slightly more than elder's. Additionally, a negative correlation between risk of fatal accident and price was discovered, similarly as in Andersson (2005) for the Swedish car market.

With environmentally friendly transportation being at the forefront of both political and academic debate, many researchers look into the pricing of fuel-efficient vehicles. Applying hedonic pricing regressions on the Swiss car market over the period 2000 to 2011, Alberini et al. (2014) attempt to find whether fuel economy labels have any additional effect on car prices. This depends on whether consumers value fuel efficiency benefits or "misperceive" them (Anderson et al., 2011). Studies found that producers set higher prices for goods that would otherwise be identical, if they have fuel economy label (Houde, 2018).

2.2. Hedonic Regressions on Other Composite Goods

A popular area of research where the hedonic pricing technique is frequently used is house prices. For example, Sirmans et al. (2005) investigate the marginal contributions that house characteristics have on prices. Similarly to the hedonic modelling used in the car markets, linear or semi-logarithmic[3] models are typically used, and these models assume that consumers derive utility and value from the many different housing attributes (Malpezzi et al., 1980). However, while many papers also choose to use the OLS and WLS, a number of researchers decided to rather apply a quantile regression. For example, Zietz et al. (2007) argue that the observed variation of results from previous studies may be explained by the fact that attributes are not priced equally across the house prices distribution.

[3] The dependent variable would typically be the natural logarithm of price ($\text{Ln}\, Price$). The coefficient estimates would then provide the percentage change in price for a marginal (one unit) change in each variable.

Therefore, they apply a quantile regression to find the implicit prices of housing attributes for different points of the house prices spectrum, and alleviate the insufficiencies of an OLS regression. Additionally, as quantile regressions use the entire sample, truncation is not a problem (Heckman, 1979). The same variations in the valuation of different house characteristics across the price distribution are found by Coulson and McMillen (2007) using their hedonic indices, and this effect would likely apply to the car market as well.

2.3. Theory and Hedonic Regressions

Several authors go over the theory behind hedonic pricing technique and hedonic price indices. A primary example is the work of Rosen (1974), who formalized this research area on a theoretical level. The model that Rosen (1974) presented assumes that producers create a composite good (such as a car) in such a way, that it has the ultimate characteristics that consumers desire. This conceptual foundation of Rosen was later used in many subsequent studies.

A useful review of hedonic regression methods was done by de Haan and Diewert (2013). The authors look into the standards of hedonic modelling, estimation and the theory behind them, review some of the practical issues that can be encountered, the methods to calculate hedonic price indices, as well as their advantages and disadvantages. The calculation of hedonic price indices is also examined by Diewert (2003). The author investigates whether quantity weights should be used in hedonic regressions, and concludes that the use of weights when modelling hedonic regressions can be highly recommended. Diewert (2003) also states that in general, linear hedonic regressions should be avoided as they are “difficult to justify on theoretical grounds”. A detailed discussion and explanation of a wide range of hedonic techniques can be found in the book by Aizcorbe from 2014. The book can be highly useful for a researcher attempting to apply hedonic methods, as it contains hedonic theory, basic intuition, explanations of practical approaches, their advantages and disadvantages, as well as a large array of examples for each approach.

Further hedonic regression theory is investigated by Machado and Santos Silva (2001, 2006), through examining the implications of identification with averaged data for hedonic regression studies. The authors point out that when models with averaged data, and an endogenous selection into groups, are estimated, WLS should be used as weights not only provide efficiency, but consistency as well. Finally, Machado and Santos Silva (2006) suggest that the question of endogeneity can be empirically tested by comparing the OLS and WLS estimators (White, 1980).

3. Data

This paper uses a novel and unique dataset in order to provide answers to the main research questions. As it is entirely new and I am the sole creator, the dataset hasn’t been used in any research paper before and will thus provide a unique and contemporary insight into the UK car market. The entire dataset was manually collected (data point by data point) from a wide range of both physical and online sources. It contains extensive information on a wide range of vehicle attributes for almost all of the car models in the UK car market, for the time period 2008-2018. Apart from sales and prices, there is information on vehicle physical characteristics (such size or maximum speed), electric vehicle characteristics (e.g. battery capacity and charging time), vehicle safety attributes (such as safety ratings or the presence of airbags), vehicle equipment (e.g. presence of electric windows or the number of speakers) and others (such as country of origin, or the vehicle's segment[4]). As such, the dataset contains information on hundreds of car models from 46 different car manufacturers and 11 different countries. Each year of my dataset represents more than 99% of the UK car market, and the only vehicles that are not included are special niche vehicles, such as very fast and expensive sports cars sold only in small quantities. Including such distinct vehicles into the dataset would increase the number of data points and make my data set contain the UK car market in its entirety, but such an approach is ill-advised, as these vehicles represent extremes, and would most likely skew the overall data and thus, the final results. Altogether, the dataset contains 2733 observations (where each one car model in one year represents one observation), 85 different vehicle characteristics, 142 different variables and therefore, more than 400 000 unique data points. The data collection process and variable construction will now be examined in more detail.

3.1. Collection of the Data

As in any car market, vehicle models enter and exit the UK car market often. Sometimes, new models enter as a new generation of an already existing model, which at the same time exits the market. For example, Ford Focus was in the UK car market for the entire 2008-2018 period examined, but it went through two generation changes and hence changes in characteristics and equipment. On top of that, each vehicle model also usually has an extensive range of trim[5] levels, where the differences in characteristics, equipment and price between the most basic trim and the top trim are vast. Therefore, it had to be ensured that models will be consistent across years and thus comparable. For example, comparing the most basic trim of Škoda Fabia with the top trim of BMW i3 would not yield much useful information, even though both are of segment B (small cars). Consequently, to ensure consistency, only the physical characteristics, equipment and price of the most basic trim are added for each vehicle model in the dataset. Additionally, each change of vehicle generation in the market is accounted for with updated characteristics, equipment and price of the most basic trim of that vehicle. This allows for a direct comparison of all the car models in the market, comparison of the market at different points in time, and hence, a meaningful regression analysis and results inference.

[4] Cars are generally divided into segments, based on their type and characteristics, e.g. A: mini cars, B: small cars, etc.

[5] Trim levels for a given car model specify which equipment and special features are included as standard. Higher trims are offered for a higher price tag, but they have more equipment and features fitted as standard, compared to lower trims.

As stated, the information for this dataset was collected from a wide range of sources. The first attributes to collect were sales and price. Sales represent the popularity of each model, its market weight, and its performance in the market, while price is of upmost interest to both consumers and dealers, and should embody the "richness" of characteristics and equipment of each car model. The data for sales (in units sold) comes from the manufacturers' websites, the Society of Motor Manufacturers and Traders (SMMT) and from the Mark Lines automotive data portal. Additionally, the market share of each model for each year, the vehicle's segment, the model manufacturer, and the vehicle's country of origin (set as the country that a consumer perceives the car to be from) are also obtained. Data for car prices comes from purchasing the old physical issues of New Car Prices and Specs Guide of the Auto Express magazine, where the basic trim list prices of each model for each year were manually transferred into the dataset.

Next, the actual physical characteristics (e.g. performance, efficiency, size), equipment and safety features of each vehicle model had to be acquired. Since no single source provides complete information on all possible physical characteristics, information from several sources was combined in order to fill gaps in the data and obtain a complete data set. A large portion of the data on vehicle physical characteristics was collected from manufacturers' websites, the Auto Express car reviews magazine, and the Parkers online database of vehicle assessments and specifications. Since these sources often provided only limited information on high-end and electric vehicles, BBC's Top Gear car magazine and the UK Database of Electric Vehicles respectively are used in these cases to collect the missing data.

The equipment of each model represents the level of special features and general gear that comes as standard on its most basic trim. This can increase the safety of a vehicle (such as ABS[6]), comfort (e.g. electric windows), or provides extra information and entertainment (e.g. display). This information was collected using the Parkers database, the Auto Express vehicle specifications magazine, and the Canadian Auto123 website for vehicle reviews.

Finally, car safety characteristics represent the safety ratings of various vehicle attributes, as well as the presence of critical safety gear (such as airbags). This information is obtained from the European New Car Assessment Programme[7] (Euro NCAP), which performs independent crash tests of new cars, after which it grants them a safety rating of various attributes (e.g. driver safety, child safety, etc.). Euro NCAP also provides information on the presence of different safety gear (e.g. airbags, pretensioners) on the most basic trim level of each car, and whether additional critical gear can be purchased as an option, or not at all.

[6] The anti-lock braking system is an anti-skid safety feature. It prevents the wheels from completely locking up during braking, which allows them to maintain tractive contact with the road surface and not go into a skid.

[7] Based in Belgium, it was founded in 1996 for the UK Department for Transport, and is backed by a number of European governments, the International Automobile Federation (FIA), as well as the European Union.

3.2. Construction of Variables

Most of the physical characteristics, equipment and safety attributes collected come directly as numbers, and therefore can be readily used as variables. However, several characteristics are not, while others suffer from missing data points. These missing values would cause issues in any subsequent regressions, as any observation that has at least one missing value is automatically removed from a regression, hence hampering a complete examination of the UK vehicle market due to the loss of power and precision. Therefore, these issues must be investigated and solved.

Since the functional form has to be semi-logarithmic for the purposes of this paper, natural logarithm of sales and prices as dependent variables are constructed from the data on sales and prices. Furthermore, dummy variables for each car segment (segment A being the base), manufacturer (Vauxhall being the base), and country of origin (UK being the base), are created. Additionally, year dummies are constructed for each year between 2008 and 2018 (2008 being the base), equal to 1 if a car model was available in the UK car market in a given year, and 0 otherwise.

Most physical characteristics of the car models do not require any manipulation, with the exception of engine type (e.g. diesel) and whether the car is a CV or an AFV. Dummy variables are created for these (with petrol engine and CV category respectively being the base). Next, a horsepower per ton variable is constructed from engine power and weight, to consider how much weight the engine has to pull. Furthermore, a size variable is constructed by multiplying length, width and height of each model. Since the equipment available for each observation is already imputed into the dataset as dummy variables, no extra variables are constructed for vehicle equipment.

Finally, a couple of variables have missing values for several observations, and these observations would be completely removed from any regression, even though they have useful information about the UK car market for the remaining 141 variables. The safety attributes from Euro NCAP suffer frequently from this issue as, for example, a two-seater sports car will have missing data for rear airbags, simply because a two-seater doesn't have any rear compartment. Therefore, to take into account the number of seats and avoid missing data, a variable "airbags per seat" is constructed, with the same approach for seatbelt pretensioners and loadlimiters. Lastly, the average Euro NCAP safety rating and interior noise variables suffered from missing data, and thus the regression imputation method was used to complete the missing data. This method uses complete information to impute data, by replacing the missing values with predicted scores from a regression equation, and thus completes the dataset in its entirety.

4. Descriptive Statistics

Looking into the dataset descriptive statistics allows for an examination of how the UK car market (and thus the CV and AFV models) developed between 2008 and 2018.

4.1. Physical Characteristics and EV Attributes

Over the last 11 years, there has been a significant increase in the average quality of vehicles in terms of physical attributes, especially in terms average car performance and efficiency (as can be seen in figure 1). The average engine power increased from 124 bhp to 155 bhp, while the average maximum speed increased from 185 km/h to 195 km/h, and the time it takes to accelerate from 0 km/h to 100 km/h decreased from 12.13 seconds to 10.51 seconds. This striking increase in raw power and performance is matched by an equally significant improvements in efficiency and positive environmental impact. Average fuel consumption per 100 km decreased from 6.72 liters to 5.72 liters, and average CO_2 emissions decreased from 166 g/km to 136 g/km. Furthermore, even though the average vehicle became more powerful while also using less fuel, the average maximum range still saw a noteworthy improvement, from 921 km to 988 km. This and other improvements over the 11-year period examined mean that there have been large improvements in vehicle quality, as also seen by the steady increase in car prices in figure 1.

