ADVISORY PANEL ON CONSUMER PRICES – TECHNICAL

Transformation of UK consumer price indices: second-hand cars

Status: Work in progress
Expected publication: For publication alongside minutes

Purpose

1. In 2021 ONS obtained data for second-hand cars from Auto Trader, an online automotive sales advertising business. In this paper we look at how we can use these new data to produce detailed, informative statistics regarding price changes in the second-hand car market and the impact of including new price indices for second-hand cars in UK consumer price statistics.

Actions

2. Members of the Panel are invited to:
   a) consider the appropriateness of proposed methods for calculating second-hand car indices in the UK

Background

3. Accurate measurement of price changes in the second-hand car market is challenging because cars depreciate as they are used. The same car in January 2021 and December 2021 cannot be treated as comparable, as it is an older car and will typically have travelled more miles – there has therefore been a decline in quality.

4. In 2017, Eurostat undertook a review of the approaches taken in EU countries to measure price change for second-hand cars, identifying highly divergent results across EU member states, largely driven by the different methodologies used.

5. Our current UK method involves pricing a sample of 35 models of one-year-, two-year- and three-year-old cars using retail prices from a trade guide. January prices are taken straight from the guide, but in subsequent months prices are interpolated for two- and three-year-old cars to ensure that cars of the same age and mileage are priced each month.

6. Further details of current methods, as well as the Eurostat review, can be found in APCP-T(18)04 Measuring changes in used car prices.

7. The transformation discussed in this paper regards the price indices for second-hand cars only; the calculation of COICOP weights for second-hand cars remains unchanged and is out of scope of this work. These weights are subject to the annual updating of the basket and corresponding weights; and calculated based on net sales to exclude household to household transactions.

8. Second-hand cars had a weight of 12 parts per thousand (1.2%) in CPIH in 2021, and 16 parts per thousand in CPI (1.6%).

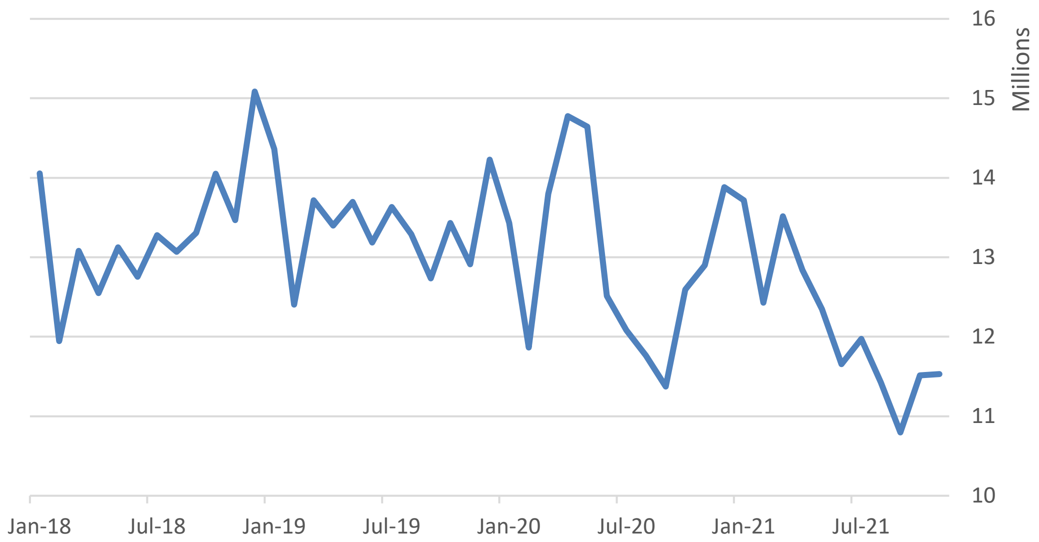
New data

9. In 2021, ONS obtained access to data for second-hand cars from Auto Trader, an automotive sales advertising business, dating back to January 2018. These data are web-provided, therefore no explicit data on sales revenues are available and the data are comparable to

those that would be obtained through web-scraping (though some additional detail is provided in the data that could not be obtained through web-scraping). While the business is an advertising business and not a sales business, it is currently the number one vehicle website in the UK[1].

10. Approximately 13 million price quotes are observed for used cars in the dataset each month (Figure 1), though this is somewhat seasonal and has also been affected by the coronavirus pandemic.

Figure 1: Monthly used car row counts for all listings over time (millions)



11. The data are highly informative including variables such as: date and advertised price, category (car, van or bike), condition (new, nearly new and used), mileage, year of registration, make, model, mark, trim, engine size, fuel type, body type and number of doors. A full data dictionary describing these variables can be found in **Annex A**.

12. Additional data cleaning is carried out to filter unwanted products such as bikes and vans, as well as identifying potential data errors. Further details of the data cleaning carried out prior to this analysis are provided in **Annex B.** Additional, more advanced, outlier detection techniques are being considered for use on our range of alternative data sources, but these methods will be brought to the Panel for consideration at a later date.

New methods

13. Our previous work, and corresponding international guidance, has pointed towards multilateral methods being most appropriate for producing price indices using large, dynamic datasets. There are, however, challenges when trying to use multilateral methods with our data for the second-hand car market.

