Office for National Statistics

Government Buildings
Cardiff Road
Newport
NP10 8XG

Michael Hardie, Matthew Corder
Tel: (+44) 01633 455923
Email: Michael.harde@ons.gov.uk
Tel: (+44) 02037 411759
Email: matthew.corder@ons.gov.uk
www.ons.gov.uk

30 November 2022

Simon Hayes and Rhys Phillips
Head of Current Economic Conditions Division, and Head of Sterling Markets Division
Bank of England
Threadneedle Street
London
EC2R 8AH

Dear Simon and Rhys,

This letter covers changes to the suite of consumer price inflation statistics in 2023.

i. Issue

A request for the Bank of England to assess the changes to the Retail Prices Index (RPI) proposed by UK Statistics Authority (the Authority) for implementation in 2023.

ii. Action

To note the contents of this letter, which is being sent to acknowledge the requirements under Section 21 of the Statistics and Registration Service Act 2007. In previous years it has initiated the Authority's consultation with the Bank over whether any proposed changes to the RPI would constitute a fundamental change in the index which would be materially detrimental to the interests of holders of relevant index-linked gilts, and hence trigger the redemption clause.

iii. Timing

For response by 2 January 2023.

iv. Context

In November 2020 we published a response to the consultation on reforming the Retail Prices Index methodology. This consultation was run jointly with HM Treasury (HMT) and sought views on the manner in which we should implement our plans to reform RPI (by bringing in the methods and data sources of the Consumer Prices Index including owner occupiers' Housing costs, CPIH). HMT sought views on whether and when the change should be made between 2025 and 2030. The (then) Chancellor responded that, while he saw the statistical arguments of our intended approach to reform, he was unable to consent to the implementation of such a proposal before the maturity of the final specific index-linked gilt in 2030.

It is UK Statistics Authority policy to address the shortcomings of the RPI in full at the earliest practical time. Considering the outcome of the consultation, the change that we propose can legally and practically be made by the Authority in February 2030. Since the earlier 2012 consultation on RPI, our policy has been to only make "routine" changes to RPI unless such changes would impede the development of our other consumer price inflation statistics.

There are increasing challenges associated with maintaining the RPI's current methodology until 2030; challenges which are particularly acute in the context of the large-scale transformation project for CPIH and CPI, which we will be delivering from 2023 onwards. To not make such changes to RPI would require us to maintain a separate data collection for RPI, duplicating efforts and increasing the risk of error in these indices. Taking into account also our intention to address the shortcomings of RPI in full in 2030, we are developing a framework to determine which changes we will propose for RPI for implementation from 2024 onwards.

We propose several changes to the RPI this year, based primarily on the view that:-

1. In order to not make such changes, resource would need to be diverted to maintain price collection for RPI separately, which would hinder the development of our other consumer price inflation statistics, increase the risk of error in RPI, CPIH and CPI, and would not reflect value for money, and
2. that these changes do not reflect a partial reform of RPI as, whilst these changes represent a significant improvement in the methods and data sources for a specific part of the basket, the shortcomings of the RPI that we have previously identified remain; we maintain that the RPI is a poor measure of inflation, and our policy remains to address its shortcomings in full in 2030, as described above.

We have made these decisions to allow us to begin the delivery of the transformation of our consumer price inflation statistics in 2023, as planned. These decisions may or may not be in line with the decision-making framework for RPI changes which we are currently developing, taking into account feedback from our user groups. These decisions also should not be seen as establishing a precedent for the kinds of changes that will be made to RPI in future years, as we implement further transformational changes into our consumer price inflation statistics. We discuss the planned changes in more detail below.

Impact of COVID-19 on our routine processes for consumer price inflation statistics

Since our last letter to you in December 2021, we have largely been able to carry out our routine activities without the need for the contingencies that were implemented in response to the coronavirus (COVID-19) pandemic. This includes the continuity of data collection, methods for dealing with prices that we have not been able to collect, and the updating of annual weights to reflect significant changes in consumer spending patterns. Contingency plans were communicated through correspondence with the Deputy Governor in April 2020, then again through our usual annual correspondence in November 2020 and December 2021. Nevertheless, there remain some activities where we have not yet fully resumed our pre-pandemic approach. I will describe these in more detail in the sections that follow.

v. Ongoing collection of prices

Over the course of the COVID-19 pandemic, we were flexible with our approach to data collection in response to changing Government guidance. We have now resumed our price collection in physical stores, reflecting the absence of movement restrictions in 2022.

vi. Annual basket and weights update

The standard procedures for the annual updating of the baskets are well rehearsed. Although fixed within each year, the contents of the baskets of goods and services and their associated expenditure weights are updated annually to ensure that they are representative of household spending patterns. The updating mitigates potential biases that might otherwise occur from not allowing for changing consumer expenditure habits.

The annual update of the suite of consumer price inflation statistics will take effect with the February 2023 indices, which will be published on 22nd March 2023. An accompanying article describing the changes to the baskets will be published on the ONS website about a week earlier. A full description of the reweighting and updating process can be found in the basket article and the weights article for 2022. The items in the Consumer Prices Index including owner-occupiers' Housing costs (CPIH), Consumer Prices Index (CPI) and Retail Prices Index (RPI) baskets will be updated in 2023, so that they remain representative of consumer spending and trends.

However, our suite of consumer price indices use a Lowe index methodology, meaning that expenditure data from before the base period are used to weight consumption classes. Given the sizeable shifts in consumption patterns observed during 2020 and 2021, our usual weights update process was adapted for 2021 and 2022, for the CPIH and CPI, to ensure that the expenditure weights used adequately reflected expenditure in the base year. The RPI process remained, and continues to remain, unchanged, as discussed in the correspondence for the year 2021.

Our approach for the construction of CPIH and CPI weights for 2022 will depend on international guidance, as well as the most up to date Household Final Consumption Expenditure data (HFCE) released in December 2022. However, we expect that the approach described in the following two paragraphs will be used for the upcoming weights update process.