			Vehicle Physical Attributes												EV-Specific Attributes		
Year	**Statistic**	**Price (£)**	**Displacement (cm³)**	**Engine Power (bhp)**	**Maximum Speed (km/h)**	**Acceleration (0 to 100 km/h in seconds)**	**Interior Noise (dB)**	**Fuel consumption combined (l/100km)**	**Combined CO_2 (g/km)**	**Length (mm)**	**Seating Capacity**	**Trunk Capacity (liters)**	**Fuel Tank Capacity (liters)**	**Range (km)**	**Battery Capacity (kWh)**	**Charge Power (kW)**	**Vehicle Consumption (Wh/km)**
2008	Mean	20 061	1825	124	185	12,13	66,87	6,72	166	4375	4,93	408	60,70	921	9,60	3,70	125,00
	Median	16 905	1599	112	180	12,30	67,00	6,60	160	4428	5,00	405	60,00	863	9,60	3,70	125,00
	Standard Dev.	8 684	478	44	21	2,19	1,99	1,38	34	351	0,54	140	11,05	175	0,00	0,00	0,00
2009	Mean	20 254	1823	126	186	12,07	67,12	6,59	163	4368	4,93	403	60,23	934	9,60	3,70	125,00
	Median	16 802	1598	111	180	12,20	67,00	6,40	157	4417	5,00	402	60,00	884	9,60	3,70	125,00
	Standard Dev.	8 965	488	45	21	2,18	2,08	1,39	35	349	0,54	137	10,98	177	0,00	0,00	0,00
2010	Mean	21 576	1822	127	187	11,92	66,92	6,46	159	4382	4,90	399	60,31	955	16,00	3,70	64,00
	Median	17 090	1598	113	182	11,90	67,00	6,10	149	4426	5,00	397	60,00	896	16,00	3,70	64,00
	Standard Dev.	9 830	501	44	20	2,03	2,06	1,37	34	342	0,58	139	11,22	186	0,00	0,00	0,00
2011	Mean	22 875	1824	129	187	11,84	66,79	6,29	154	4372	4,87	399	60,08	964	22,00	4,43	138,75
	Median	17 990	1598	112	182	11,90	67,00	6,00	140	4407	5,00	385	60,00	906	16,00	3,70	161,50
	Standard Dev.	10 681	522	48	21	2,14	2,19	1,46	37	341	0,55	138	11,27	207	9,00	1,09	37,38
2012	Mean	23 116	1815	130	187	11,79	66,90	6,06	147	4371	4,87	389	58,79	965	28,33	4,77	140,50
	Median	18 388	1594	111	181	11,95	67,00	5,70	137	4407	5,00	385	59,50	914	19,00	3,70	161,50
	Standard Dev.	10 614	546	49	21	2,18	2,24	1,41	35	346	0,56	126	11,29	212	14,44	1,49	32,33
2013	Mean	23 223	1782	130	187	11,61	66,68	5,85	141	4359	4,82	392	57,64	967	31,65	7,70	145,88
	Median	18 035	1586	110	180	11,90	67,00	5,50	134	4385	5,00	385	57,00	918	31,00	5,15	162,00
	Standard Dev.	10 964	544	51	22	2,28	2,35	1,33	33	336	0,55	124	10,90	217	14,15	4,40	26,94
2014	Mean	24 195	1779	134	189	11,33	66,59	5,77	139	4372	4,82	396	57,28	978	33,89	8,27	157,57
	Median	19 221	1586	113	182	11,50	66,00	5,40	130	4406	5,00	391	56,00	924	40,00	6,60	162,00
	Standard Dev.	11 615	539	53	23	2,29	2,37	1,31	32	330	0,55	123	10,80	220	13,62	4,70	11,47
2015	Mean	26 020	1822	142	191	11,04	66,43	5,82	139	4386	4,82	403	57,29	977	35,87	9,07	163,83
	Median	19 795	1586	114	185	11,25	66,00	5,35	130	4402	5,00	402	55,00	920	40,50	7,00	165,00
	Standard Dev.	13 584	594	60	24	2,41	2,45	1,35	33	338	0,55	124	10,99	225	13,24	4,96	4,17
2016	Mean	27 525	1816	146	193	10,83	66,20	5,70	136	4404	4,82	410	57,17	993	35,87	9,07	163,83
	Median	20 995	1593	115	185	11,00	66,00	5,30	129	4425	5,00	410	55,00	964	40,50	7,00	165,00
	Standard Dev.	14 607	576	62	24	2,39	2,45	1,31	31	325	0,55	125	10,89	230	13,24	4,96	4,17
2017	Mean	31 016	1842	152	194	10,58	65,98	5,74	136	4427	4,83	421	57,13	987	50,74	11,43	173,29
	Median	21 620	1591	118	188	11,00	66,00	5,30	128	4434	5,00	416	55,00	956	41,00	11,00	168,00
	Standard Dev.	17 773	619	67	24	2,39	2,42	1,34	32	324	0,51	125	10,84	221	28,15	5,92	14,69
2018	Mean	32 341	1842	155	195	10,51	65,90	5,72	136	4435	4,83	422	57,05	988	61,31	11,96	181,29
	Median	22 895	1591	118	190	10,90	65,00	5,30	129	4442	5,00	422	55,00	963	42,20	11,00	168,00
	Standard Dev.	18 479	620	69	24	2,44	2,44	1,30	32	319	0,51	126	10,80	217	30,30	5,47	22,33

Fig. 1 *Example of descriptive statistics for physical and EV-specific characteristics, for all cars in the UK market 2008-2018.*

Looking at alternative fuel vehicles, their share of models increased from 1% to 8%. This increase represents a large increase of hybrids (~0% to 5%) and full EVs (~0% to 3%). This process of gradual switching to cleaner vehicles can also be seen in the development of EV specific physical characteristics in figure 1. Every examined characteristic has substantially improved, some in the realms of hundreds of percent. The average charge power has increased by 223%, and charge time decreased by 21%. And while average vehicle consumption increased by 45%, average battery capacity increased significantly more (by 539%), increasing the average maximum range. All of this reflects the demand for cleaner, less polluting cars with lower effect on global warming and lower carbon footprint.

4.2. Safety Characteristics

Examining the development of safety characteristics in figure 2 shows that vehicles in the UK car market experienced solid improvements in safety features. This is apparent from the Euro NCAP ratings – each average rating either improved or stayed roughly the same. The mean rating for adult safety increased slightly from 86 to 88, while average child safety rose from 74 to 80, and pedestrian safety from 50 to 67. Furthermore, the amount of safety gear also experienced a significant increase. While gear such as front driver airbags and pretensioners was already included as standard on almost 100% of models in 2008, significant improvements went through for side airbags, and rear seatbelt pretensioners and loadlimiters. While in 2008, driver and passenger side airbags for head, chest and pelvis were present only on 84%, 78% and 11% of models respectively, this has increased to 95%, 95% and 33% in 2018. Furthermore, a significant improvement in safety also went through for the rear seats, with side head airbags being offered as standard on 58% of models (increase from 42%). Similarly, rear seatbelt loadlimiters are on 65% of models (increase from 48%), portraying an overall improvement in average safety of car models offered in the UK car market.

		Safety Ratings and Gear										Equipment											
Year	Statistic	NCAP Adult Safety	NCAP Child Safety	NCAP Pedestrian Safety	Front Driver Airbag	Side Head Driver Airbag	Side Chest Driver Airbag	Side Pelvis Driver Airbag	Side Head Rear Airbag	Driver Belt Pre-tensioner	Rear Belt Load limiter	Traction Control	Hill Start Assist	Forward Collision Warning	Traffic Sign Recognition	Display	Cruise Control	Number of Speakers	USB Jack	Parking Sensors	Rear View Camera
2008	Mean	86	74	50	0,99	0,84	0,78	0,11	0,42	1,00	0,48	0,67	0,13	0,02	0,00	0,35	0,51	6,33	0,12	0,21	0,10
	Median	82	80	50	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	10	9	13	0,02	0,26	0,34	0,19	0,49	0,00	0,50	0,44	0,23	0,04	0,00	0,46	0,50	2,12	0,21	0,34	0,18
2009	Mean	85	75	51	0,99	0,83	0,79	0,10	0,41	1,00	0,48	0,69	0,15	0,02	0,00	0,36	0,52	6,33	0,15	0,22	0,11
	Median	85	80	50	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	10	8	13	0,02	0,28	0,34	0,19	0,48	0,00	0,50	0,42	0,26	0,05	0,00	0,46	0,50	2,12	0,26	0,34	0,19
2010	Mean	86	76	53	1,00	0,86	0,81	0,13	0,40	1,00	0,47	0,75	0,17	0,04	0,00	0,40	0,54	6,46	0,20	0,26	0,13
	Median	88	80	50	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	9	7	13	0,01	0,25	0,31	0,22	0,48	0,00	0,50	0,38	0,29	0,07	0,00	0,48	0,50	2,17	0,33	0,38	0,23
2011	Mean	87	77	54	1,00	0,86	0,83	0,13	0,33	0,99	0,43	0,79	0,24	0,04	0,00	0,43	0,54	6,46	0,26	0,29	0,16
	Median	89	80	50	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	9	7	12	0,01	0,23	0,28	0,23	0,44	0,01	0,49	0,33	0,36	0,07	0,01	0,49	0,50	2,12	0,38	0,41	0,27
2012	Mean	87	78	55	1,00	0,89	0,86	0,13	0,32	0,99	0,39	0,83	0,31	0,04	0,01	0,50	0,57	6,44	0,31	0,31	0,19
	Median	89	80	53	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	0,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	8	6	13	0,00	0,19	0,24	0,22	0,44	0,01	0,48	0,28	0,43	0,08	0,02	0,50	0,49	2,08	0,43	0,43	0,31
2013	Mean	87	79	58	1,00	0,90	0,89	0,12	0,33	0,98	0,40	0,85	0,33	0,06	0,01	0,54	0,59	6,39	0,36	0,34	0,23
	Median	89	80	58	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	1,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	7	6	12	0,01	0,18	0,20	0,21	0,44	0,04	0,48	0,25	0,44	0,11	0,02	0,50	0,48	2,10	0,46	0,45	0,36
2014	Mean	87	80	59	1,00	0,93	0,91	0,12	0,33	0,98	0,42	0,90	0,35	0,07	0,01	0,59	0,63	6,47	0,43	0,36	0,25
	Median	89	80	62	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	1,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	7	5	11	0,01	0,13	0,16	0,22	0,44	0,04	0,49	0,19	0,46	0,13	0,02	0,48	0,46	2,14	0,49	0,46	0,38
2015	Mean	87	79	62	1,00	0,94	0,92	0,17	0,40	0,98	0,49	0,91	0,39	0,08	0,03	0,63	0,68	6,64	0,47	0,41	0,31
	Median	89	81	65	1,00	1,00	1,00	0,00	0,00	1,00	0,00	1,00	0,00	0,00	0,00	1,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	7	6	11	0,01	0,11	0,14	0,29	0,48	0,04	0,50	0,17	0,48	0,15	0,05	0,47	0,43	2,22	0,50	0,48	0,43
2016	Mean	87	80	64	1,00	0,95	0,94	0,24	0,47	0,98	0,55	0,94	0,42	0,11	0,04	0,69	0,71	6,82	0,50	0,45	0,39
	Median	89	81	66	1,00	1,00	1,00	0,00	0,00	1,00	1,00	1,00	0,00	0,00	0,00	1,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	6	5	10	0,01	0,10	0,11	0,37	0,50	0,04	0,50	0,11	0,49	0,20	0,08	0,43	0,42	2,30	0,50	0,50	0,48
2017	Mean	88	80	66	1,00	0,96	0,96	0,33	0,57	0,99	0,64	0,97	0,40	0,13	0,07	0,74	0,76	6,99	0,49	0,49	0,41
	Median	89	82	67	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	0,00	0,00	0,00	1,00	1,00	6,00	0,00	0,00	0,00
	Standard Dev.	6	5	9	0,01	0,09	0,09	0,44	0,49	0,03	0,46	0,06	0,48	0,23	0,14	0,38	0,37	2,32	0,50	0,50	0,48
2018	Mean	88	80	67	1,00	0,95	0,95	0,33	0,58	0,99	0,65	0,96	0,41	0,14	0,07	0,76	0,77	7,06	0,49	0,50	0,41
	Median	90	82	67	1,00	1,00	1,00	0,00	1,00	1,00	1,00	1,00	0,00	0,00	0,00	1,00	1,00	6,00	0,00	1,00	0,00
	Standard Dev.	6	5	8	0,01	0,09	0,09	0,44	0,49	0,03	0,45	0,07	0,48	0,24	0,14	0,37	0,36	2,41	0,50	0,50	0,48

Fig. 2 *Example of descriptive statistics for safety ratings, gear and equipment, for all cars in the UK market 2008-2018.*

4.3. Equipment

Overall, the quality of cars has increased substantially over the last 11 years. The average amount of equipment fitted as standard increased significantly – every equipment type is fitted to considerably more models as standard in 2018, compared to 2008, as seen in figure 2 above. The largest increases can be seen for traction control (from being standard on 67% of models in 2008 to 96% in 2018), hill assist (13% to 41%), display (35% to 76%), cruise control (51% to 77%), USB jack (12% to 49%), parking sensors (21% to 50%) and rear view camera (10% to 41%). Furthermore, equipment that was not standard on almost any model in 2008 also significantly increased. For example, while in 2008 forward collision warning and traffic sign recognition was present as standard on only 2% and 0% of car models respectively, in 2018 this increased to 14% and 7%. All in all, cars in the UK market experienced a vast improvement in quality over the period examined. Their overall performance and efficiency have increased, while providing higher levels of safety and significant improvements in equipment and gear fitted as standard.