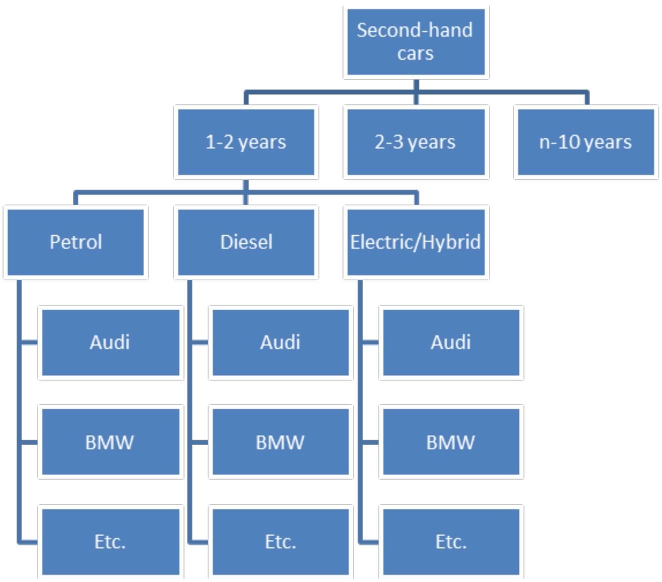
14. An initial challenge is product match. Every car in the dataset is unique and only listed for a short period of time. It therefore cannot be followed through time as we would typically follow products in a CPI collection. The cars that remain on the market for a longer period are likely to be unrepresentative of cars purchased and will be constantly changing in quality if they're

[1] https://www.similarweb.com/site/autotrader.co.uk/#overview

still in use. To resolve this, we propose deriving a product identifier based on a pre-determined set of characteristics, whereby prices will be averaged across cars of similar age and of the same specification.

15. There is also an absence of quantity information when using these web-provided data. We have previously identified that a GEKS-Jevons could be used in the absence of quantity information, but this would result in both popular and rare cars having similar influence in the index so is not our preferred method. As each car registration is unique, in this analysis we limit the dataset to each cars' latest available listing (see paragraphs 27 – 32); when a listing drops out, we consider this a proxy for a sale and propose a quantity of 1 (q=1). Summing the quantities for each product definition means that products that have more items (unique car registrations) assigned to them each month will receive a greater importance within a weighted index.
16. The GEKS-Törnqvist (aka CCDI) and the Geary Khamis indices are the most commonly considered index methods for use with alternative data sources internationally, while the hedonic family of price indexes are also considered suitable by some, particularly in the case of second-hand cars. Our work in choosing the most appropriate index number method is ongoing; for the purpose of this analysis, we focus on a GEKS-Törnqvist (with a mean splice and 25-month window) because of its reduced run-time comparatively to other methods.
17. To aid interpretation of the second-hand cars price index we produce lower-level indices stratifying by age (1-2 years, 2-3 years etc.), fuel-type (petrol, diesel, and electric/hybrid) and make (Audi, BMW etc.) as shown in Figure 2.

Figure 2: New, detailed hierarchy for second-hand cars



18. We aggregate the elementary aggregate indices using a Lowe aggregation based on quantity shares from the previous year. For example, a price index for Audi cars in 2022 would be given a weight equivalent to their share in the 2021 data relative to all car makes. These derived weights are broadly comparable to figures obtained through the DVLA, though the DVLA data is for *all registered* cars, including new cars, whereas the Auto Trader data captures cars purchased in a particular year and can be filtered to just used cars.

Defining a product

19. To define what we mean by a product for second-hand cars there is a trade-off that needs to be addressed between product match and product homogeneity. If the definition is too narrow, there will be a lowered product match rate over time, and if too broad the indices may be influenced by changes in composition. This challenge is discussed in the context of clothing in ONS 2021: [Product grouping: measuring inflation in dynamic clothing markets](#).

20. Figures 3a and 3b present indices and growth rates (respectively) based on three different levels of homogeneity when defining a product (differences in the make, model, mark and trim of cars is explained in **Annex C**):

- Scenario 1 is the most detailed, including the; age, fuel-type, make, model, mark, trim, engine size, mileage bin, transmission, and body type (e.g., 2-year-old, Petrol, Honda Civic Mk10 [mark] SE [trim], 1.0L, 30,000 miles <= 40,000 miles, Manual, Hatchback)

- Scenario 2 uses the same specifications but is agnostic to trim (e.g., 2-year-old, Petrol, Honda Civic Mk10 [mark], 1.0L, 30,000 miles <= 40,000 miles, Manual, Hatchback)

- Scenario 3 uses the same specifications but is agnostic to both trim and mark (e.g., 2-year-old, Petrol, Honda Civic, 1.0L, 30,000 miles <= 40,000 miles, Manual, Hatchback)

Figure 3a: Index values for second-hand cars using new data and methods, January 2018 = 100

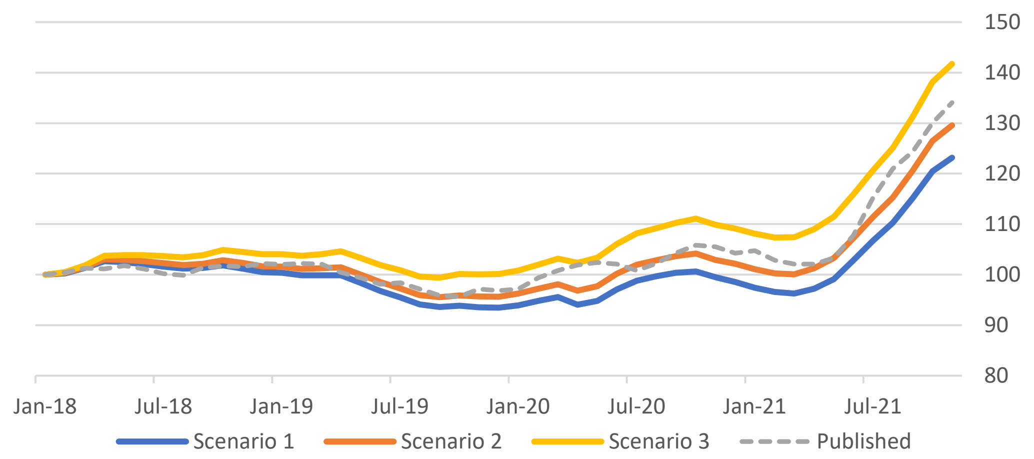
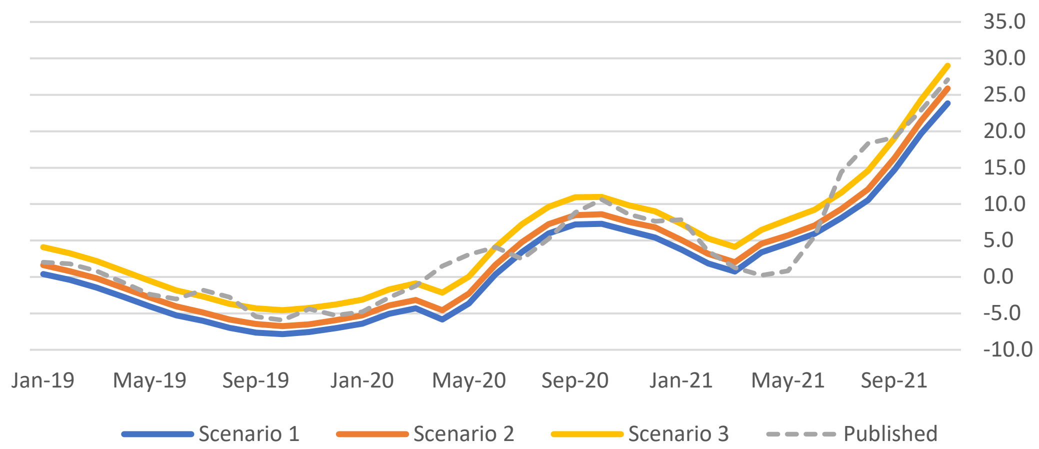