Despite the absence of COVID-19 pandemic restrictions in 2022, it is still necessary to maintain our adapted weights updating process. This is because the HFCE data that we would typically use refer to the year 2021, when some COVID-19 restrictions were in place. Therefore, spending in the base year, which is 2022, is likely to be distributed in a different way to that suggested by the 2021 data.

For the upcoming 2023 weights for CPIH and CPI, we will therefore update the 2021 HFCE in consumption classes where there have been notable shifts in consumer spending since 2021. The most recent HFCE estimates for the first 3 quarters of 2022 will be used as the primary data source for these adjustments, with other data sources (such as retail sales and financial transaction data) used to supplement HFCE estimates where there is limited information. Quarter 4 spending will need to be estimated by applying the typical seasonal growth. This approach is coherent with previous Eurostat guidance provided for the internationally comparable Harmonised Index of Consumer Prices (HICP, which is the same measure as our CPI).

As with last year's weights update, the RPI will continue to be weighted using annual Living Costs and Food survey data at a lag of 6 months. This means that the 2023 weights will be based on spending patterns between July 2021 and June 2022 and will therefore reflect spending patterns immediately after the period of movement restrictions that were imposed during the course of the COVID-19 pandemic.

Changes in the contents of the baskets and their associated weights are not significant beyond their primary aim of ensuring the continuous and proper representation of consumer expenditure habits.

vii. Location rotation and re-enumeration

Approximately 140 locations in the UK are visited each month as part of the local price collection. These locations are carefully selected according to retail turnover and controlled to avoid overlaps with current locations within the sample. To maintain the sample of locations, each year 30 locations are refreshed, either by excluding a location and replacing it with a new one (rotation) or refreshing the list of outlets in the existing location (re-enumeration). Due to movement restrictions in 2020, it was not possible for our field collectors to carry out this routine process for 2021 and, due to ongoing difficulties, a reduced number of 14 locations were re-enumerated for 2022 with no locations being rotated. We have however since refreshed a full 30 locations again in readiness for our 2023 collections.

Other changes to the suite of consumer price statistics in 2023

viii. Changes to the price collection and lead times for foreign holiday items

We are proposing changes to three of our foreign holiday items (hotels, self-catering, and cruises) which appear in the RPI, CPI and CPIH. Primarily, we plan to collect prices six months in advance of the holiday, rather than collecting all prices for a given year in an arbitrary order within August to February the year before. This will: bring these items in line with our other similar foreign holiday items (such as city breaks and late-booked holidays), ensure greater consistency within a given price relative, more closely match consumer behaviour and enable us to conduct real time checks. Secondarily, we plan to collect the prices of holidays taking place in index week, rather than holidays taking place on/around the first of the month, to more closely align with the majority of the basket. We also plan to drop child prices from our collection for hotels, as data limitations mean that they form only a small part of our sample.

We have conducted an impact analysis by re-working a pilot collection, in which prices were collected three months before the holidays took place. This re-working comes with caveats (more detail in the Annex). This analysis is expected to overestimate the volatility of the proposed changes. Also, due to the current index method[1], extreme price relatives in one month will affect every other subsequent month in that year, and this should be taken into consideration when interpreting the impacts. Therefore, these estimates should be regarded as upper bounds. The analysis found that the mean monthly effect on the RPI all-items index level (Jan-87 = 100) for 2017-2019 was +0.004 index points, while the mean effect on the all-items annual rate was -0.016

[1] Under the current methodology, price relatives are formed by comparing prices in the current month with prices for the same holiday in the same month a year earlier. This differs from the rest of the basket where price relatives are based in January of the same year. Seasonal weights are then used to calculate a rolling-year average of price relatives, resulting in a smoothed and non-seasonal index series. The methodology is described in more detail in section 9.5.7.1 of our Consumer Price Indices Technical Manual.

percentage points, with a mixture of upward and downward effects observed. The changes introduced here can be implemented in Feb-24 data at the earliest, as the agreed method requires a year's worth of base prices before indices can be calculated. We intend to discontinue the traditional collections, starting with the one which would usually occur Aug-23 to Feb-24.

A more detailed analysis of the impact is provided in Annex A of this letter.

ix. Changes to the measurement of rail fares price indices

The following section describes changes we aim to make to the rail fares index in CPIH, CPI and RPI. We have yet to publish our full impact analysis and some internal decision-making processes have yet to be completed. We will only proceed with these changes if we are satisfied that we have robust data sources, methods, systems and processes in place. We will publish our final decision in January next year.

Currently the rail fares price index is calculated annually for Great Britain, by applying the increase in the cap on rail fares as reported each year (typically based on the previous year's July annual rate for the Retail Prices Index plus a fixed amount) in the month in which it comes into effect, which is usually the following January (March in 2021 and 2022). Northern Ireland price changes are sourced separately and are aggregated together with the constructed index for GB using appropriate weights.

From February 2023 (published in March 2023) we intend to make use of rail fares data for Great Britain (GB), provided to us by the Rail Delivery Group and sourced from the rail industry's Latest Earnings Networked Nationally Overnight (LENNON) ticket revenue system; this system does not retail (as in, it does not act as a ticket selling platform). As they are transaction level data, they contain information on the cost and quantities of purchased tickets along with associated metadata related to the journey travelled. We have been provided rail transaction data for GB dating back to January 2019. Northern Ireland data will continue to be provided separately and will become a separately published index in the CPIH, CPI and RPI baskets.

The size, information, and dynamic nature of new data mean that new methods, systems, and processes are required to enable quality and robust calculation of price indices. More details on our proposed methods for the rail fares index can be found in our June publications. Impact analysis for this item is presented in Annex B of this letter.