5. Methodology

The methodology used to answer the main research questions will now be described in the following sections.

5.1. Effect of Different Car Attributes on Price and The Key Vehicle Characteristics Influencing the Sales of AFVs

5.1.1. Hedonic Pricing

Most of the research questions of this paper use and apply the revealed preference hedonic pricing technique, formalized by Rosen (1974). The hedonic technique assumes that there exists differentiated, or composite goods (such as cars), which are made out of a set of characteristics or attributes, and each characteristic is considered to have its own unique (although unobservable) price. The market demand and supply of cars determines each characteristic's marginal contribution to overall price, and hedonic regressions can then be used to find out the value of these unobserved prices (de Haan and Diewert, 2013). If we allow Z to represent a car, which is made out of *n* measurable characteristics z_i, such as $Z = (z_1, z_2, \dots, z_n)$, then the hedonic price function can be defined as:

$$P(Z) = P(z_1, z_2, \dots, z_n) \tag{1}$$

where $P(Z)$ is the market price of the car.

These characteristics are objectively measurable (e.g. length, or engine horsepower) and information about them is publicly available, however, each consumer is likely to value these characteristics differently.

Following Andersson (2005), a consumer's utility function may be expressed as $U(x, Z)$, where *x* are all the goods consumed, except for the car *Z*. The standard assumptions of utility functions are applied, such as e.g. strict concavity. If we normalize the price of all the goods *x* to unity (where $x = money$), then the consumer's budget constraint is $y = x + P(Z)$, where *y* is income. Each consumer maximizes his utility based on the budget constraint, and thus "chooses the amount of each characteristic where marginal costs are proportional to the marginal rate of substitution between the characteristics and money": $P_{z_i} = \frac{U_{z_i}}{U_x}$ (Andersson, 2005). The amount that a consumer is willing to spend for a car with unique characteristics is defined by $\theta(Z; y, u)$, where $u = U(y - \theta, Z)$. As θ is the willingness to pay, it will be increasing with better characteristics z_i, but at a decreasing rate. When θ is differentiated with respect to the characteristics z_i as in Rosen (1974), the marginal rate of substitution between the characteristics and money is obtained: $\theta_{z_i} = \frac{U_{z_i}}{U_x}$. Since θ is the price that the consumer is willing to pay for a car Z with specific set of characteristics z_i, and $P(Z)$ is the minimum price of the car with such characteristics, utility is maximized when:

$$\theta(Z^*; y, u^*) = P(Z^*) \tag{2}$$

$$\theta_{z_i}(Z^*; y, u^*) = P_{z_i}(Z^*), \qquad i = 1,2, \dots, n \tag{3}$$

where Z^* and u^* are the optimum quantities.

This means that the consumer consumes at a point where the price he is willing to pay for an amount of characteristic z_i is exactly equal to its market (minimum) price (i.e. point of tangency). An example of this equilibrium for characteristic z_1 can be seen in figure 3 (but the same applies also for all other characteristics $z_2, z_3, \dots, z_n$). The problem on the producer side is symmetric to the described consumer side.

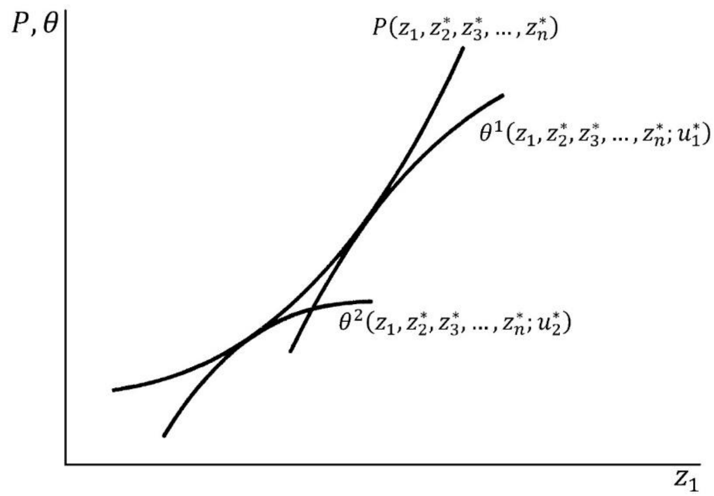

Fig. 3 *Equilibrium of the consumer side for z_1 – two consumers with different willingness to pay (Rosen, 1974).*

5.1.2. Functional Form

Before we run any regressions and find out the effect of different vehicle attributes on vehicle price, the functional form of the econometric hedonic model has to be decided. The theory doesn't explicitly favor any one functional form over the other, and generally concludes that decisions should be made empirically, on a case-by-case basis by the researcher (Andersson, 2005; Reis and Santos Silva, 2006). Empirically, researchers have explored and applied a wide range of different functional forms, however, the three best known hedonic specifications are the fully linear, the semi-logarithmic and the fully logarithmic model. To make a final decision, I follow Schamel and Anderson (2003), and Reis and Santos Silva (2006), and apply the heteroskedasticity robust RESET[8] test to the three functional forms. The results (in Appendix A) showed that the semi-logarithmic model (e.g. $\ln P_j = \beta_0 + \sum_{i=1}^{I} \beta_i z_{j,i} + \epsilon_j$) is the most appropriate, with lowest levels of heteroskedasticity and omitted variable bias. This result is well supported by the literature, as hedonic pricing models in the semi-logarithmic functional form are often used by economists and

[8] The Ramsey Regression Equation Specification Error Test (RESET) is a general model specification test. It has power against several alternatives and is robust to both non-normality and heteroskedasticity (Reis and Santos Silva, 2006).

are well justified from the econometric point of view (e.g. Griliches, 1961; Andersson, 2005). Moreover, for high-tech goods, the semi-logarithmic functional form is usually the most appropriate, as it has the highest chance to reduce heteroskedasticity (since prices are usually distributed log-normally) (Diewert, 2003).

5.1.3. Adaptive Lasso

Before running any regressions, there is an issue of too many characteristics present. There are altogether 142 variables available for regression and such large number of variables can lead to a number of problems, such as multicollinearity, parameter estimation issues and thus subsequent problems in interpretation, or overfitting (He et al., 2012). Especially multicollinearity between various vehicle characteristics is a well-known issue in hedonic regressions (de Haan and Diewert, 2013; Alberini et al., 2014). Therefore, to improve the estimation with such large number of variables, it is important to only include those that are important and relevant on economic grounds and that do not suffer from high levels of multicollinearity. Thus, I apply the adaptive least absolute shrinkage and selection operator (adaptive Lasso) methodology before every regression run, and for each research question.

The adaptive Lasso is an econometric regularization[9] method that is used to prevent multicollinearity and for variable selection. When the problem of too many variables arises, adaptive Lasso can be used to find out which variables are relevant. The method works in such a way that those variables that are irrelevant are shrunk exactly to zero, while the important ones aren't. These selected variables can be used further in the regressions, while the ones shrunk to zero can be concluded to be unimportant. Taking an OLS model, the adaptive Lasso estimator is:

$$\hat{\beta}_n^{AL} = \arg\min_{\beta} \sum_{j=1}^{n} \left(y_j - x_j\beta\right)^2 + \lambda_n \sum_{i=1}^{I} \lambda_{n,i} |\beta_i| \tag{4}$$

where y_j is the dependent variable, $x_j = (x_{j,1}, x_{j,2}, \ldots, x_{j,I})'$ is the vector of independent variables, $\lambda_n > 0$ is the tuning parameter, $\lambda_{n,i} = \frac{1}{(|\hat{\beta}_{n,i}|)^{\gamma}}$ is the adaptive weights vector, $\hat{\beta}_{n,i}$ is an initial estimate of the coefficients, γ is a positive constant for adjustment of the adaptive weights vector, set between $\frac{1}{3}$ and $\frac{10}{3}$, β_i are the estimated coefficients.

The most important parameter of equation (4) is the tuning parameter λ_n. This chosen parameter determines the level of penalization applied to coefficients – i.e. the boundary relevance above which the coefficient is not shrunk to zero. The adaptive weights vector $\lambda_{n,i}$ adjusts the regularization and applies stronger penalization to smaller coefficients. This is one of the reasons why the adaptive Lasso is a superior method compared to the ordinary Lasso. Additionally, the adaptive Lasso method has the oracle properties[10], and is much less sensitive to the choice of the tuning parameter

[9] The process of regularization introduces additional information, in order to prevent the problem of overfitting.

[10] Oracle properties state that a procedure must identify a right subset of true variables, and it has to have optimal estimation rate.

λ_n. Further methods, such as the variance inflation factor (VIF) analysis, were also run to ensure that multicollinearity is reduced as much as possible, without the risk of significant omitted variable bias.

5.1.4. Ordinary and Weighted Least Squares Regressions

The goal of my first research question is to determine the relationship of various vehicle characteristics with vehicle price, and in a hedonic pricing model, estimation of regression coefficients is usually done using the ordinary least squares (OLS) (Czembrowski and Kronenberg, 2016). As the classical OLS regression is often used as a starting point, I start with estimating an OLS regression for the entire UK car market and the entire period, using the optimal semi-logarithmic functional form found above. The OLS regression takes the form of:

$$\ln P_j = \beta_0 + \sum_{t=1}^{T} \beta_t d_{j,t} + \sum_{i=1}^{I} \beta_i z_{j,i} + \epsilon_j, \qquad j = 1,2,\dots,n \tag{5}$$

P_j is the market price of car model j, $d_{j,t}$ is a dummy variable equal to 1 if car j was in the market in year t, $z_{j,i}$ is the value of characteristic i for car j, β_t is the coefficient for time dummy t, β_i is the coefficient for characteristic i, ϵ_j is the error term, T is the number of time periods, I is the number of vehicle characteristics, n is the number of cars.

Therefore, this allows for finding out the effect of various vehicle characteristics on vehicle price, while controlling for changes in time. However, as OLS typically suffers from extensive heteroskedasticity of the error term (as found using the Breusch-Pagan and the White tests), the regression was recalculated using the heteroskedasticity-robust standard errors. A similar method was used for further research questions; but when finding how the effect of vehicle characteristics on car prices differs between CVs and AFVs, the dataset was divided into CV and AFV sub datasets. Furthermore, when pinpointing the most important car characteristics for AFV consumers, the OLS model is:

$$\ln S_a = \beta_0 + \sum_{t=1}^{T} \beta_t d_{a,t} + \sum_{i=1}^{I} \beta_i z_{a,i} + \sum_{e=1}^{E} \beta_e x_{a,e} + \epsilon_a, \qquad a = 1,2,\dots,A \tag{6}$$

S_a is the no. of sales of AFV model a, $d_{a,t}$ is a dummy variable equal to 1 if AFV a was in the market in the year t, $z_{a,i}$ is the value of vehicle characteristic i for AFV a, $x_{a,e}$ is the value of electric vehicle characteristic e for AFV a.