Figure 3b: Annual growth rate for second-hand cars using new data and methods, %



21. These results suggest that defining a product more tightly tends toward lower inflation than when the definition of a product is broader, and this pattern of results was found consistently across different strata. This could be because of compositional effects, where the quality within a product definition that is too heterogeneous may be increasing over time (particularly in the case of scenario 3). However, it could be that defining the product too narrowly is leading to a lower product match rate, and this could result in a smaller sample of products being used to construct the index (particularly in scenario 1).

22. To further understand the results, we can implement the "Match Adjusted R Squared" (MARS) method as described in Chessa (2019; PDF, 860KB), though we cannot effectively weight our product match score in the absence of sales quantities for each unique car registration (item).

23. MARS assesses how well a product has been defined using the match rate (based on the proportion of individual items in the current month that could be allocated to a product definition that existed in the chosen base period) and a homogeneity metric (based on the proportion of explained variance in product prices, relative to the total variance in item prices, R squared).

24. Our MARS scores for all 3 scenarios are high (Table 1). Scenario 1 sees an extremely high R squared value suggesting a high product homogeneity but suffers a lower product match score than the other scenarios. Scenario 3 sees the highest product match but suffers a lower R squared value, suggesting lower homogeneity than for the other scenarios. Scenario 2 appears to manage this trade-off the best, resulting in the highest MARS score (see **Annex D** for more detail).

Table 1: Average MARS scores, product match scores and R squared values for Scenarios 1-3, Jan 2019 – Nov 2021

	MARS	**Product Match**	**R Squared**
Scenario 1	**0.89**	**0.90**	**0.98**
Scenario 2	**0.93**	**0.96**	**0.96**
Scenario 3	**0.91**	**0.98**	**0.93**

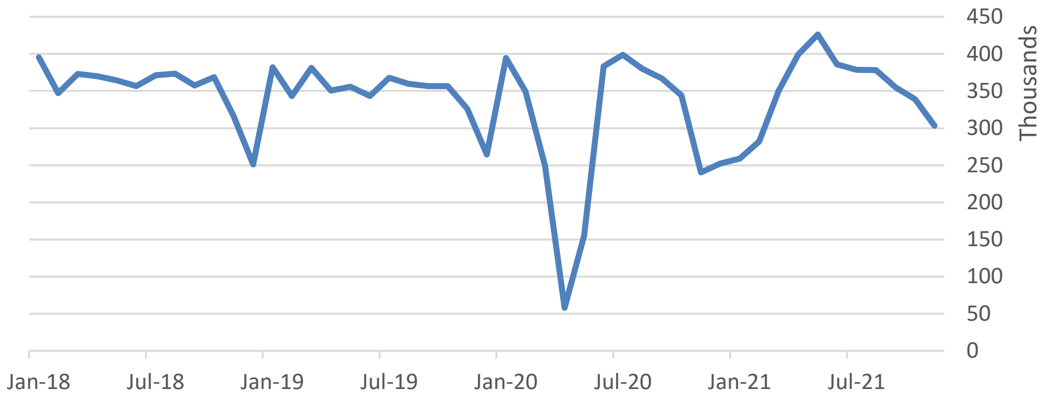
25. It could be considered that when the product match score crosses a certain threshold that the impact of churn is not a cause for concern, and more weight should be placed on the homogeneity metric to reduce the impact of compositional effects. Furthermore, the product match score in our work is not optimal, as it is unweighted and relies on a fixed base period. A fixed base doesn't necessarily encompass true product match in our data when calculating multilateral indices, as their calculation doesn't rely on having matched products with a single fixed-base period. In future work we will consider use of a rolling base period to measure product match.

26. The product match score does, however, give us something of an indication of remaining churn in the data and, when calculated consistently, gives us a reference for how different product definitions perform relative to each other. We have therefore opted to use Scenario 2 for the remaining analysis based on it having a consistently higher MARS score when considering homogeneity and product match equally.

Prices indices for latest listings compared to price indices for all advertised cars