Further analysis suggests that, had the new rail fares indices been implemented in RPI between January 2019 and August 2022, the all-items annual rate would only have been affected in May 2020, reducing the rate by 0.1 percentage points to 1%. This analysis should be treated as indicative; for example, the imputations made for unavailable items over the COVID-19 pandemic have not been recreated based on the revised rail fares data.

x. Changes to the measurement of second-hand cars price indices

The following section describes changes we aim to make to the second-hand cars index in CPIH, CPI and RPI. We have yet to publish our full impact analysis and some internal decision-making processes have yet to be completed. We will only proceed with these changes if we are satisfied that we have robust data sources, methods, systems and processes in place. We will publish our final decision in January next year.

The current methodology and data for our second-hand cars index is described more fully in Section 9.5.3 of our Consumer Price Indices Technical Manual. This involves drawing a sample of prices for 1, 2 and 3 year old cars from a trade guide and making a quality adjustment to account for the age of the car.

Our new second-hand cars data are provided by Auto Trader, currently the largest and most visited vehicle advertising website in the UK. As well as the price they include extensive metadata on each listed car, and have provided us with data dating back to January 2018. As these data are for advertised listings, they do not contain explicit sales revenue information.

As with the new rail fares index, the size, information, and dynamic nature of these new data mean that new methods, systems, and processes are required to enable quality and robust calculation of price indices. More details on our proposed methods for the used cars price index can be found in our June publications. Impact analysis for this item is presented in Annex B.

Further analysis suggests that, had the new second-hand cars indices been introduced to RPI between January 2019 and August 2022, the all-items rate would have been reduced by 0.1 percentage points in 7 months, and increased by 0.1 percentage points in 11 months with a further larger increase of 0.2 percentage points in July 2022. The average impact over the period was to increase the annual rate by 0.011 percentage points. As with the rail fares analysis, these analysis are indicative as, for example, the imputations for unavailable items have not been recreated based on the new indices.

Moreover, our analysis suggests that, had both the changes outlined in viii and ix been implemented together in January 2019, the impact on the annual rate of all-items RPI inflation would have been to increase the annual rate on average by 0.009 percentage points between the implementation date and August 2022. Overall, 19 months were impacted, with the annual rate 0.1 percentage points lower in 8 months, 0.1 percentage points higher in a further 10 months, and 0.2 percentage points higher in July 2022. In terms of cumulative growth, by August 2022 the changes increased RPI from a published index of 345.2 to 345.3 (January 1987=100), with an average impact of 0.03 index points over the period.

Ongoing quality improvements and developments in consumer price statistics

Quality improvement continues to be integral to our work in Prices Division and across ONS. We continue to maintain high standards through external accreditation of our suite of inflation statistics and housing index (ISO 9001:2015) and have also recently assessed our outputs using Statistical Quality Maturity Models (SQMMs) achieving one of the highest scores across all ONS statistics assessed. We have a quality improvement plan in place and are making some substantial improvements to the way that we process, calculate and quality assure our statistics over the coming years, our highest priority deliverables are outlined in this section.

xi. Quality improvements in the Living Costs and Food survey (LCF)

The Living Costs and Food survey (LCF) data underly the construction of weights in the RPI and are also used to inform weights at the lowest levels of the CPIH and CPI hierarchies (where HHFCE data are not granular enough to provide this information).

In July 2021, the Office for Statistics Regulation (OSR) published a report assessing the Living Cost and Food Survey (LCF) for compliance with the Code of Practice for Statistics. This report identified two requirements and three considerations for improvements to be made to ensure that the LCF and its outputs continue to be fit for purpose. A progress report in response to OSRs assessment was published in March 2022 detailing our ongoing work and actions.

This work will ultimately improve the quality of our suite of consumer price statistics and reduce the risk of errors in the calculation of expenditure weights.

xii. Development of the private rental series

We have investigated new methodologies to improve our private rental prices statistics, which feed into the Owner Occupier Housing (OOH) component of the CPIH as well as the private rental components of the RPI, CPIH and CPI. We have extended the timeline in our latest development plan given the importance of these statistics to allow time to appropriately quality assure the new methodology and develop a sustainable system on a strategic platform to support their ongoing production.

xiii. Transformation of consumer price statistics using new data and methods

We continue our endeavour to obtain point-of-sale transaction (scanner) data feeds from UK retailers and have made substantial progress in accessing both historic and regular feeds of these data, currently covering up to 50% of the grocery sector. These data will be aggregated with traditionally collected data in calculating price indices for food and non-alcoholic beverages, alcohol, and tobacco. We plan to continue to improve our coverage of this sector using scanner data in the coming months and years.

As part of our wider transformation, we will begin incorporating alternative data sources into the CPIH and CPI from 2023, through the sources described in viii and ix.

We are also investigating the use of web-scraped data collection for some other sectors, but due to complexities of methodologies needed for use with these data they will not be incorporated into our headline statistics until 2024 at the earliest. Experimental indices will continue to be published as part of our research series in the interim.

xiv. Improvements to the scrutiny process

The current scrutiny process is an important part of our production as it prevents incorrect prices from impacting the published RPI, CPIH and CPI figures. Due to the human nature of many of the errors that occur, and the need to make decisions that don't always have a clear outcome, this cannot be easily automated. An intensive manual process is therefore completed monthly to investigate any observations deemed as outliers, to either validate, correct, or remove the associated prices and/or metadata.

As detailed in our letter of 1 December 2021, this scrutiny process was reviewed over the course of 2021, and we concluded that the manual process was still of fundamental importance, but some improvements could be made. We are currently finalising these improvements and working to embed this process into production. Therefore, as per our previous correspondence, in November 2022 we plan to:

a) move the scrutiny onto a new system with an improved user interface and functionality, including improved reporting of changes

b) bring some of the process forward so that it occurs earlier in the production round, allowing time for additional checking
c) improve training and supporting information provided to price collectors, particularly regarding accurate assigning of indicator codes (used to provide additional information as to whether a product is on sale or a replacement, for example)

These changes will allow us to streamline the process and improve traceability of any changes made, ultimately providing additional reassurances over the quality of price quotes and quality adjustments being made in our suite of consumer price statistics.