A similar question aiming to find the key vehicle characteristics that influence the sales of alternative fuel vehicles was examined in previous papers using stated preference data. One of the goals of this study is to see whether previous conclusions hold using the constructed revealed preference dataset for the UK car market.

Although the OLS regression is a useful starting point, it cannot be considered optimal as it values each observation equally. In reality, different cars have different sales, market share and consumer popularity. To account for this, each observation should be weighted by its relative importance in the market (optimally its sales, or market share). This is usually achieved by applying the weighted least squares (WLS) regression, as proposed by Griliches (1961). Its use is also frequently recommended in the presence of heteroskedasticity, as it will lead to consistent, more efficient estimates (Machado and Santos Silva, 2001). Most frequently used form of weights is the car market share (e.g. Murray and Sarantis, 1999; Crawford and Neary, 2019). Therefore, the WLS regression is applied to research questions 1 and 2 in addition to the OLS, using the market share of each car model as weights. The WLS estimator is $\hat{\beta}_{WLS} = (\boldsymbol{X'WX})^{-1}\boldsymbol{X'WY}$, where $\boldsymbol{W}$ is a matrix with weights w_i on the diagonal and zeroes everywhere else.

5.1.5. Quantile Regression

Although that the WLS regression has many advantages over the OLS, it cannot account for the fact that vehicle characteristics may be valued differently and have a different effect at different points of the car price distribution. The implicit/hedonic prices of vehicle characteristics at high car prices (upper quantiles[11]) may be affected by the high willingness to pay of the rich consumers. Liao and Wang (2011) call this the quantile effects. As the OLS and WLS regressions would not provide any useful information on this phenomenon, a quantile regression is applied in order to get a more complete picture about the effect of various vehicle characteristics on car prices.

The quantile regression was formalized by Koenker and Bassett in their 1978 seminal paper. Quantile regression allows the estimated coefficients to vary with the quantile chosen, and therefore it represents a perfect tool to employ when quantile effects are present. Additionally, quantile regression benefits from higher robustness to outliers and non-normal errors, and is semiparametric. A standard quantile regression estimator can be defined as:

$$\hat{\beta}_q = \arg\min_{\beta_q \in \mathbb{R}} \sum_{j=1}^{n} \rho_q\left(y_j - x_j'\beta_q\right) \tag{7}$$

where $\hat{\beta}_q$ is the vector of coefficient estimates, q is the quantile that is going to be estimated, with $q \in (0,1)$, n is the number of observations, ρ_q is the loss function, given by $\rho_q(u) = u\big(q - I(u < 0)\big) = \begin{cases} u(q-1), & u < 0; \\ uq, & u \geq 0. \end{cases}$

It can be argued that a similar effect to the quantile regression can be achieved by dividing the dependent variable (car prices) into subsamples according to the unconditional distribution, and then applying a standard OLS regression. However, this approach should be avoided in favor of the quantile regression (which employs the full data sample), since truncation of the dependent variable can create biased parameter estimates (Heckman, 1979).

As noted by Machado and Santos Silva (2013), the estimation of quantile regression standard errors may be problematic due to the presence of heteroskedasticity, and I thus employ the MSS quantile regression

[11] Quantiles are cut-off points that divide a distribution into continuous intervals, e.g. quartiles creating 4 parts (25%).

heteroskedasticity test proposed by Machado and Santos Silva (2000). This is computed as n times the R^2 of the auxiliary regression of $\rho_q\left(\hat{u}_{i_q}\right)$ on a constant and the functions of x. The test then compares the statistic to the critical value from the $\chi^2_{(J-1)}$ distribution. As heteroskedasticity was found, the model is re-estimated using the pairs bootstrap standard errors, as suggested by Buchinsky (1995), to correct for this. Therefore, quantile regressions are applied to equation (5) at quantiles 0.1, 0.25, 0.5, 0.75 and 0.9, using the bootstrapped standard errors with 3000 replications.

5.1.6. Weighted and Penalized Weighted Quantile Regressions

Similarly to the OLS regression, quantile regressions give equal weight to each observation. As discussed before, this is not optimal, and thus a weighted quantile regression, proposed by Huang et al. (2015), is employed. By adding market share weights into the quantile regression, the data on the sample can be used more effectively and therefore, the resulting estimates are more robust, reliable and efficient (Xiong and Tian, 2019). Since I am applying the adaptive Lasso methodology before every regression (as discussed in section 5.1.3.), and the variable selection is thus performed using a penalization approach, the weighted quantile regressions employed are in fact penalized weighted quantile regressions (Xiong and Tian, 2019). If we define λ_q to be a tuning parameter for quantile q, and $\lambda_{q,i}$ to be the adaptive weights vector, then the penalized weighted quantile regression estimate can be defined as:

$$\hat{\beta}_{q_{pw}} = \arg\min_{\beta_q \in \mathbb{R}} \sum_{j=1}^{n} w_j(x_j, q)\rho_q(y_j - x_j'\beta_q) + \lambda_q \sum_{i=1}^{I} \lambda_{q,i}|\beta_i| \tag{8}$$

The weights $w_j(x_j, q)$ are the market share of each car model in the UK car market. Similarly to the ordinary quantile regressions, the penalized weighted quantile regressions are estimated for quantiles 0.1, 0.25, 0.5, 0.75 and 0.9.

5.2. Hedonic Price Indices of the UK Car Market

The final research question to be covered aims at creating quality constant hedonic price indices for the UK car market, including the hedonic Laspeyres, Paasche, and Fisher price indices. The application of hedonic pricing regressions to calculate hedonic price indices came into interest again with the report by Boskin et al. (1996). The authors pointed out that the correct evaluation of the current economic situation through accurate price indices is critical for many policy decisions. Feldstein (2017) argues that inaccurate price indices cause overestimation of the inflation rate, and consequently an underestimation of living standards and aggregate output. As hedonic regressions relate variation in prices across goods and over time to differences in goods' characteristics, they can be used to construct price indices that control for changes in these characteristics (i.e. changes in quality), as well as for car models entering and exiting the market. Typically, one of two ways is used when constructing hedonic price indices using hedonic regressions – the dummy variable method, and the imputation method.

5.2.1. The Dummy Variable Method

The better known and more traditional quality-constant hedonic price index is the dummy variable (DV) price index. The method uses a classic hedonic regression approach, regressing vehicle prices on all vehicle characteristics and time dummy variables for each year. The price indices are then calculated using the coefficients of the time dummy variables for each period (Aizcorbe, 2014). However, the method suffers from several issues and probably the biggest one is that the hedonic coefficients stay the same over time (Pakes, 2003). As the characteristics of products, or the distribution of consumer preferences can change over time, so would the characteristics' coefficients. This change of coefficients over time is not accounted for by this method, since it pools over time periods.

5.2.2. The Imputation Method

The method that overcomes this problem is the imputation method, which runs separate regressions for each time period, in order to be less restrictive and allow for a change in coefficients over time (Aizcorbe, 2014). Since the aim of quality-constant hedonic price indices is to solve the issue of missing car models (i.e. models entering and exiting the market during the period analyzed), in the imputation method, a price index formula is chosen (e.g. Laspeyres price index) and a weighted hedonic price regression is then used to estimate (impute) the predicted values for the missing prices in each year. There are 3 different ways of constructing hedonic price indices using the imputation method – the single imputation method, the double imputation method, and the full imputation method.

Each of the 3 imputation methods replaces the missing (unobserved) prices of the missing car models with predicted prices from a weighted hedonic pricing regression, while some imputation methods also replace observed prices. The single imputation method is the simplest case of imputation. The method uses all of the prices that are directly observed in the dataset, and use the hedonic pricing regressions to impute all the missing prices. This method however suffers from the issue of introducing variance to the price index, as the price relative includes a mix of actual and predicted prices (Pakes, 2003; Triplett, 2006). In the case where the residuals contain omitted variables, the observed prices will contain the influence of these omitted variables, while the imputed prices will not. The double imputation method predicts (imputes) both the observed and unobserved prices for the disappearing models, and only includes actual prices for models that stayed in the market in both periods. This is considered a better approach than the single imputation method (Aizcorbe, 2014). The full imputation method replaces all observed and unobserved prices of all vehicle models with predicted prices from the hedonic regressions. This method is not usually recommended, as it replaces all actual prices with predicted prices from the hedonic regressions, and thus removes a significant amount of information (Aizcorbe, 2014). Therefore, considering everything, the double imputation method was chosen as the preferred approach for constructing the hedonic price indices. Taking Laspeyres as an example:

$$Hedonic\ Laspeyres\ Price\ Index = \sum_{s=1}^{S}\sum_{d=1}^{D}\left(\frac{\left(P_{s,t_m} * Q_{s,t_0}\right) + \left(\hat{P}_{d,t_m}\left(z_{d,t_0}\right) * Q_{d,t_0}\right)}{\left(P_{s,t_0} * Q_{s,t_0}\right) + \left(\hat{P}_{d,t_0}\left(z_{d,t_0}\right) * Q_{d,t_0}\right)}\right) * 100 \qquad (9)$$

where S is the no. of car models s that stayed between period t_0 and t_m, D is the no. of car models d that disappeared between period t_0 and t_m, P are car prices, Q are the sales, $\hat{P}$ are the predicted prices, z are the car characteristics.

The hedonic price indices constructed include the quality-constant hedonic Laspeyres, Paasche and Fisher indices (as it adjusts for the bias found when using Laspeyres or Paasche indices (Feenstra, 1988)), and standard price indices for comparison. Since the optimal functional form of the weighted hedonic regressions is, once again, semi-logarithmic, each predicted price coefficient must be converted back from its logged measure. This conversion involves a bias adjustment as "even if the coefficients are unbiased estimates of the logged indices, taking the exponent of the coefficients does not give unbiased estimates of the indices" (Aizcorbe, 2014). Thus, I apply and adjustment of Wooldridge (2009) that doesn't rely on normality of residuals: $\frac{1}{n}\sum \exp\left(\epsilon_{i,t_m}\right)$. The numerical importance of the bias adjustment is not agreed on by researchers. Triplett (2006) argued that it is likely small, while Pakes (2003) discovered that in his dataset, the bias adjustment was very large, at about 20%. Applying the Wooldridge (2009) method, the bias adjustment in this paper was found to be moderate, at around 4%.

6. Results

6.1. Effect of Different Vehicle Attributes on Vehicle Price – Results

6.1.1. OLS and WLS

Results of the OLS regression are used as a comparison for the main regression of interest – the WLS. Since the dependent variable (price) of both regressions is in the logarithmic form, each coefficient estimated represents the percentage change in the dependent variable after a 1-unit change in the independent variable. To get an accurate percentage change, I use the standard transformation for semi-logarithmic models, $\bar{\beta}_i = e^{\hat{\beta}_i} - 1$. As expected, applying market share weights (WLS) leads to more intuitive, better results, and thus those will be primarily examined[12].

The arguably most interesting category of vehicle attributes is the vehicle physical characteristics. However, while many of the coefficients of the results in figure 4 are for dummy variables, the coefficients on physical

[12] The OLS and WLS regressions also included time dummies, car segments and country of origin dummies as control variables.

characteristics aren't. The coefficient of a dummy variable can by easily interpreted and compared to other coefficients, as they are all in the same units (e.g. presence of a display increases car price by 3.66%). Physical characteristics variables are however in different units, and thus cannot be directly compared – i.e. is an increase of displacement by 1 dm^3 a lot or not (and thus is the effect on price large or not)? Therefore, to overcome this problem, and allow for a comparison between the effects of various variables, I calculate the scaled effect of each variable, which measures the effect of a 10% increase from the weighted average of a characteristic variable (as shown in figure 5).