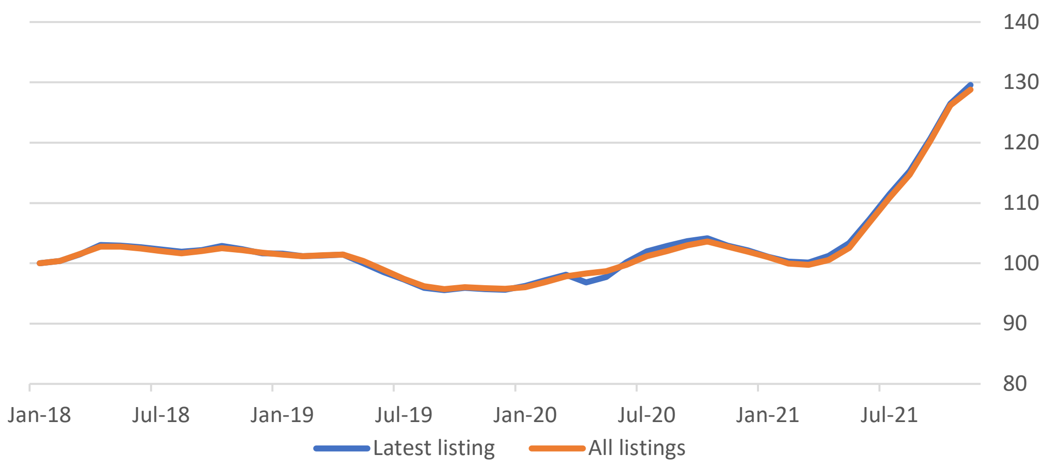
27. Our data for second-hand cars are somewhat unique in that each car will (typically) only be sold once within a given window (in our case 25 months). They are based on advertised prices, meaning that a unique car may be advertised in more than one consecutive period, but will not have been sold until the latest point of advertisement. The price may have changed during the car's period of advertisement, despite the car having not been purchased in this time.
28. To ensure that price changes for unsold cars do not unduly influence the price index we look at the impact of filtering the dataset so that only the latest available list price for each unique car registration is available (when sold in consecutive periods), as a proxy for when cars have been sold. Figure 4 shows the monthly row counts when applying this filtering.

Figure 4: Monthly used car row counts for unique ('latest') listings over time (thousands)



29. When using all listings (Figure 1) we observed an increase in the number of cars listed on the website in April 2020, this represents the supply of second-hand cars listed on Auto Trader. When filtering to the latest listing, this trend is reversed. Demand was likely low for second-hand cars during this month as it was the first national lockdown, as a result of the coronavirus pandemic, where people were only allowed to leave their homes for 'essential' purposes. Therefore, reducing to the latest listing appears to provide an approximate for when a car has been sold. In Figure 5 we compare the index constructed using data for just the latest listing to an index calculated based on all advertised prices.

Figure 5: Index values for just the latest listing compared to all listings, Jan 2018 = 100



30. The difference between these measures in this case is negligible in most periods; this is likely because cars are only listed on the website for a short period (average 40 days), so will often fall inside the same month of index calculation. If they are available in more than one time period than it may be simply that their price doesn't typically change between these periods.

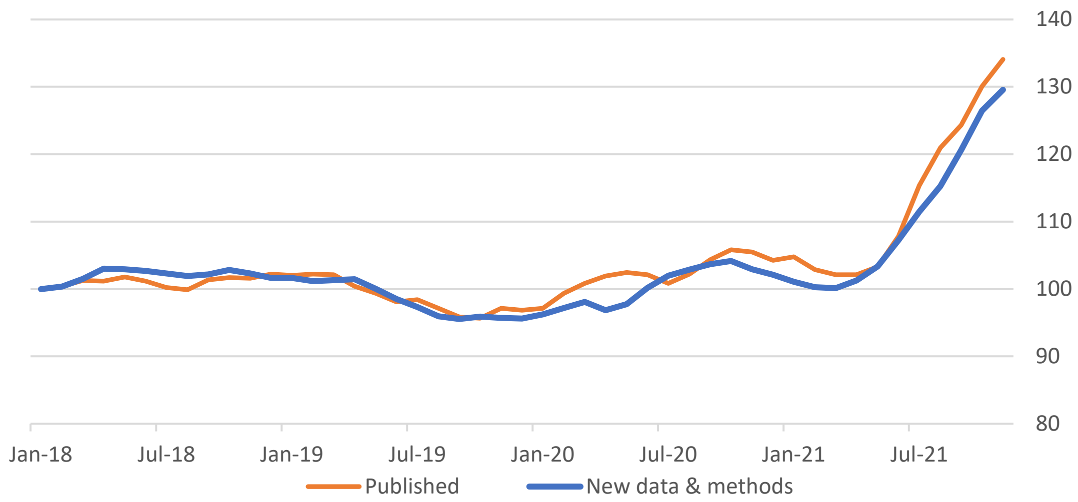
31. An exception to this is observed in April 2020, the first full month of the restrictions imposed due to the coronavirus pandemic, where we see a fall in price in the latest-listing measure but not the all-listings measure. Demand for second-hand cars was likely low during the initial months of the pandemic due to the resultant economic shock and the movement restrictions that were imposed.[2] This could very likely have resulted in a temporary fall in car prices. This may not have been identified through the measure based on all available listings because many prices were carried forward from previous months due to cars not being sold.

32. Given the measure based on the latest available listing appears more responsive to sudden changes in price, we propose this is the most suitable method for calculating second-hand car indices using these data. Furthermore, using the latest listing will more accurately capture demand and will avoid the price movement issue previously discussed (paragraph 27).

Comparing new and existing indices for second-hand cars

33. Figure 6 shows our currently published index compared to the index produced using new data and methods. Broadly, the indices show similar trends throughout this time period. The index based on new data and methods shows a slower rate of growth (and greater deflation in some periods).

Figure 6: Comparison of the second-hand cars index using new methods and data to the published series, Jan 2018 = 100