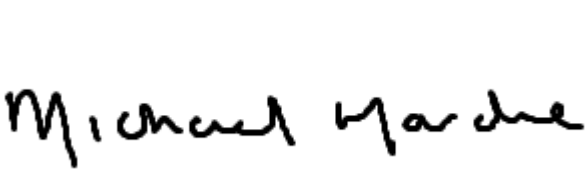
xv. Barcode pilot collection

Since July 2021 our field collectors have been collecting barcodes alongside prices for a small sample of retailers and items, with the aim of providing an effective means of matching products collected in the field with those provided to us through scanner data. Matching products between the field collection and scanner data can serve as an important cross-checking tool in the short-term. The barcode collection will become part of the live CPI local collection from January 2023.

Please let me know if you have any queries or would like to discuss any of the changes raised in this letter further.

A copy of this letter goes to Simon Kirby, Carleton Webb and Simon Dolan at the Bank of England, to Thomas Yeomans at the Treasury and to Sam Beckett, Mike Keoghan, Grant Fitzner, Jason Zawadzki and Chris Payne here at ONS.

Yours sincerely,



Michael Hardie
Matthew Corder
Head of Prices Division, Office for National Statistics

Annex A – Impact analysis for changes to foreign holidays items

Data and method

The proposed changes will not alter the existing methodology, but rather change how far in advance prices are collected. However, the best data we have for conducting an impact analysis, a pilot collection with a 3-month lead time that ran between 2016 and mid 2022, was collected based on a methodology which allowed quality changes in the holidays being priced within a year. Therefore, we needed to decide how to repurpose this data so it could give a reasonable impact analysis.

As part of the existing methodology[1], we compare prices of holidays taking place in a given month to the price for an identical holiday taking place a year ago. However, applying this level of strictness to the pilot data results in too few data points to conduct a meaningful comparison. Therefore, for the purpose of this analysis, we tolerated changes in room type, airline quality and departure day. This provided us with sufficient data points for hotels and self-catering (cruises lacks sufficient data points for a meaningful analysis), but as a result, we expect these tolerated quality changes to generate more extreme price relatives than we would observe if the change were implemented. Furthermore, even with this tolerance, some countries in some months had very few price relatives, leading to volatile indices. This is sometimes compounded by small sample sizes in some months for the published versions of these three items, which the reworked pilot data is compared to.

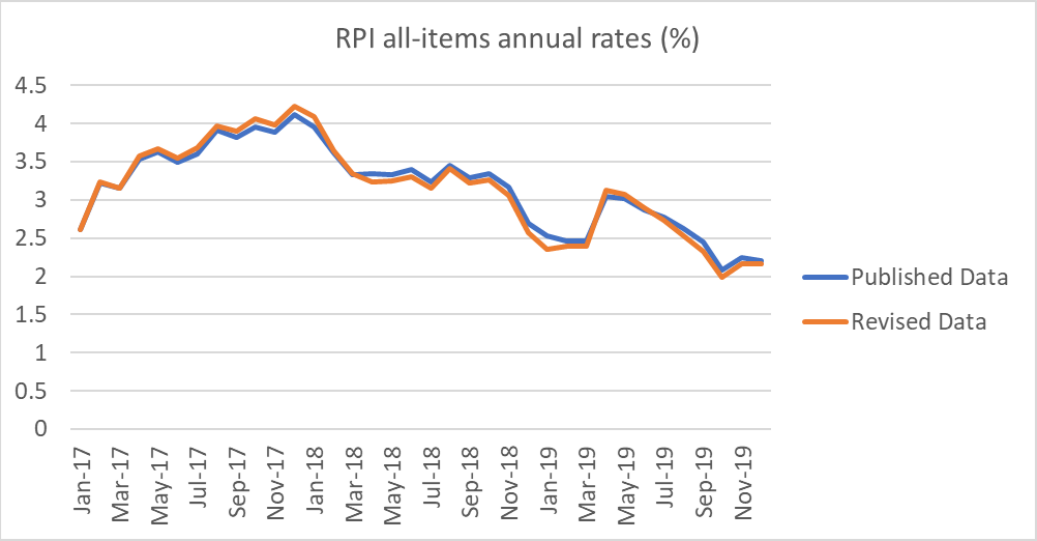
This volatility can be amplified by the agreed foreign holidays methodology, as an extreme price relative in one month will push that month's index up (assuming the relative shows a price increase), and then all subsequent months in that year will use that month's index as part of their rolling-year average, so will be higher as well.

Results

This analysis uses the years 2017 to 2019, due to pandemic restrictions disrupting meaningful analysis for 2020 onwards, and focuses on the RPI, which is the measure of inflation in which these foreign holiday items have the largest weight in pre-covid years (in 2019 the three items had a combined weight of 3.4% in RPI, 2.9% in CPI and 2.3% in CPIH). As noted above, this analysis only considers the effects from hotels and self-catering, due to insufficient data for cruises, however cruises has a weight less than half that of either hotels or self-catering.

[1] See section 9.5.7.1 of our Consumer Price Indices Technical Manual.

Figure 1: impact analysis effect on RPI all-items annual inflation rate (to 1 d.p.)



In 19 out of the 36 months shown above, there were effects on the RPI annual rate when rounded to 1 d.p., with 9 months seeing a +0.1pp impact and 10 months seeing a -0.1pp impact. The upward or downward tendencies within years are likely due to the interaction between data limitations and the existing methodology, as mentioned previously. For example, an extremely high price relative in April of one year can contribute to an unusually high monthly index in April, as well as in May, June, and subsequent months.

The average monthly effect on the unrounded RPI annual rate is -0.016 percentage points over those 36 months, while the average monthly effect on the RPI index level (Jan 1987=100) is +0.004 index points.