When comparing different engine types, the diesel engine is on average cheaper compared to petrol, reducing the price of a car by 3.05%. On the other hand, if consumers want to buy an AFV, the price of the car would be almost 13% higher. Looking at the scaled effects of car characteristics in figure 5, the largest effect on price comes from curb weight, where a 10% increase from weighted average increases car price by almost 5%. In general, larger, and thus heavier cars display a much larger market price tag compared to smaller vehicles. Other characteristics that significantly increase car market prices include vehicle performance (maximum speed by 3.87% and engine power by 1.71%), safety ratings (by 1.71%), range (by 1.61%) and interior noise (by 1.41%). The strong positive effect of vehicle performance is connected to the positive effect of CO_2 emissions, as more powerful cars tend to have higher

	Ln Price	OLS (robust)	WLS
Car Characteristics	Diesel Engine	0,0201 (0,0130) †	-0,0300 (0,0122) **
	Alternative Fuel Vehicle	0,1206 (0,0340) ***	0,1221 (0,0316) ***
	Displacement *(dm^3)*	-0,0039 (0,0165)	0,0076 (0,0161)
	Engine Power *(bhp)*	0,0023 (0,0002) ***	0,0016 (0,0002) ***
	Maximum Speed *(km/h)*	0,0018 (0,0006) ***	0,0021 (0,0005) ***
	Acceleration *(s)*	0,0013 (0,0040)	-0,0061 (0,0032) *
	Interior Noise *(dB)*	0,0020 (0,0015)	-0,0022 (0,0013) *
	CO_2 Emissions *(g/km)*	-0,0007 (0,0006)	0,0014 (0,0004) ***
	Maximum Range *(100 km)*	-0,0010 (0,0062) †	0,0166 (0,0055) ***
	Size *(m^3)*	0,0225 (0,0064) ***	0,0094 (0,0052) *
	Number of Doors	-0,0300 (0,0049) ***	-0,0242 (0,0038) ***
	Trunk Capacity *(10 l)*	0,0005 (0,0003) **	0,0017 (0,0003) ***
	Fuel Tank Capacity *(l)*	0,0013 (0,0012)	-0,0031 (0,0012) **
	Curb Weight *(10 kg)*	0,0034 (0,0005) ***	0,0037 (0,0003) ***
Safety	NCAP Average Rating	0,0009 (0,0005) *	0,0022 (0,0004) ***
	Airbags per Seat	0,0317 (0,0098) ***	0,0441 (0,0078) ***
	Pretensioners per Seat	-0,0483 (0,0189) **	-0,0467 (0,0160) ***
	Loadlimiters per Seat	-0,0227 (0,0176)	-0,0467 (0,0145) ***
Equipment	Traction Control	0,0285 (0,0124) **	0,0208 (0,0084) **
	Hill Start Assist	-0,0081 (0,0085)	-0,0112 (0,0071) †
	Deflation Warning System	0,0116 (0,0083)	0,0068 (0,0072)
	Forward Collision Warning	-0,0563 (0,0117) ***	-0,0117 (0,0113)
	Traffic Sign Recognition	-0,0456 (0,0198) **	-0,0160 (0,0203)
	Lane Assist	0,0178 (0,0108) *	0,0139 (0,0098)
	Display	0,0339 (0,0079) ***	0,0359 (0,0064) ***
	Automatic Air Conditioning	0,0270 (0,0117) **	0,0732 (0,0105) ***
	Electric Mirrors	0,0103 (0,0148)	0,0612 (0,0105) ***
	Heated Mirrors	0,0097 (0,0132)	-0,0026 (0,0081)
	Heated Steering Wheel	-0,0012 (0,0140)	0,0138 (0,0143)
	Automatic Wipers	0,0329 (0,0110) ***	0,0542 (0,0109) ***
	Automatic Lights	-0,0222 (0,0096) **	-0,0157 (0,0093) *
	Radio	0,0985 (0,0373) ***	0,0357 (0,0298)
	Number of Speakers	0,0108 (0,0023) ***	0,0047 (0,0017) ***
	Remote Audio Control	0,0179 (0,0096) *	-0,0229 (0,0067) ***
	USB Jack	0,0184 (0,0079) **	-0,0106 (0,0067) †
	Parking Sensors	0,0363 (0,0113) ***	0,0179 (0,0090) **
	Rear View Camera	0,0351 (0,0100) ***	0,0382 (0,0086) ***
	Heated Seats	-0,0463 (0,0103) ***	-0,0248 (0,0100) **

Fig. 4 *Selected results of the robust OLS and WLS regressions for finding the effect of vehicle attributes on price. Notes: ***significant at 1%, **significant at 5%, *significant at 10%, †significant at 15%.*

Variable	Scaled Effect for +10% from Weighted Average	Absolute Scaled Effect for +10% from Weighted Average
Curb Weight (kg)	0,048	0,048
Maximum Speed (km/h)	0,038	0,038
Combined CO_2 (g/km)	0,018	0,018
NCAP Average Rating	0,017	0,017
Engine Power (bhp)	0,017	0,017
Range (km)	0,016	0,016
Fuel Tank Capacity (liters)	-0,016	0,016
Interior Noise (dB)	-0,014	0,014
Size L*W*H (m^3)	0,011	0,011
Number of Doors	-0,010	0,010
Acceleration (0 to 100 km/h in seconds)	-0,008	0,008
Trunk Capacity (liters)	0,007	0,007
Airbags per Seat	0,006	0,006
Number of Speakers	0,003	0,003
Loadlimiters per Seat	-0,002	0,002
Pretensioners per Seat	-0,002	0,002
Displacement (cm^3)	0,001	0,001

Fig. 5 *The scaled effect of car characteristics on vehicle price. Note: red variables are insignificant in WLS.*

emissions, and therefore display a higher price. Much smaller effects on price of below 1% can be seen for acceleration, trunk capacity, airbags per seat and number of speakers. Lastly, we can see an unexpected negative effect for loadlimiters and pretensioners, however, this non intuitive effect is very small, where a 10% increase from weighted average would reduce the car price by a negligible 0.2%.

As expected, presence of extra equipment as standard increases the overall price of a vehicle. The equipment that is found to increase car prices the most is automatic air conditioning (by 7.59%), fully electric mirrors (by 6.31%), and automatic wipers (by 5.57%). Other significant equipment includes the rear-view camera (3.89%), display (3.66%), traction control (2.1%), and, marginally, the number of speakers (0.47% for an extra speaker). Several equipment features display a non-intuitive negative effect; however, the magnitude and significance are very low. Furthermore, advanced equipment, such as traffic sign recognition, is found to be insignificant, likely because this equipment is rarely added as standard, leading to a high variation in the effect on price and thus inconclusive results.

6.1.2. Quantile and Weighted Quantile Regressions

The bootstrapped quantile and penalized weighted quantile regression were run to examine how the effect of vehicle characteristics changes at different points of the car price distribution. The weighted quantile regression (figure 6[13]) in general shows better results (as expected from the literature review), and thus will be the one examined.

Looking at the results for the physical characteristics, there are a few interesting findings in the quantiles examined. We can see that higher performance (engine displacement and power) increases the market price significantly more so for more expensive vehicles. Higher displacement values do not have any significant effect on cheaper cars, but at the 75th and 90th percentile, an increase in engine displacement by 1 dm^3 increases the car price by 6.26% and 11.19% respectively. The same is true for the increasingly positive effect of engine power on car price. The most likely reason for this larger effect of performance for more expensive cars, is that richer consumers put greater importance to vehicle performance, while buyers of cheap vehicles rather look at other factors, such as e.g. utility, or low running costs. Similarly, the effect of vehicle maximum range gets stronger as we go up the quantiles. At the 10th quantile, an increase in range of 100 km doesn't affect the price, while at the 25th quantile the effect is 1.54%, and at 75th it is almost 4%. This means that richer consumers are willing to pay an increasingly higher price for greater range, but only up to a point, as at the 90th quantile, the price increase drops to only 1.75%. Rich buyers thus value this characteristic significantly less, likely valuing other characteristics, such as performance, more. Lastly,

[13] The bootstrapped quantile and penalized weighted quantile regressions also included time dummies, car segments and country of origin dummies as control variables.

an interesting quantile effects can be seen for trunk capacity. As the trunk capacity characteristic represents a utility attribute of a vehicle, and it would be expected to have the highest effect on car prices for the cheapest vehicles, which is exactly what is found. The increase in car price for a 10 liter increase in trunk capacity is the largest at the 10th and 25th percentiles (0.21% and 0.19%), while for more expensive cars this effect is reduced and stable, at around 0.14%.