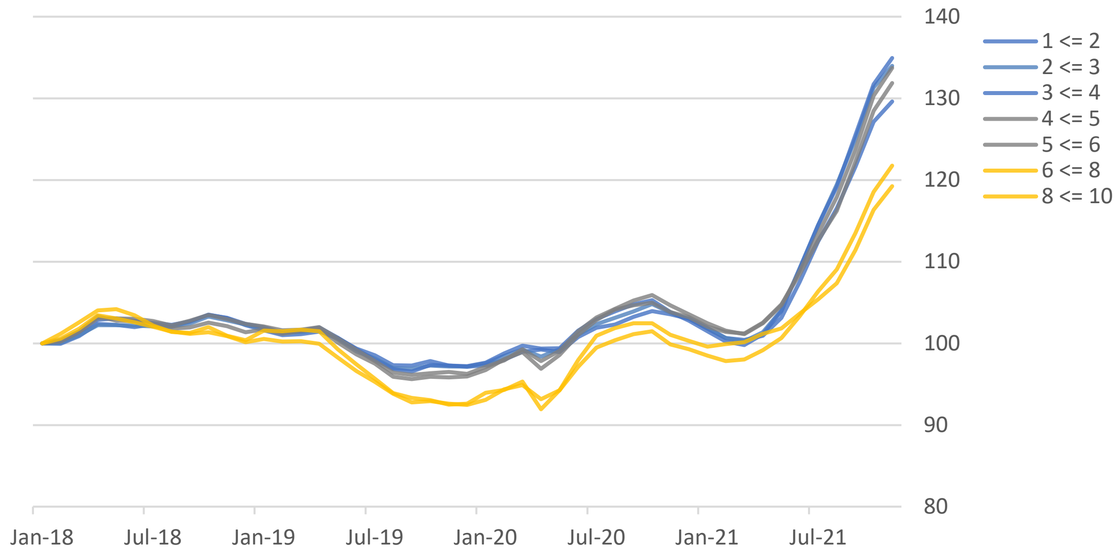


34. Our indices are now inclusive of a huge variety of different ages and models of cars, fuel types and mileages. These differences in the data and methods naturally lend themselves to different results. Of particular interest here is in the age of the car as, while our current sample captures a range of makes and models, indices are only produced for 2- and 3-year-old cars.

[2] https://www.ons.gov.uk/economy/grossdomesticproductgdp/bulletins/gdpmonthlyestimateuk/april2020

35. Figure 7 shows that newer cars (blue lines) tend towards higher inflation than older cars (yellow lines), that could help explain some of the difference between our new aggregate index and the currently published figures (though given the differences in samples and methods, the new indices for similar-aged cars are remarkably close to the previously published indices).

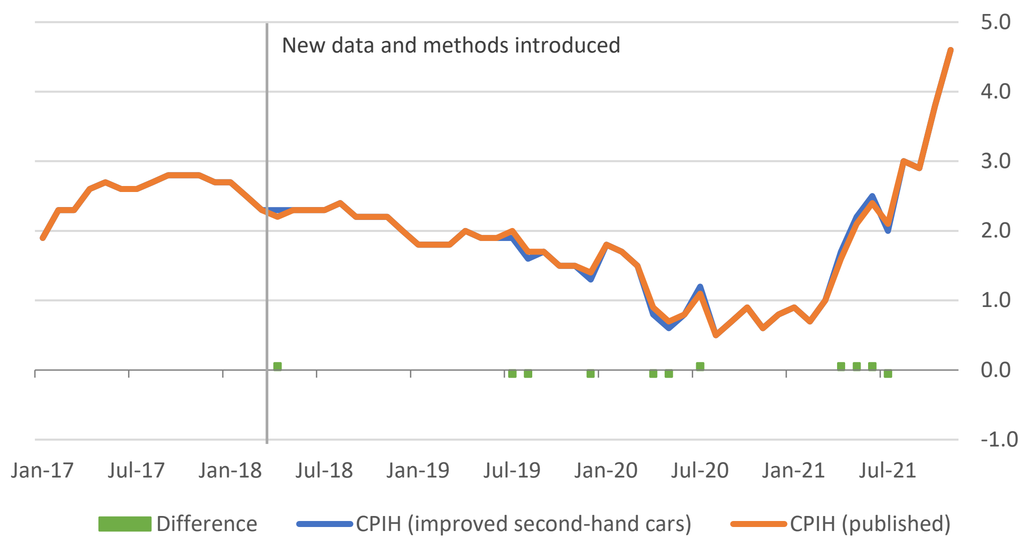
Figure 7: Stratum indices for different ages (in years) of car, Jan 2018 = 100



Impact of new data and methods for second-hand cars on headline consumer price statistics

36. The aggregate index for CPIH (Figure 8) has been produced between January 2017 and November 2021, including the new second-hand cars index from February 2018 onwards so that growth rates in the year of introduction can be seen, as well as annual growth in the years following introduction. The new index is aggregated together with published series using the existing annual weights and chain-linking methodology.

Figure 8: Impact of new data and methods for second-hand cars on CPIH annual growth rate (%)



37. While our new index for second-hand cars shows a slightly lower rate of inflation than our currently published index, the impact on our headline indices as a result of this change is marginal, even in the year of introduction. The maximum difference is 0.1 percentage points, and this is also the maximum impact on CPI.

38. Note that since March 2020 there have been a number of unavailable items that have been imputed in some periods based on price movements of the headline index. For this impact analysis we haven't recalculated these imputations due to the complexity of their calculations, but we would expect the impact of recalculating these to be negligible based on the minimal impact of these new data and methods on the headline indices.

39. Impacts of the introduction of these data are discussed further in **APCP-T(22)03 Transformation of UK consumer price indices: Impact Analysis, 2022**.

Future work

40. There is still some remaining work to refine our indices for this category as we ready for our first publication of experimental statistics using these data in May 2022. We do not expect any of the remaining work to significantly change the impacts presented in this paper.

41. In particular we are:
    a) trying to identify an alternative source of data to use to weight stratum level indices, based on expenditure, particularly to remove household to household sales
    b) making some ongoing improvements to the data cleaning and filtering, such as considering whether filters should be conditional on each strata (e.g., whether newer cars need stricter bounds on mileage), as well as broader outlier detection work
    c) looking at how these results compare to those of other index number methods, such as the time dummy hedonic method
    d) considering whether to remove hybrid and electric vehicles (until their consumption increases) due to their small sample sizes that are resulting in volatile elementary aggregate indices – though these have little impact on higher aggregates due to the small weight (~ 3%) these cars currently have
    e) considering whether the partial post code we have been provided can be used to produce regional indices for second-hand cars
    f) completing a more in-depth check into our chosen method for defining a unique product to ensure consistency and the appropriateness of homogeneity and product match, in particular, looking at using a rolling base period when calculating product match scores.