As for cumulative effect, in December 2019, the RPI index level was 291.9 in the published data, compared to 291.7 in the revised data, a decrease of 0.2 index points.

Interpretation

The volatility observed in this impact analysis seems likely to stem, in a large part, from the data limitations. Investigations into specific months of largest difference suggest they are mainly due to small numbers of quotes for one or more countries in the pilot collection, or the main live collection, or both.

A further reason to regard this impact analysis as an upper bound is that the pilot collection was conducted on a 3-month lead time, while the proposed new lead time is 6-months, and 6-months on average will be closer to the sort of lead times seen in the traditional collection process – which range from a few weeks to 17 months.

Annex B – Impact analysis for changes to second-hand cars and rail fares items

Background

We are currently undertaking an ambitious programme of transformation across our consumer price statistics, including identifying new data sources, improving methods, developing systems and establishing new processes. The work has been in progress for several years, reflecting the complexity and intricacy of what we are striving to achieve.

Our latest plans detail how updates will be made annually as part of a continuous improvement cycle, and that our first introduction, in 2023, should prioritise the inclusion of new data for rail fares and second-hand cars. These new data will offer a substantial improvement in the coverage and granularity of our published indices.

The second-hand cars data are provided by Auto Trader, currently the largest and most visited vehicle advertising website in the UK. As well as the price, they include extensive metadata on each listed car, with a database dating back to January 2018. As these data are for advertised listings, they do not contain explicit sales revenue information.

The rail fares data for GB are provided to us by the Rail Delivery Group and are sourced from the rail industry's Latest Earnings Networked Nationally Overnight (LENNON) ticket revenue system; this system does not retail (as in, it does not act as a ticket selling platform). As they are transaction level data, they contain information on the cost and quantities of purchased tickets along with associated metadata related to the journey travelled. We have been provided rail transaction data for GB dating back to January 2019. Northern Ireland data will continue to be provided separately and will become a separately published index in future consumer prices publications.

The size, information, and dynamic nature of new data mean that new methods, systems, and processes are required to enable quality and robust calculation of price indices. More details on our proposed methods for these categories can be found in our June publications. This article updates the indices shown in these publications to include data up until August 2022 and shows the combined impact of introducing these methods and data into our headline consumer price statistics. We also include details of some key improvements that we have made to our methods and to how the indices are structured.

The estimates in the present analysis are indicative, there are still some minor changes being made to the systematisation of the indices, though these are expected to have a negligible impact on the figures presented. There are three changes still to be implemented:

I. transitioning the system into a monthly production state
II. data cleaning
III. filtering and imputation procedures for multilateral index number methods

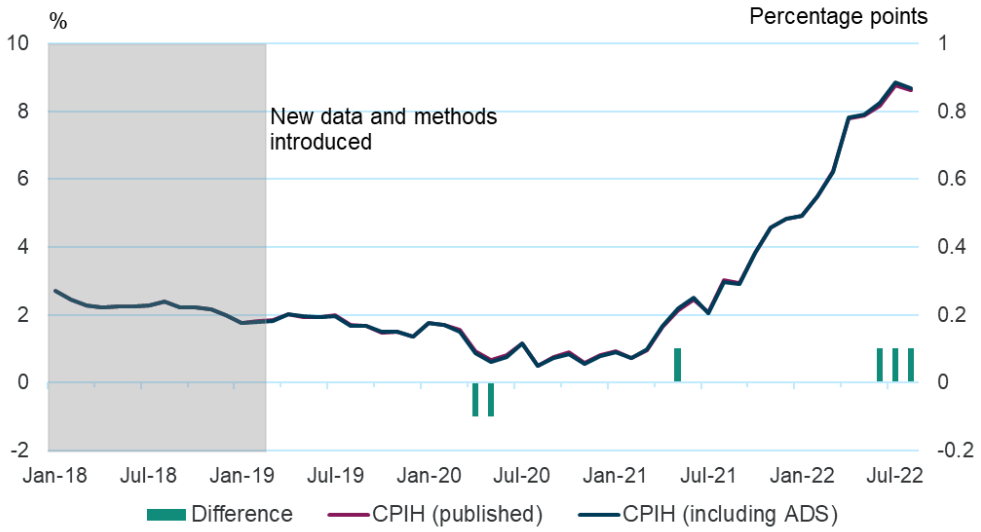
Indicative impact of ADS on annual CPIH and CPI inflation rate

As GB rail fares data are only available from 2019, to show the impact of introducing the rail and second-hand cars data simultaneously, as we plan to in 2023, we have chosen to also include the second-hand car data from 2019 (despite having it from 2018). This revised series from 2019 is chain-linked to the published series pre-2019 to show the indicative impact on the expected growth rate in the year of introduction. While our present analysis covers the period to August 2022, for the November publication we will only include impact analysis to June 2022, as previously agreed with stakeholders.

The existing, published weights for rail fares and second-hand cars are used to aggregate the new data and methods into our published consumer price statistics. The weight for rail fares in CPIH is small and has fallen over the last 4 years, ranging from 0.8% of CPIH in 2019 to just 0.3% of CPIH in 2022. The weight for second-hand cars is slightly larger; 1.4% of CPIH in 2019, 1.2% of CPIH in 2020-2021 and 2.0% of CPIH in 2022.

Figure 1 shows the impact of using these new data and methods to calculate CPIH between January 2018 and August 2022, had they been introduced from February 2019. Over this period the average, unrounded difference has been neutral (0 percentage points), though when rounded a small contribution of 0.1 percentage points can be seen in six individual months following introduction. The largest unrounded upward difference is 0.077 in July 2022, the largest unrounded downward difference is 0.076 in May 2020.