	Ln Price	WLS	Weighted (Penalized) Quantile Regression				
			0.10	0.25	0.50	0.75	0.90
Car Characteristics	Diesel Engine	-0,0300 (0,0122) **	-0,0424 (0,0180) **	-0,0181 (0,0136)	-0,0058 (0,0133)	-0,0672 (0,0147) ***	-0,0503 (0,0194) ***
	Alternative Fuel Vehicle	0,1221 (0,0316) ***	0,0764 (0,0466) *	0,0496 (0,0353)	0,1202 (0,0345) ***	0,0891 (0,0382) **	0,1107 (0,0502) **
	Displacement *(dm^3)*	0,0076 (0,0161)	-0,0133 (0,0237)	0,0135 (0,0180)	-0,0125 (0,0176)	0,0607 (0,0194) ***	0,1061 (0,0255) ***
	Engine Power *(bhp)*	0,0016 (0,0002) ***	0,0012 (0,0004) ***	0,0020 (0,0003) ***	0,0022 (0,0003) ***	0,0024 (0,0003) ***	0,0023 (0,0004) ***
	Maximum Speed *(km/h)*	0,0021 (0,0005) ***	0,0010 (0,0007)	0,0003 (0,0006)	0,0013 (0,0006) **	0,0008 (0,0006)	0,0011 (0,0008)
	Acceleration *(s)*	-0,0061 (0,0032) *	-0,0189 (0,0047) ***	-0,0149 (0,0036) ***	-0,0092 (0,0035) ***	-0,0065 (0,0039) *	-0,0064 (0,0051)
	Interior Noise *(dB)*	-0,0022 (0,0013) *	0,00004 (0,0019)	-0,0002 (0,0015)	-0,0023 (0,0014) †	-0,0018 (0,0016)	-0,0015 (0,0021)
	CO_2 Emissions *(g/km)*	0,0014 (0,0004) ***	0,0008 (0,0006)	0,0013 (0,0005) ***	0,0019 (0,0005) ***	0,0031 (0,0005) ***	0,0009 (0,0007)
	Maximum Range *(100 km)*	0,0166 (0,0055) ***	0,0102 (0,0081)	0,0153 (0,0062) **	0,0217 (0,0060) ***	0,0384 (0,0066) ***	0,0173 (0,0087) **
	Size *(m^3)*	0,0094 (0,0052) *	0,0166 (0,0077) **	0,0147 (0,0058) **	0,0009 (0,0057)	0,0167 (0,0063) ***	0,0208 (0,0083) **
	Number of Doors	-0,0242 (0,0038) ***	-0,0059 (0,0056)	-0,0025 (0,0043)	-0,0113 (0,0042) ***	-0,0170 (0,0046) ***	-0,0271 (0,0061) ***
	Trunk Capacity *(10 l)*	0,0017 (0,0003) ***	0,0022 (0,0004) ***	0,0019 (0,0003) ***	0,0013 (0,0003) ***	0,0015 (0,0003) ***	0,0014 (0,0004) ***
	Fuel Tank Capacity *(l)*	-0,0031 (0,0012) **	-0,0028 (0,0018) †	-0,0046 (0,0014) ***	-0,0060 (0,0013) ***	-0,0080 (0,0015) ***	-0,0041 (0,0019) **
	Curb Weight *(10 kg)*	0,0037 (0,0003) ***	0,0031 (0,0005) ***	0,0026 (0,0004) ***	0,0041 (0,0004) ***	0,0041 (0,0004) ***	0,0031 (0,0005) ***
Safety	NCAP Average Rating	0,0022 (0,0004) ***	0,0030 (0,0005) ***	0,0028 (0,0004) ***	0,0025 (0,0004) ***	0,0020 (0,0004) ***	-0,00004 (0,0006)
	Airbags per Seat	0,0441 (0,0078) ***	0,0271 (0,0116) **	0,0326 (0,0088) ***	0,0376 (0,0086) ***	0,0586 (0,0095) ***	0,0348 (0,0124) ***
	Pretensioners per Seat	-0,0467 (0,0160) ***	0,0157 (0,0236)	-0,0686 (0,0179) ***	-0,0285 (0,0175) †	-0,0356 (0,0194) *	-0,0534 (0,0254) **
	Loadlimiters per Seat	-0,0467 (0,0145) ***	-0,0692 (0,0213) ***	-0,0460 (0,0162) ***	-0,0613 (0,0158) ***	-0,0415 (0,0175) **	-0,0265 (0,0230)
Equipment	Traction Control	0,0208 (0,0084) **	-0,0039 (0,0124)	0,0022 (0,0094)	0,0092 (0,0092)	0,0143 (0,0101)	-0,0054 (0,0133)
	Hill Start Assist	-0,0112 (0,0071) †	-0,0031 (0,0105)	-0,0014 (0,0080)	-0,0079 (0,0078)	-0,0167 (0,0086) *	0,0116 (0,0113)
	Deflation Warning System	0,0068 (0,0072)	0,0376 (0,0106) ***	0,0268 (0,0080) ***	0,0024 (0,0078)	-0,0071 (0,0087)	0,0045 (0,0114)
	Forward Collision Warning	-0,0117 (0,0113)	-0,0398 (0,0167) **	-0,0474 (0,0127) ***	-0,0046 (0,0124)	0,0174 (0,0137)	-0,0015 (0,0180)
	Traffic Sign Recognition	-0,0160 (0,0203)	-0,0050 (0,0299)	0,0369 (0,0227) †	-0,0078 (0,0222)	-0,0315 (0,0245)	-0,0044 (0,0322)
	Lane Assist	0,0139 (0,0098)	0,0202 (0,0144)	0,0154 (0,0109)	0,0216 (0,0107) **	0,0142 (0,0118)	0,0165 (0,0155)
	Display	0,0359 (0,0064) ***	0,0426 (0,0095) ***	0,0363 (0,0072) ***	0,0264 (0,0070) ***	0,0091 (0,0078)	0,0029 (0,0102)
	Automatic Air Conditioning	0,0732 (0,0105) ***	0,1088 (0,0154) ***	0,1097 (0,0117) ***	0,0746 (0,0114) ***	0,0899 (0,0126) ***	0,0780 (0,0166) ***
	Electric Mirrors	0,0612 (0,0105) ***	0,0767 (0,0149) ***	0,0684 (0,0113) ***	0,0972 (0,0111) ***	0,0992 (0,0122) ***	0,0763 (0,0161) ***
	Heated Mirrors	-0,0026 (0,0081)	0,0278 (0,0120) **	0,0005 (0,0091)	-0,0299 (0,0089) ***	-0,0094 (0,0098)	-0,0158 (0,0129)
	Heated Steering Wheel	0,0138 (0,0143)	0,0259 (0,0210)	-0,0006 (0,0159)	0,0022 (0,0155)	-0,0246 (0,0172)	-0,0128 (0,0226)
	Automatic Wipers	0,0542 (0,0109) ***	0,0548 (0,0160) ***	0,0434 (0,0122) ***	0,0383 (0,0119) ***	0,0291 (0,0131) **	0,0267 (0,0173) †
	Automatic Lights	-0,0157 (0,0093) *	-0,0364 (0,0137) ***	-0,0182 (0,0104) *	-0,0037 (0,0101)	0,0073 (0,0112)	-0,0087 (0,0147)
	Radio	0,0357 (0,0298)	0,0032 (0,0438)	0,0120 (0,0333)	-0,0228 (0,0325)	0,0353 (0,0359)	0,1388 (0,0472) ***
	Number of Speakers	0,0047 (0,0017) ***	0,0044 (0,0025) *	0,0052 (0,0019) ***	0,0068 (0,0019) ***	0,0021 (0,0021)	0,0038 (0,0027)
	Remote Audio Control	-0,0229 (0,0067) ***	-0,0277 (0,0099) ***	-0,0211 (0,0075) ***	-0,0283 (0,0074) ***	-0,0198 (0,0081) **	-0,0028 (0,0107)
	USB Jack	-0,0106 (0,0067) †	-0,0014 (0,0098)	0,0052 (0,0074)	-0,0073 (0,0073)	-0,0142 (0,0080) *	-0,0065 (0,0106)
	Parking Sensors	0,0179 (0,0090) **	0,0284 (0,0132) **	0,0179 (0,0100) *	0,0152 (0,0098) †	0,0366 (0,0108) ***	0,0247 (0,0142) *
	Rear View Camera	0,0382 (0,0086) ***	0,0390 (0,0127) ***	0,0296 (0,0096) ***	0,0128 (0,0094)	0,0442 (0,0104) ***	0,0239 (0,0137) *
	Heated Seats	-0,0248 (0,0100) **	0,0187 (0,0147)	-0,0181 (0,0111) †	-0,0269 (0,0109) **	-0,0649 (0,0120) ***	-0,0443 (0,0158) ***

Fig. 6 *Selected results of weighted penalized quantile regression, for finding the effect of vehicle attributes at different points of the car price distribution. Notes: ***significant at 1%, **significant at 5%, *significant at 10%, †significant at 15%.*

Safety characteristics display two opposing effects – the effect of safety ratings decreases, while the effect of airbags per seat increases. At the 10th quantile, an increase of safety ratings by 10 points and an extra airbag increases car prices by 3.05% and 2.75% respectively, while at 75th quantile, this is 2.02% and 6.04%. This could mean buyers of cheap vehicles are willing to pay more for higher safety ratings compared to features such as airbags, while buyers of more expensive vehicles put less importance on ratings and rather expect greater amount of tangible safety features.

Examining the equipment, the size of the coefficients and their significance gradually decreases as we move up the quantiles. While at the 10th percentile, there are 11 equipment features of at least a 5% significance, at the median there are 9, and at the 90th percentile only 4. This means that equipment in general has the largest effect on vehicle price in the cheapest cars, and this effect decreases as cars become more expensive. An explanation for this can be the fact that at very high price levels, the price of additional equipment is negligible in relation to the overall price of the vehicle. Additionally, a lot of rich buyers likely put much greater value on intangible characteristics, such as vehicle style, or the social status the car represents. An example of this are the automatic wipers. At the 10th percentile, automatic wipers increase price by 5.63% on average, and the result is significant at a 1% level. However, this reduces as we go up the quantiles, and at the 90th this is only 2.71%, and significant only at a 15% level.

6.2. Effect of Different Vehicle Attributes on Car Price (CVs vs. AFVs) – Results

The second research question aimed to examine how the influence of different vehicle attributes on car prices differs between CVs and AFVs. As in section 6.1.1., the main regression of interest is the WLS, and this can be seen in figure 7[14], along with the heteroskedasticity-robust OLS regression for comparison.

Looking at the vehicle physical characteristics, improvements in AFV performance (engine power, maximum speed, acceleration) influence the car market price much more strongly than improvements to CVs. An increase of 10 bhp and 10km/h in a CV increases the car price by 2.02% and 1.92% respectively, while in an AFV, this increase is much higher at 5.65% and 4.81%. For acceleration, this effect is even more pronounced. A decrease of acceleration time to 100 km/h by 1 second doesn't have a significant effect on the CV price, but increases AFV price by 6.19%. This suggests that improving performance attributes of AFVs is more complicated and expensive for the manufacturers compared to the CVs, which results in AFVs being significantly more expensive than CVs for comparable physical characteristics. Furthermore, as AFVs represent cleaner, environmentally friendly transportation, reduction in CO_2 emissions by 10 g/km is connected to a further premium of 1.61% paid on AFVs. The most surprising effect involves maximum range, where greater range reduces the price. Although unexpected, this result can be explained by the inclusion of hybrid vehicles into the

Ln Price		OLS (robust)		WLS	
		CVs	AFVs	CVs	AFVs
Car Characteristics	Diesel Engine	0,0021 (0,0146)		-0,0375 (0,0127) ***	
	Displacement *(dm³)*	0,0130 (0,0179)		0,0117 (0,0166)	
	Engine Power *(bhp)*	0,0023 (0,0002) ***	0,0062 (0,0009) ***	0,0020 (0,0003) ***	0,0055 (0,0012) ***
	Maximum Speed *(km/h)*	0,0020 (0,0006) ***	0,0027 (0,0016) *	0,0019 (0,0005) ***	0,0047 (0,0019) ***
	Acceleration *(s)*	0,0045 (0,0040)	-0,0540 (0,0201) ***	-0,0033 (0,0033)	-0,0601 (0,0166) ***
	Interior Noise *(dB)*	-0,0010 (0,0016)	0,0134 (0,0073) *	-0,0025 (0,0014) *	0,0166 (0,0068) **
	CO_2 Emissions *(g/km)*	-0,0009 (0,0007)	-0,0015 (0,0007) **	0,0012 (0,0005) ***	-0,0016 (0,0007) **
	Maximum Range *(100 km)*	-0,0072 (0,0075)	-0,0161 (0,0069) **	0,0153 (0,0058) ***	-0,0186 (0,0085) **
	Size *(m³)*	0,0190 (0,0065) ***	-0,0110 (0,0214)	0,0073 (0,0053)	0,0167 (0,0196)
	Number of Doors	-0,0321 (0,0048) ***	0,1178 (0,0287) ***	-0,0234 (0,0039) ***	0,0747 (0,0311) **
	Trunk Capacity *(10 l)*	0,0004 (0,0002) *	-0,0068 (0,0023) ***	0,0016 (0,0003) ***	-0,0081 (0,0034) **
	Fuel Tank Capacity *(l)*	0,0014 (0,0013)		-0,0029 (0,0013) **	
	Curb Weight *(10 kg)*	0,0034 (0,0005) ***		0,0037 (0,0003) ***	
Safety	NCAP Average Rating	0,0009 (0,0005) **		0,0021 (0,0004) ***	
	Airbags per Seat	0,0342 (0,0099) ***	-0,1140 (0,0515) **	0,0456 (0,0080) ***	-0,1848 (0,1079) *
	Pretensioners per Seat	-0,0345 (0,0184) *		-0,0443 (0,0163) ***	
	Loadlimiters per Seat	-0,0336 (0,0184) *		-0,0478 (0,0148) ***	
Equipment	Traction Control	0,0234 (0,0123) *		0,0231 (0,0085) ***	
	Hill Start Assist	-0,0020 (0,0084)	-0,2692 (0,0502) ***	-0,0176 (0,0073) **	-0,4119 (0,0543) ***
	Deflation Warning System	0,0098 (0,0086)	0,1122 (0,0589) *	0,0125 (0,0073) *	0,0955 (0,0407) **
	Forward Collision Warning	-0,0818 (0,0124) ***		-0,0131 (0,0118)	
	Traffic Sign Recognition	-0,0291 (0,0197) †		-0,0139 (0,0213)	
	Lane Assist	0,0285 (0,0109) ***	0,0952 (0,0500) *	0,0171 (0,0101) *	0,0243 (0,0864)
	Display	0,0280 (0,0081) ***	0,1000 (0,0353) ***	0,0332 (0,0107) ***	0,1009 (0,0359) ***
	Automatic Air Conditioning	0,0160 (0,0115)	-0,0219 (0,0761)	0,0679 (0,0107) ***	0,0184 (0,0600)
	Electric Mirrors	0,0152 (0,0137)		0,0585 (0,0103) ***	
	Heated Mirrors	0,0078 (0,0131)		-0,0031 (0,0083)	
	Heated Steering Wheel	0,0022 (0,0142)	0,0463 (0,0706)	0,0150 (0,0147)	-0,0921 (0,0492) *
	Automatic Wipers	0,0419 (0,0118) ***		0,0443 (0,0112) ***	
	Automatic Lights	-0,0229 (0,0103) **	0,0871 (0,0508) *	-0,0129 (0,0097)	0,1873 (0,0517) ***
	Radio	0,0778 (0,0374) **		0,0448 (0,0300) †	
	Number of Speakers	0,0125 (0,0024) ***	0,0550 (0,0134) ***	0,0046 (0,0018) ***	0,0435 (0,0110) ***
	Remote Audio Control		-0,2708 (0,1067) **		-0,4062 (0,1488) ***
	USB Jack	0,0248 (0,0080) ***		-0,0092 (0,0068)	
	Parking Sensors	0,0321 (0,0118) ***		0,0196 (0,0092) **	
	Rear View Camera	0,0296 (0,0104) ***	0,1958 (0,0578) ***	0,0325 (0,0089) ***	0,3767 (0,0554) ***
	Heated Seats	-0,0397 (0,0106) ***		-0,0194 (0,0102) *	

Fig. 7 *Results of the robust OLS and WLS regressions for comparing the effect of vehicle attributes on the price of CVs and AFVs. Notes: ***significant at 1%, **significant at 5%, *significant at 10%, †significant at 15%.*

[14] The OLS and WLS regressions also included time dummies, car segments and country of origin dummies as control variables. Several variables had to be removed for the AFV dataset, due to AFV irrelevance (e.g. displacement) and multicollinearity issues.

examined AFV sub dataset. Hybrids have greater range than electric or hydrogen cars, and a lower price. This would cause a negative relationship of price and range, where cheaper hybrids have greater range compared to EVs.