David Moran & Helen Sands
Prices Division, Office for National Statistics
January 2022

List of Annexes

Annex A	Second-hand cars data dictionary
Annex B	Data cleaning and filtering
Annex C	Make, model, mark and trim explained
Annex D	Product match and MARS results

Annex A – Second-hand cars data dictionary

Variable	Description	Example
stock_item_id	Unique identifier for vehicle listing	2c9299cf73e81eb50173ec0d4e346f21
category	Type of vehicle	Car, Bike, Comm
mileage	Odometer reading (overall vehicle mileage)	38,500
price	Listing price of vehicle	8995
public_reference	Unique identifier for vehicle listing	202008142467658
year_of_registration	Year vehicle first registered	2018
vehicle_condition	Vehicle condition	Used, nearly new, new, virtual
standard_make	Manufacturer of vehicle	Ford, Volvo, Honda etc.
standard_model	Specific model made by manufacturer	Fiesta, V70, Civic etc.
vat_excluded	If VAT is included in price (relevant to commercial vehicles)	y/n
derivative_badge_engine_size_text	Engine size	1.0, 2.0, 1.5 etc.
derivative_badge_engine_size_unit	Engine size unit	Litre / Cubic Centimetre
derivative_body_type	Type of vehicle body (hatchback, saloon)	Hatchback, saloon, SUV etc.
derivative_doors	Number of doors	3,5 etc.
derivative_drivetrain	How the wheels are powered	Front-wheel drive, Four-wheel drive
derivative_emissions_standard	European emissions standard	EU1, EU2, EU£, EU4, EU5, EU6
derivative_fuel_type	Fuel type	Petrol, diesel etc.
derivative_mark	Generation of a particular car model which identifies a redesign	Mk8, Mk5, C117
derivative_seats	Number of seats	5, 7 etc.
derivative_transmission	Gearbox type (manual / automatic)	Manual / Automatic
derivative_trim	Additional specification type for a particular model	GTI, SE, Zetec etc.
insurance_group_1_to_50	Insurance groups from 1 (cheapest) -50 (most expensive)	06E, 38A
is_retailer	Identifier for private or trade vehicle listing	y/n
writeoff_category	Insurance write-off categories	C / D / N / S
postcode_district	First half of postcode	SE22, M15, PE1 etc.
date_of_registration	Date of first registration	2019-09-30
colour	Unique vehicle body colour	Black, Red, Silver etc.
derivative_badge_engine_power_text	Engine Power	150, 190, 130 etc.
derivative_badge_engine_power_unit	Engine Power Unit	BHP/PS

Engineered variables

Variable	**Description**	**Example**
collection_date	Date created from filename path. This is the day the data is received.	2018-02-07
retailer	‘autotrader’ added for every row (for use when aggregating with other data sources)	autotrader
mileage_per_month	Derived from ‘mileage' & 'date_of_registration'. From the 'date_of_registration' column create a column for how many months old the car is. Needed for Time Dummy Hedonic method.	489, 750 etc.
mileage_bins	Mileage bins created for stratification and product grouping. Bins based on knowledge that on average a car drives 7,400 miles a year and that the newer the car the more significant the difference in mileage would be (thus smaller to larger bins as mileage increases)	- 0 <= 5000 - 5000 <= 10,000 - 10,000 < = 20,000 - 20,000 <= 30,000 - 30,000 <= 40,000 - 40,000 <= 60,000 - 60,000 <= 80,000 - 80,000 <= 100,000 - 100,000 <= 140,000 - 140,000 <
age_month	Age in number of months derived from last ‘collection_date’ available and ‘date_of_registration’. This means final age of car is determined when the listing drops out of the data.	92, 26 etc.
age_bins_year	Age bins created for stratification derived from ‘age_month’	- 0 <= 1 - 1 <= 2 - 2 <= 3 - 3 <= 4 - 4 <= 5 - 5 <= 6 - 6 <= 8 - 8 <= 10 - 10 <= 12 - 12 <= 14 - 14 <= 16 - 16 <= 18 - 18 +
fuel_type_bins	Created for stratification, derived from ‘derivative_fuel_type’	- petrol = 'Petrol', 'Bi Fuel' - diesel = Diesel - hybrid/electric = 'Petrol Hybrid', 'Petrol Plug-in Hybrid', 'Diesel Hybrid', 'Electric', 'Diesel Plug-in Hybrid', 'Hydrogen'
productid_ons	A concatenation of selected variables to product group and produce indices from. The three scenarios in this paper are created here	

Annex B - Data cleaning and filtering

1. As discussed in this paper (paragraphs 27 - 32) there is a potential issue with the data to resolve where we have multiple listings of a unique vehicle over time. This may be a problem as the price can change over time (increase or decrease) which we wouldn't want to bias our index. To carry out the analysis in this paper we provided two options: 'all listings' and 'latest listing'. To implement the 'latest listing' option we filtered the data on 'stock_item_id' (the unique listing identifier) to return the latest listing for each unique car registration.

2. The data include vehicles that are not cars. To exclude these data, we have filtered on the variable 'category' to only keep 'cars', this excludes the other two options of 'bike' and 'comm'. The 'comm' category is an abbreviation of commercial which captures vans in the data. By excluding 'comm' and 'bike' this removes 10.83% of the data.

3. To ensure we are only using *second-hand* cars we have set a filter on the 'vehicle condition' variable to keep 'used' and 'nearly new', this excludes 'new' and 'virtual'. 'virtual' is a condition which only applies to bikes. Some investigation was carried out on the 'nearly new' condition to ensure this does not capture any cars we would consider 'new'. The average mileage for this condition is 3226 miles compared to 5 miles for new cars which does suggest these are not new, on average. To ensure no data are included that we may consider new, a lower limit has been filtered on mileage (see 5.).