Figure 1: Published CPIH and CPIH including ADS 12-month rates (%) including differences (percentage points), UK, January 2018 to August 2022

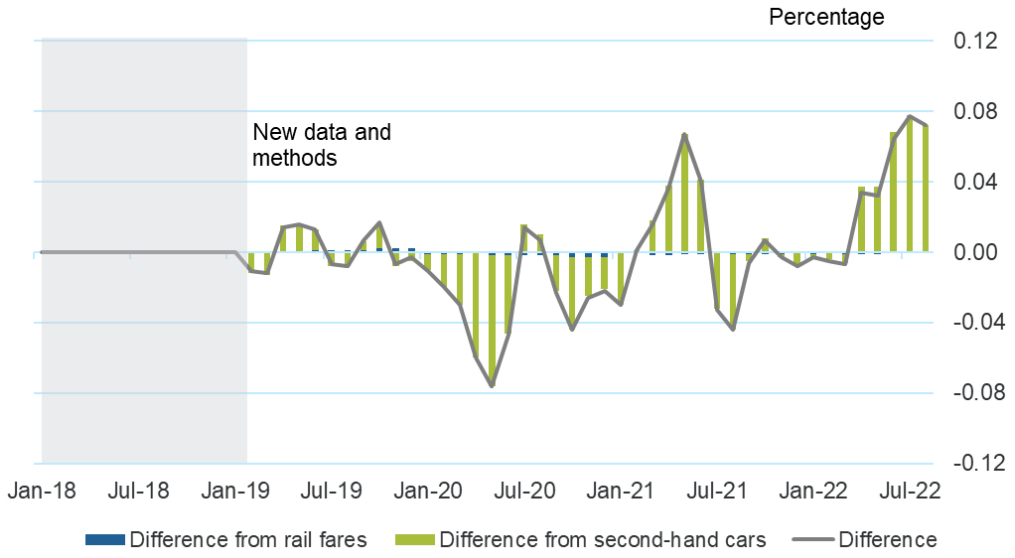


For CPI, there were 10 instances over the same period, but these again resulted in a neutral average difference on the index.

Through 2020 and 2021 there were several unavailable items due to the coronavirus (COVID-19) pandemic that had to be imputed in some periods. For this impact analysis, we haven't recalculated these imputations for the revised series, but we would expect the impact of these to be small (as the headline rate was used to impute missing items during this period, and there is little change to the headline rate using these new data).

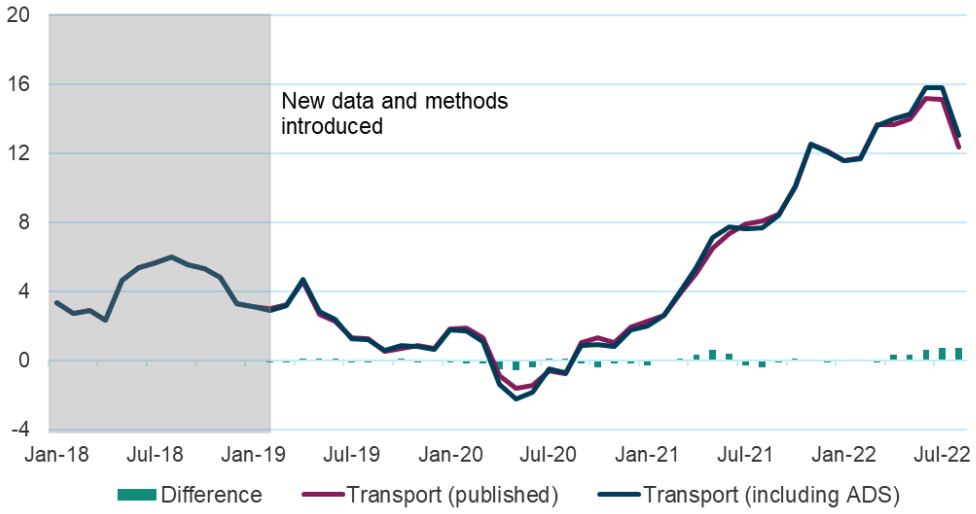
The primary contributor to the difference between the published CPIH (and CPI) 12-month inflation rate and the comparable inflation rate with the revised indices came from second-hand cars, as seen in Figure 2.

Figure 2: Contributions from revised second-hand car and rail fares indices to the difference in the CPIH 12-month rate (percentage points), UK, January 2018 to August 2022



Differences in the index for the transport division of the CPIH, of which second-hand cars and rail fares account for between 14% and 18%, are intuitively larger, ranging between -0.6 and 0.7ppts. These differences are shown in Figure 3.

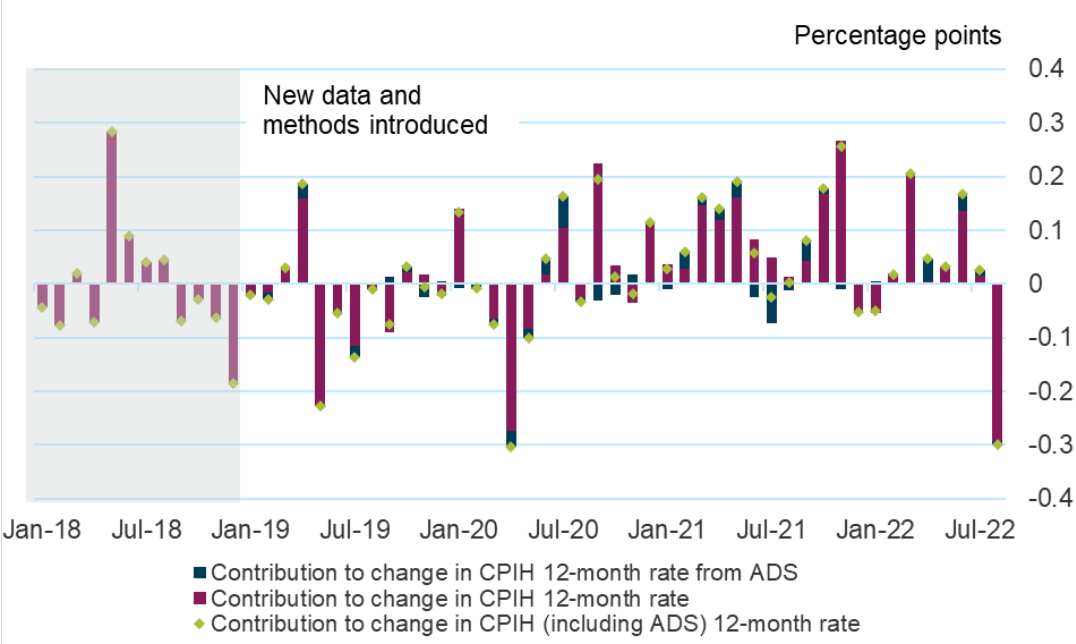
Figure 3: Published transport index and transport including ADS index, 12-month rates (%) and differences (percentage points), UK, January 2018 to August 2022