The situation with equipment is similar to the physical characteristics – more equipment generally increases the car price much more, if the car is an AFV. This effect is the strongest for the rear-view camera – its presence as standard increases the CV price by 3.3%, but on an AFV, this increase skyrockets to 45.75%. Similar difference can be seen for automatic lights (no effect on CV price vs. 20.6% increase for an AFV), the infotainment display (3.38% vs. 10.62%), the deflation warning system (1.26% vs. 10.02%) and the number of speakers (0.46% vs. a 4.45%). This phenomenon once again portrays the situation where AFV manufacturers claim a higher premium for extra characteristics and features, compared to the same ones on a CV, resulting in a higher final price.

6.3. Hedonic Price Indices of the UK Car Market – Results

The third research question aims at constructing quality constant hedonic price indices for the UK car market. For this reason, the hedonic Laspeyres, Paasche and Fisher price indices have been constructed, along with, several standard price indices for comparison. The results of all the price indices (2008 = 100) can be seen below in figure 8.

The most basic price indices calculated are the average and weighted average price indices. These simply follow the development of the average car price, with the difference that the weighted index weights each vehicle by its market share, giving a better approximation of the situation in the market, and suggesting a price increase of 42.34% between 2008 and 2018. However, it suffers from the flaw of being as much a price index as a quantity index.

Year	Average Price Index	Weighted Average Price Index	Laspeyres Price Index	Paasche Price Index	Fisher Price Index	Hedonic Laspeyres Price Index	Hedonic Paasche Price Index	Hedonic Fisher Price Index
2008	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
2009	100,96	98,27	100,21	100,15	100,18	100,05	107,14	103,53
2010	107,55	102,42	103,71	102,57	103,14	94,57	101,17	97,82
2011	114,02	106,89	104,52	103,71	104,11	95,67	98,82	97,23
2012	115,23	109,04	108,01	105,79	106,89	98,76	99,85	99,30
2013	115,76	108,66	108,49	105,68	107,08	98,45	98,20	98,32
2014	120,60	112,50	116,67	111,91	114,26	106,51	99,08	102,73
2015	129,70	117,16	120,14	111,44	115,71	102,48	100,80	101,64
2016	137,20	125,13	120,75	115,81	118,26	104,85	102,10	103,47
2017	154,61	135,04	122,68	125,13	123,90	112,44	107,91	110,15
2018	161,21	142,34	135,02	133,16	134,09	111,45	110,39	110,92

Fig. 8 *The price indices calculated for the UK car market between 2008 and 2018, where the year 2008 is the base (100).*

The Laspeyres and Paasche price indices solve this issue. As seen in figure 8, both indices show similar results, where the difference comes from the fact that Laspeyres slightly overstates the inflation, and the Paasche slightly understates it. Therefore, the Fisher price index is calculated as the optimal, geometric mean between the Laspeyres and Paasche. The Fisher index indicates that the price level increase in the UK car market in 2008-2018 was 34.09%, or 2.98% annually. However, these indices suffer from an important problem of not accounting for quality change – i.e. for disappearing and newly entering car models. Therefore, the quality constant hedonic Laspeyres, Paasche and Fisher price indices were constructed, in order to get accurate price index estimations.

Hedonic price indices apply the discussed double imputation method in order estimate price changes for all vehicle models. Thus, hedonic price indices are quality constant, and show the closest estimation to the real evolution of the pure prices in the UK car market – how much of the price increase was due to quality change, and how much due to other factors. As above, hedonic Fisher price index is the geometric mean between Laspeyres and Paasche, and thus represents the optimal hedonic price index calculated. According to the hedonic Fisher price index, the car price level increased by 10.92% in 2008-2018, or 1.04% annually. Such result is relatively comparable, albeit lower, to the UK car market (1977-1991) results of Murray and Sarantis (1999), at 1.7% annually. The 1.04% growth is in stark contrast to the weighted price index results, which exhibit annual growth of 3.59%. This suggests that about 70% of the price increase in UK passenger cars in 2008-2018 was caused by improvements in car quality – i.e. quality improvements were responsible for a price increase of 2.55% per year. As there have been extensive improvements in overall vehicle quality in terms of physical characteristics, safety, and the amount of equipment, the discovery that quality improvements drive the car price growth is in fact not surprising.

6.4. The Key Vehicle Characteristics Influencing the Sales of AFVs – Results

The last research question aims to find the key characteristics and features that influence sales of AFVs (i.e. which attributes are the most important for AFV consumers). This was explored in the literature before, but only using stated preference data, typically concluding that performance, range, environmental friendliness and price are the most important (e.g. Ewing and Sarigöllü (2000), Larson et al. (2014)). Thus, the goal of this research question is to check whether these conclusions hold when applied to revealed preference data and whether the findings of stated preference AFV studies can be relied upon. Therefore, a robust OLS regression is applied on a sub dataset of AFVs, with $\ln Sales$ as the dependent variable, and car characteristics and features as independent variables.

The results of the robust OLS regression in figure 9 show several interesting findings in terms of the characteristics that affect AFV sales the most in the UK car market (2014-2018). For example, presence of most of the equipment as standard, such as automatic air conditioning, rear view camera or number of speakers, understandably significantly increases vehicle sales. However, most characteristics suffer from a similar deficiency as the results in section 6.1.1. – the effects cannot be directly compared in terms of magnitude. We see that 1 bhp increase in engine power increases sales by 8.44%, but is 1 bhp change large or not? Is it larger, smaller or comparable to a 100 dm^3 increase in vehicle size? Thus, to remove this problem of comparison, the scaled effects were calculated as in section 6.1.1.. These scaled effects for a 1% increase from each characteristic's weighted average can be seen in figure 10, allowing for the identification of characteristics that have the largest effect on AFV sales.

Looking at figure 10, we can see that the characteristics with the largest effect on sales are performance (in the form if engine power), comfort (vehicle size), environmental friendliness (CO_2 emissions), and maximum range. An increase of 1% from the weighted average of performance increases AFV sales by 14.11%. This finding is supported by the results for interior noise, likely caused by the fact that higher performance vehicles are generally noisier. Furthermore in terms of magnitude, an improvement of 1% from the weighted average of size/comfort increases AFV sales by 9.86%, a reduction of CO_2 emissions (i.e. cleaner AFVs) cause an increase of 7.36%, and an increase of maximum range increases AFV sales by 6.72%. On the other hand, vehicle price in this case is found with a smaller effect – a decrease in sales of 2.63% for a 1% increase in price from the weighted average. In a nutshell, the most important AFV characteristics are performance, maximum range, environmental friendliness and size/comfort. Thus, the revealed preference data of the UK car market (2014-2018) supports the findings of the stated preferences past literature, reaching similar conclusions in general. This suggests that using stated preference AFV data can provide a solid approximation to the real situation in the car market, which is especially useful when actual market data is not available.

	Ln Sales	OLS (robust)
	Price *(£1000)*	-0,0879 (0,0229) ***
Year	2015	-0,5418 (0,4132)
	2016	-0,5610 (0,4807)
	2017	0,1129 (0,3730)
	2018	0,0021 (0,3978)
Car Characteristics	Engine Power *(bhp)*	0,0810 (0,0104) ***
	Maximum Speed *(km/h)*	-0,0026 (0,0132)
	Acceleration *(s)*	0,4428 (0,1312) ***
	Interior Noise *(dB)*	0,5678 (0,0771) ***
	CO_2 Emissions *(g/km)*	-0,1110 (0,0156) ***
	Maximum Range *(km)*	0,0077 (0,0010) ***
	Size *(m^3)*	0,7586 (0,1751) ***
	Wheelbase *(mm)*	-0,0247 (0,0042) ***
	Turning Circle *(m)*	-2,9800 (0,4937) ***
	Number of Doors	-0,9574 (0,3045) ***
	Trunk Capacity *(l)*	0,0072 (0,0017) ***
Safety	NCAP Average Rating	0,0687 (0,0328) **
	Airbags per Seat	-2,9741 (0,7921) ***
Equipment	Hill Start Assist	0,1971 (0,6114)
	Blind Spot Assist	-0,9753 (0,5012) *
	Lane Assist	-2,2961 (0,8639) **
	Automatic Air Conditioning	5,1361 (0,7888) ***
	Heated Steering Wheel	-0,4396 (0,5374)
	Automatic Lights	1,2236 (0,6709) *
	Number of Speakers	0,3117 (0,1508) **
	Rear View Camera	4,3073 (0,7493) ***

Fig. 9 *Results of the robust OLS regression for finding the effect of car characteristics on sales. Notes: ***significant at 1%, **significant at 5%, *significant at 10%, †significant at 15%.*

Variable	Scaled Effect for +1% from Weighted Average	Absolute Scaled Effect for +1% from Weighted Average
Interior Noise (dB)	0,372	0,372
Engine Power (bhp)	0,132	0,132
Size L*W*H (m^3)	0,094	0,094
Range (km)	0,065	0,065
NCAP Average Rating	0,054	0,054
Automatic Air Conditioning	0,047	0,047
Acceleration (0 to 100 km/h in seconds)	0,042	0,042
Rear View Camera	0,030	0,030
Trunk Capacity (liters)	0,030	0,030
Number of Speakers	0,020	0,020
Automatic Lights	0,007	0,007
Hill Start Assist	0,001	0,001
2017	2,93E-04	2,93E-04
2018	6,03E-06	6,03E-06
Heated Steering Wheel	-0,001	0,001
2015	-0,001	0,001
2016	-0,001	0,001
Blind Spot Assist	-0,002	0,002
Maximum Speed (km/h)	-0,004	0,004
Lane Assist	-0,009	0,009
Price (£)	-0,026	0,026
Airbags per Seat	-0,045	0,045
Number of Doors	-0,046	0,046
Combined CO_2 (g/km)	-0,071	0,071
Turning Circle (m)	-0,323	0,323
Wheelbase (mm)	-0,668	0,668

Fig. 10 *The scaled effect of car characteristics on AFV sales. Note: red variables are insignificant in OLS.*

7. Conclusions

The main objective of my study was to provide a direct insight into the UK car market, focusing on the effect of various car characteristics on prices and sales. The main goals of my empirical paper were: identifying which characteristics and equipment influence prices the most, exploring how this influence varies between CVs and AFVs, constructing quality constant hedonic price indices, and identifying key vehicle attributes that influence AFV sales. To fulfill these goals, I constructed an extensive and novel dataset for the UK car market in 2008-2018. A hedonic pricing model is then created, applying the adaptive Lasso, OLS, WLS, quantile, and penalized weighted quantile regressions, along a range of robustness tests and corrections, such as variance inflation factor analysis, omitted variable and heteroskedasticity tests, and functional form tests.