4. As we are planning to produce consumer price indices with these data, we need to consider what counts as a second-hand car sale. This means dealing with the issue of excluding the purchase of second-hand cars directly from other households. The purchase and sale of second-hand cars among households on the economic territory will cancel out, meaning that the net weight is nil for these transactions. From the Harmonised Index of Consumer Prices (Eurostat) manual:

   > '*Purchase of vehicles includes purchases by households of new vehicles and net purchases by households of second-hand vehicles from dealers. Sales of second-hand vehicles between households are excluded*'

   Therefore, we obtained a private and trade listing identifier from Auto Trader. This binary variable effectively captures all dealerships and small independent traders as a trade sale. An individual seller would be classed as private. Thus, with the binary 'is_retailer' variable we have filtered to keep only trade. This excludes 7% of the data we have already filtered as 'cars' and 'used' / 'nearly new'.

5. Filtering has been applied to some of the numerical variables to omit potential errors. This has also been done to exclude cars which are not bought by most consumers or would skew expected trends e.g., supercars, classic cars etc. Data counts below refer to data already filtered by paragraphs 1-5:

   a. Mileage:
      i. Filter to keep between 100 miles and 250,000 miles
      ii. Median mileage = 38,000 miles
      iii. 99% of data within those bounds

   b. Price:

        i. Filter to keep between £200 and £60,000
        ii. Median price = £9,200
        iii. 99% of data within those bounds
    c. Year of Registration
        i. Condition to keep year (t): t-20
        ii. 99% of data within t-20 currently
    d. Engine Size
        i. Filter to keep top 20 engine sizes which accounts for 98% of data: 0.9, 1, 1.1, 1.2, 1.25, 1.3, 1.33, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.4, 2.5, 2.7, 3
        ii. Engine size is one of our variables which concatenates into our product grouping - ‘productid’. Limiting the amount of engine sizes here prevents small group sizes for engine sizes considered outliers.

6. As ‘make’ is being used for stratification we have filtered the data to the top 25 makes which accounts for 97% of the data. This avoids having any strata with not enough data to produce a meaningful index. The top 25 makes are based on the shares within the data and are: Ford, Vauxhall, BMW, Volkswagen, Audi, Mercedes-Benz, Nissan, Peugeot, Toyota, Land Rover, Renault, Mini, Citroen, Kia, Hyundai, Honda, Fiat, Volvo, Seat, Mazda, Skoda, Jaguar, Suzuki, Porsche, Mitsubishi.

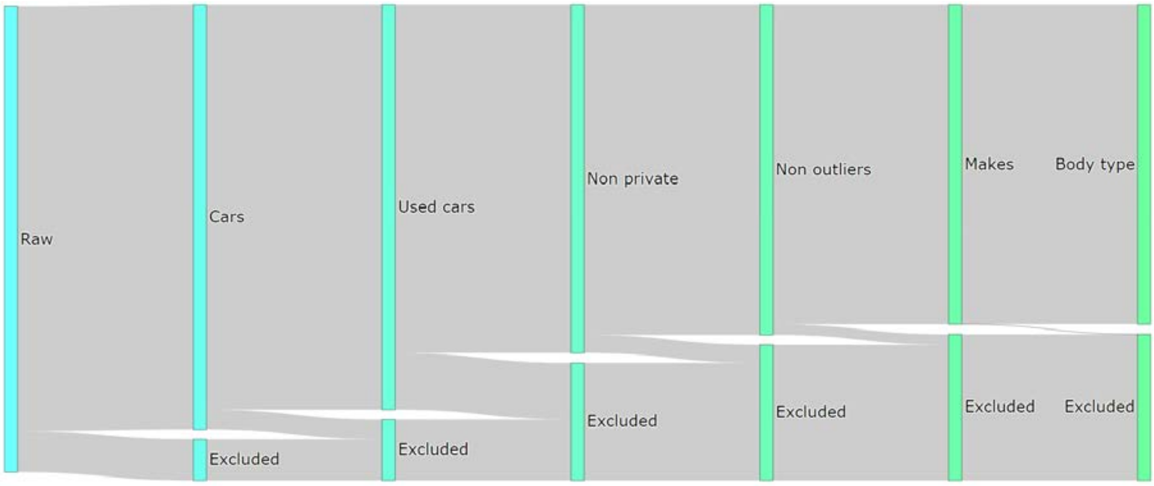
7. To further ensure we are only capturing cars we have also filtered on the ‘body type’ variable. This was done as some non-car body types were being mis-categorised within the ‘cars’ category. Thus, we have set a filter to keep the only ‘car’ body types within the dataset: 'Hatchback', 'SUV', 'Saloon', 'Estate', 'MPV', 'Coupe', 'Convertible'. This excludes a further 3.24% of the remaining data, which is mainly accounted for by nulls (3.2%).

8. A unit conversion was required for some data we had in the ‘engine size’ variable and corresponding ‘engine size unit’ variable. There were two units, litre and cubic centimetre. We converted all units to litre and converted all engine sizes, with the corresponding unit of cubic centimetre, to litres by dividing by 1000.

9. Further filtering was applied to our age variable (beyond paragraph 6c) when applied as strata to the index. In Annex A we described that age bins were created to group the precise age of the cars into more useful categories for stratification purposes (‘age_bins_years’). To ensure the indices were truly reflective of second-hand cars we decided anything less than 1 year of age should be out of scope. To ensure representivity of the indices to consumer purchases we also decided to not include cars older than 10 years, the indices for these age bins were more volatile due to sample size issues. Over the four years we have data for 75% (on average) falls within the 1-10-year-old age bins.