The transport division contains consumption categories including the purchase of vehicles (both new and second-hand cars), operation of personal transport equipment (including motor fuels) and transport services (including air fares). This means that the transport division often provides a notable contribution to the change in the 12-month inflation rate.

Figure 4 shows how much the transport division contributes to the change in the CPIH 12-month rate (purple), to the change in the CPIH (including ADS) 12-month rate (green) and by how much the contribution changes following the introduction of new data and methods for rail fares and second-hand cars (navy).

Figure 4: Transport division contributions to change in the CPIH 12-month annual rate, including impact of ADS, UK, January 2018 to August 2022



For the majority of periods, the contribution from the transport division to the change in the headline CPIH 12-month rate is either maintained or marginally amplified (either positively or negatively) following the introduction of ADS for rail fares and second-hand cars. There is one period, in July 2021, where the transport division previously had a positive contribution to the change in the headline rate but following the introduction of ADS for rail and cars had a slightly negative contribution.

Improvements made since the June 2022 method publications

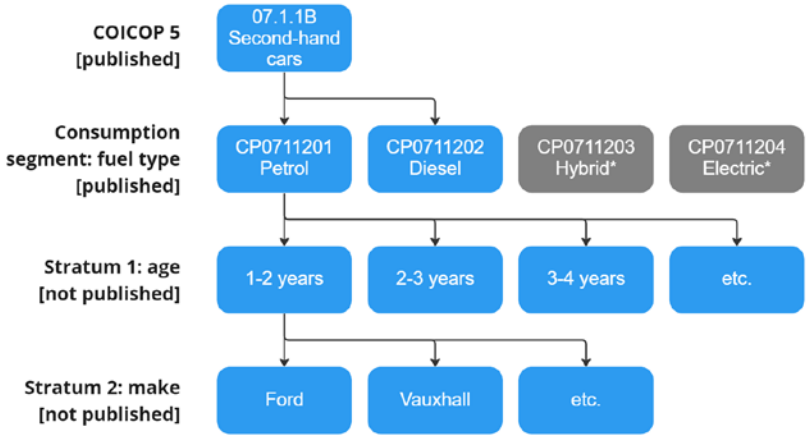
While the impact of these new data and methods on headline consumer price inflation measures is small, there are many improvements arising from use of the new data sources, such as improved coverage of prices through the month, of different car makes and ages, and of rail ticket types. We can also provide a greater level of detail in the published indices.

Since our methods publications in June 2022, we have made several further improvements to our index calculations, the most notable of which is in the ordering of the hierarchies for both rail and cars. Both hierarchies have now been restructured based on user feedback, such that users will be able to explore published rail fare indices by ticket type and published second-hand car indices by fuel type. For a general overview of our aggregation strategy, see Introducing alternative data into consumer price statistics: aggregation and weights.

Second-hand cars

For second-hand cars, we can begin to publish the inflation rate of cars with different fuel types, instead of the two existing items for 2-year-old and 3-year-old cars. The new, detailed hierarchy for second-hand cars is provided in Figure 5.

Figure 5: New, detailed hierarchy for second-hand cars



*Note that, while we have data for second-hand hybrid and electric cars, the number of cars in the sample is still small as is the weight for these in the second-hand cars market. These consumption segments will be added as part of the annual basket update once they gain a significant proportion of the second-hand car market.

Indices for petrol and diesel using new data and methods are shown in comparison to the currently published COICOP5 index for second-hand cars in Figure 6, with the annual growth rate for these indices shown in Figure 7.

While the indices show broadly similar trends to the published series, we can now see that since January 2019 the published series has been more in line with the inflation rate of petrol cars. This has not been the case since March 2022, where we have seen a sharp downward trend in the published series. While the annual growth rate for the revised series has also slowed, the decline has not been as sharp. There are many potential reasons for differences between the published and revised series; in particular the revised series have substantially increased coverage (both temporal and market coverage), meaning that the price indices are less sensitive to sudden price movements of individual makes or models of car.

Figure 6: Petrol and diesel second-hand car indices compared to the currently published COICOP5 index for second-hand cars, UK, Jan 2019 = 100