The key vehicle characteristics that influence the car prices in the UK car market were identified as size/weight, performance (speed and engine power), safety ratings, and maximum vehicle range. The effect of performance, safety, and range on price is found to be stronger for the more expensive vehicles. Furthermore, the presence of automatic air conditioning, fully electric mirrors, automatic windscreen wipers, rear-view camera, and display increases prices the most, but the effects get weaker as the price of the vehicle increases. Therefore, especially manufacturers of cheaper vehicles should focus on improving these characteristics and equipment features, in order to reduce their prices and become more competitive in the UK car market. Manufacturers and car dealers of the more expensive vehicles should also note that vehicle performance is the strongest factor affecting the price of higher-end cars, and this effect gets stronger the more expensive the car is.

Examination of how these results differ between CVs and AFVs shows that improvements in vehicle performance (engine power, speed and acceleration) affect the AFV price significantly more than the CV price. Therefore, to reduce AFV prices and increase competitiveness, manufacturers should focus their R&D on making improvements to performance cheaper. Furthermore, AFV consumers pay an extra premium when they buy a cleaner AFV, while having extra equipment on a vehicle generally increases the AFV price significantly more than the CV price. This difference between equipment costs is the strongest for the rear view camera and the infotainment display.

Analysis of the hedonic Fisher price index shows that the majority of the rise in the UK car market price level can be explained by a rise in quality. The results suggest that improvements in the overall quality of vehicles are responsible for car prices growing annually by about 2.55%. Therefore, about 70% of the observed increase in car prices between 2008 and 2018 in the UK car market was caused by improvements in vehicle quality.

Lastly, it is found that AFV consumers value performance, maximum range, environmental friendliness, and size/comfort the most. Therefore, it is advisable that manufacturers further focus their R&D on improvements in performance and vehicle range, such as improvements to batteries. Since consumers also put value on greater positive effect on the environment, advertising and marketing should target and promote cleaner AFVs, in order to encourage higher general AFV sales. Furthermore, the results also support the use of stated preference AFV data when market data is not available, as it is found to provide a good approximation to the situation in the market.

The results acquired can offer useful information to various parties. For example, the characteristics that influence vehicle prices are of interest to car dealers and manufacturers, who aim to optimize their pricing strategies and offer competitive prices, while increasing their sales in the UK. Furthermore, quality constant hedonic price indices are at the focus of many researchers, as noted by Reis and Santos Silva (2006), and Feldstein (2017). These are for example useful for the governments, allowing for correct measurements of quality change and thus accurate predictions and analysis of productivity and living standards. Additionally, confirming which vehicle characteristics are the most important to the UK AFV buyers is of great use to the UK government, manufacturers and marketers, who look to encourage the adoption of environmentally friendly and more energy efficient vehicles. Achieving higher AFV adoption rates is high on the government policy goals, especially in relation to the promised progress towards the reduction of polluting emissions and global goals of reducing climate change.

References

Aizcorbe, A. (2014). *A practical guide to price index and hedonic techniques*. 1st ed. New York: Oxford University Press.

Alberini, A., Bareit, M. and Filippini, M. (2014). Does the Swiss Car Market Reward Fuel Efficient Cars? Evidence from Hedonic Pricing Regressions, a Regression Discontinuity Design, and Matching. *SSRN Electronic Journal*.

Anderson, S., Parry, I., Sallee, J. and Fischer, C. (2011). Automobile Fuel Economy Standards: Impacts, Efficiency, and Alternatives. *Review of Environmental Economics and Policy*, 5(1), pp.89-108.

Andersson, H. (2005). The Value of Safety as Revealed in the Swedish Car Market: An Application of the Hedonic Pricing Approach. *Journal of Risk and Uncertainty*, 30(3), pp.211-239.

Auto123. (n.d.). *Vehicle Characteristics and Equipment*. Available: https://www.auto123.com/en/ [Accessed 27.11.2019].

Auto Express. (n.d.). *New Car Prices and Specs Guide*. Available at: https://www.autoexpress.co.uk/car-news/98431/new-car-prices-and-specs-guide-2018 [27.11.2019].

Blomquist, G., Levy, D. and Miller, T. (1996). Values of Risk Reduction Implied by Motorist Use of Protection Equipment: New Evidence from Different Populations. *Journal of Transport Economics and Policy*, pp.55-66.

Boskin, M.J., Dulberger, E.R., Gordon, R.J., Griliches, Z., Jorgenson, D., (1996). Toward a more accurate measure of the cost of living. Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index.

Buchinsky, M. (1995). Estimating the asymptotic covariance matrix for quantile regression models a Monte Carlo study. *Journal of Econometrics*, 68(2), pp.303-338.

Coulson, N. and McMillen, D. (2007). Dynamics of Intraurban Quantile House Price Indexes. *Urban Studies*, 44(8), pp.1517-1537.

Court, A. (1939). *Hedonic price indexes with automotive examples*. NY: Dynamics of Automobile Demand. GM Corporation.

Crawford, I. and Neary, P. (2019). New Characteristics and Hedonic Price Index Numbers. *CEPR Discussion Paper No. DP13445*. Available: https://ssrn.com/abstract=3315378 [27.11.2019].

Czembrowski, P. and Kronenberg, J. (2016). Hedonic pricing and different urban green space types and sizes: Insights into the discussion on valuing ecosystem services. *Landscape and Urban Planning*, 146, pp.11-19.

de Haan, J. and Diewert, E. (2013). Hedonic Regression Methods. In: B. Balk, E. Diewert, D. Fenwick, M. Prud'homme and J. de Haan, ed., *Handbook on Residential Property Price Indices*. Luxembourg: Eurostat.

Diewert, E. (2003). *Hedonic regressions a consumer theory approach*. Available: https://www.nber.org/chapters/c9740.pdf [27.11.2019].

Euro NCAP. (n.d.). *New Car Safety Ratings*. Available: https://www.euroncap.com/en [27.11.2019].

Ewing, G., & Sarigöllü, E. (2000). Assessing Consumer Preferences for Clean-Fuel Vehicles: A Discrete Choice Experiment. *Journal Of Public Policy & Marketing*, *19*(1), 106-118.

Feenstra, R. (1988). Quality Change Under Trade Restraints in Japanese Autos. *The Quarterly Journal of Economics*, 103(1), p.131.

Feldstein, M. (2017). Underestimating the Real Growth of GDP, Personal Income, and Productivity. *Journal of Economic Perspectives*, 31(2), pp.145-164.

Griliches, Z. (1961). *Hedonic Price Indexes for Automobiles: An Econometric of Quality Change*. Available: https://www.nber.org/chapters/c6492 [29.11.2019].

He, L., Chen, W. and Conzelmann, G. (2012). Impact of vehicle usage on consumer choice of hybrid electric vehicles. *Transportation Research Part D: Transport and Environment*, 17(3), pp.208-214.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), p.153.

Hogarty, T. (1975). Price-quality relations for automobiles: a new approach. *Applied Economics*, 7(1), pp.41-51.

Houde, S. (2018). Bunching with the Stars: How Firms Respond to Environmental Certification. *SSRN Electronic Journal*.

Huang, M., Xu, X. and Tashnev, D. (2015). A weighted linear quantile regression. *Journal of Statistical Computation and Simulation*, 85(13), pp.2596-2618.

Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), p.33.

Larson, P., Viáfara, J., Parsons, R., & Elias, A. (2014). Consumer attitudes about electric cars: Pricing analysis and policy implications. *Transportation Research Part A: Policy And Practice*, 299-314.

Liao, W. and Wang, X. (2011). Hedonic house prices and spatial quantile regression. *Journal of Housing Economics*, 21, pp.16-27.

Machado, J.A.F. and Santos Silva, J.M.C. (2000). Glejser's Test Revisited. *Journal of Econometrics*, 97, pp.189-202.

Machado, J. and Santos Silva, J. (2001). *Identification with Averaged Data and Implications for Hedonic Regression Studies*. Available: https://ideas.repec.org/p/ptu/wpaper/w200110.html [29.11.2019].

Machado, J. and Santos Silva, J. (2006). A Note on Identification with Averaged Data. *Econometric Theory*, 22(03).

Machado, J. and Santos Silva, J. (2013). *Quantile regression and heteroskedasticity*. Available: https://jmcss.som.surrey.ac.uk/JM_JSS.pdf [29.11.2019].

Malpezzi, S., Ozanne, L. and Thibodeau, T. (1980). *Characteristic Prices of Housing in Fifty-Nine Metropolitan Areas | HUD USER*. Available: https://www.huduser.gov/portal/publications/affhsg/char_prices_metro_1980.html [29.11.2019].

MarkLines. (n.d.). *Vehicle Sales Data by Country*. Available: https://www.marklines.com/en/ [30.11.2019].

Mount, T., Weng, W., Schulze, W. and Chestnut, L. (2000). *Automobile safety and the value of statistical life in the family*. Available: http://yosemite.epa.gov/ee/epa/eerm.nsf/vwAN/EE-0431-01.pdf/$file/EE-0431-01.pdf [30.11.2019].

Murray, J. and Sarantis, N. (1999). Price-Quality Relations and Hedonic Price Indexes for Cars in the United Kingdom. *International Journal of the Economics of Business*, 6(1), pp.5-27.

OEC (2017). *Economic Complexity Rankings*. Available: https://oec.world/en/rankings/country/eci/ [30.11.2019].

Pakes, A. (2003). A Reconsideration of Hedonic Price Indexes with An Application to PC's. *SSRN Electronic Journal*.

Parkers. (n.d.). *Trusted Car Reviews*. Available: https://www.parkers.co.uk/ [30.11.2019].

Reis, H. and Santos Silva, J. (2006). Hedonic prices indexes for new passenger cars in Portugal. *Economic Modelling*, 23, pp.890-908.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 34-55.

Schamel, G. and Anderson, K. (2003). Wine Quality and Varietal, Regional and Winery Reputations: Hedonic Prices for Australia and New Zealand. *Economic Record*, 79(246), pp.357-369.

Sirmans, S., MacPherson, D. and Zietz, E. (2005). Composition of Hedonic Pricing Models. *Journal of Real Estate Literature*, 13, pp1-44.

SMMT - The Society of Motor Manufacturers and Traders. (2018). *UK Automotive Industry*. Available: https://www.smmt.co.uk/industry-topics/uk-automotive/ [30.11.2019].

Top Gear. (n.d.). *Top Gear Reviews*. Available: https://www.topgear.com/ [Accessed 30.11.2019].

Triplett, J. (2006). *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*. Paris: OECD Publishing.

UK Electric Vehicle Database. (n.d.). *All Electric vehicles*. Available: https://ev-database.org/ [30.11.2019].

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48, p.817.

Wooldridge, J. (2012). *Introductory Econometrics: A Modern Approach*. 5th ed. Cengage Learning.

Wynne, M. and Rodriguez-Palenzuela, D. (2004). Measurement Bias in the HICP: What do we know and What do we need to know?. *Journal of Economic Surveys*, 18(1), pp.79-112.

Xiong, W. and Tian, M. (2019). Weighted Quantile Regression Theory and its Application. *Journal of Data Science*,17, pp.145-160.

Zietz, J., Zietz, E. and Sirmans, G. (2007). Determinants of House Prices: A Quantile Regression Approach. *The Journal of Real Estate Finance and Economics*, 37(4), pp.317-333.

Appendices

Appendix A

Method	Statsitic from RESET	Functional Form		
		Linear (Price)	**Semi-logarithmic (Ln Price)**	**Fully Logarithmic (Ln Price)**
OLS	**Test Statistic**	1001,74	30,94	91,56
	P-value	0,0000	0,0000	0,0000
WLS	**Test Statistic**	345,75	6,24	8,25
	P-value	0,0000	0,0003	0,0000

Fig. 11 *Results of the RESET tests for the OLS and WLS regressions, aiming to find the best functional form for further regressions. The linear, semi-logarithmic and fully logarithmic functional forms are included.*