Figure 9: Flow of data through the data cleaning process

Figure 9 incrementally shows the remaining data after sequentially filtering through paragraphs 2-7 in Annex B. Starting point 'Raw' is already filtered to 'latest listing' (paragraph 1):



Annex C - Make, model, mark and trim explained

- **Make** = Manufacturer e.g., Ford, Volvo etc.
- **Model** = Different type of car within manufacturers range e.g., Ford **Fiesta,** Volvo **V70**
- **Mark** = A new generation of a particular model when the model undergoes a complete redesign. Each generation (mark) roughly last around 4-6 years. e.g., Ford Fiesta **Mk5** (2002 – 2008)
- **Trim** = A particular version of a model which identifies a vehicle's level of equipment or special features. Trim packages include base level, appearance changes, sports upgrades, handling upgrades, safety upgrades etc. e.g., Ford Fiesta Mk5 **Zetec.**

>>>>>>>>> Order of granularity

Make → **Model** → **Mark** → **Trim**

Examples:

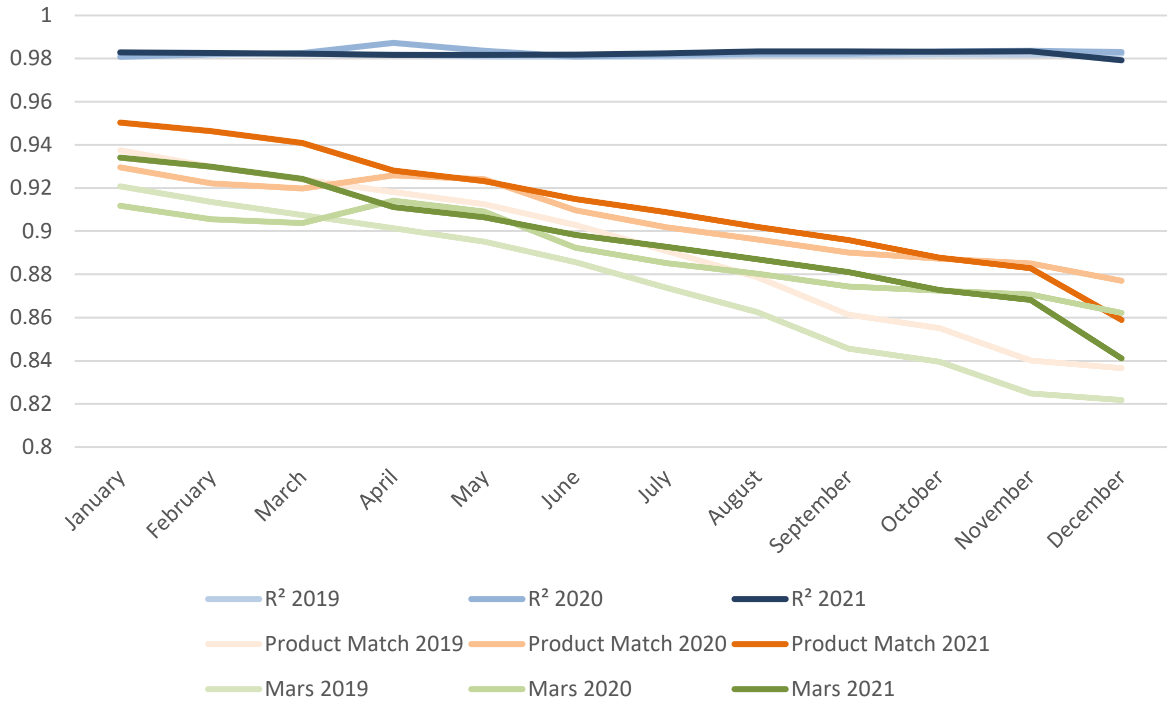
Volkswagen → **Golf** → **Mk8** *(8th Gen)* → **GTI** *(Grand Tourer Injection)*

Honda → **Civic** → **Mk7** *(7th Gen)* → **SE** *(Sports Edition)*

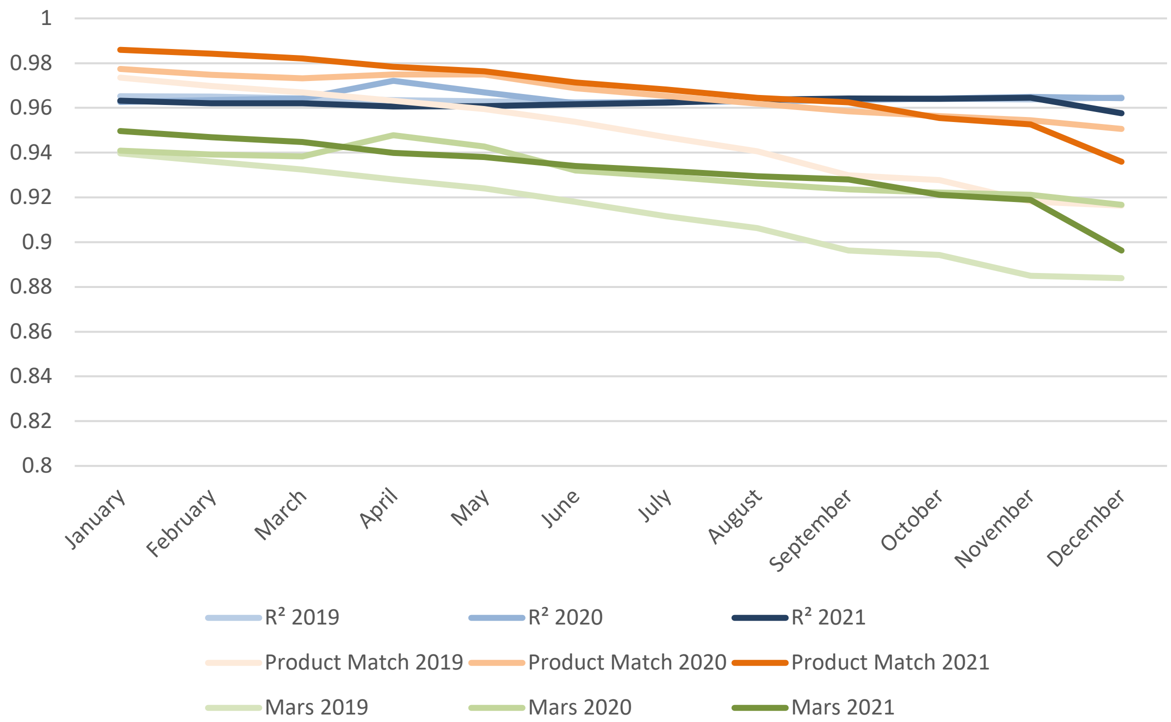
Annex D – Product match and MARS results

Each Scenario is tested with a base period of Q1 in year t-1

Scenario 1 (Model, Mark, Trim):



Scenario 2 (Model & Mark):



Scenario 3 (Model):