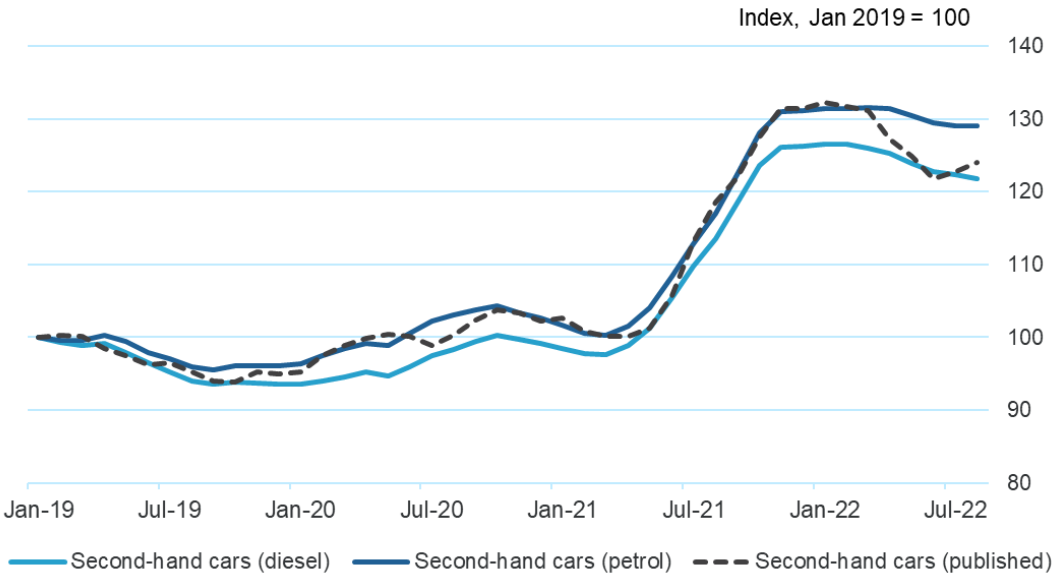
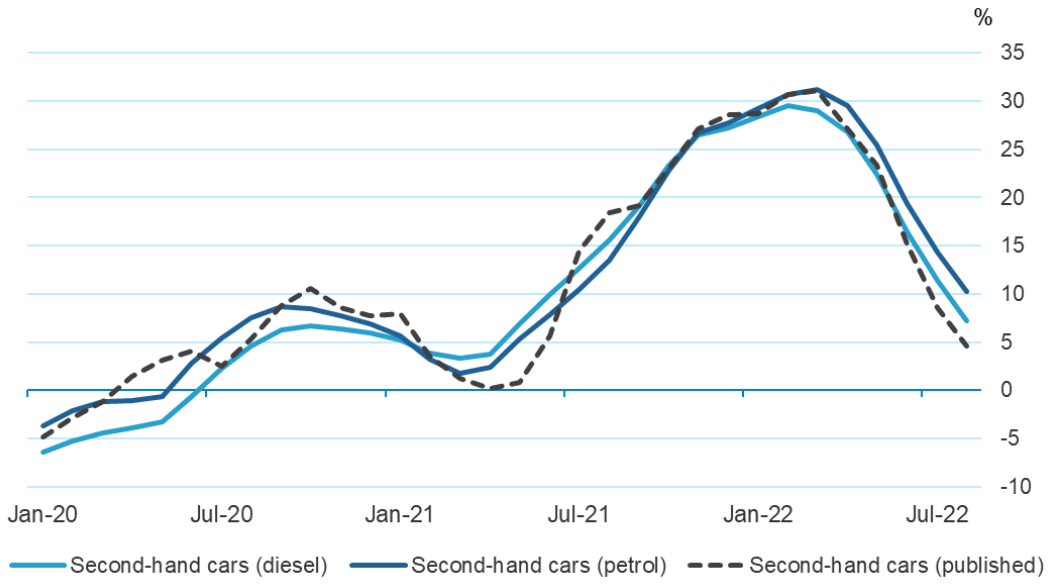


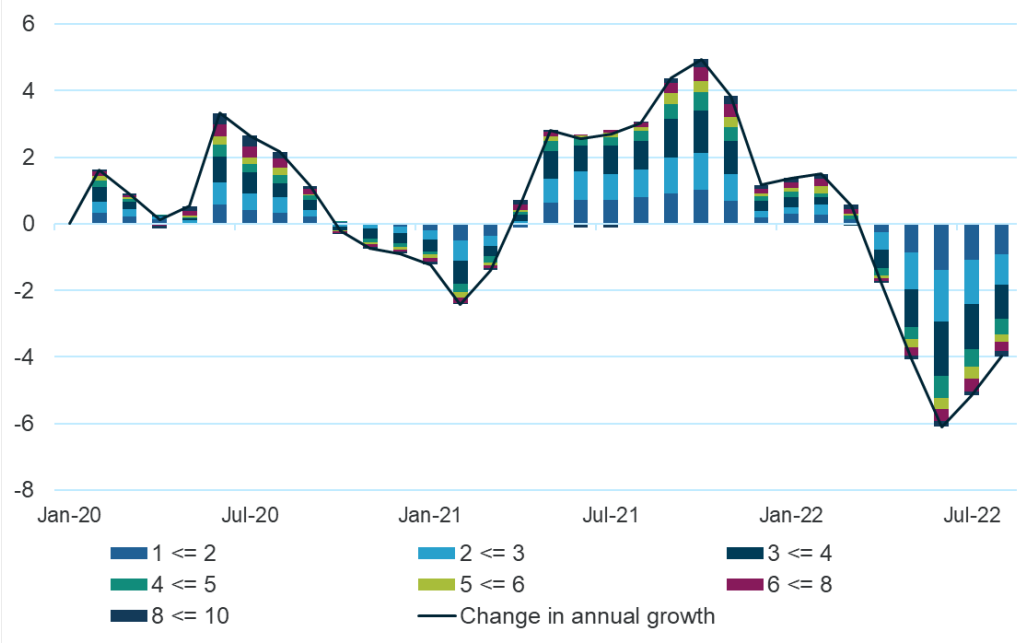
Figure 7: Petrol and diesel second-hand car indices compared to the currently published COICOP5 rate for second-hand cars, UK, Annual % change



The new petrol and diesel second-hand car indices are stratified by age (between 1 and 10 years) and make (up to ~25 makes per fuel type). We will therefore be able to provide users with more detailed information (through the monthly detailed briefing note) as to, for example, the age of second-hand cars that are contributing to the change in inflation, although these stratum level indices will not be published as part of the regular inflation tables.

Contributions from different ages of petrol cars to the change in the 12-month growth rate can be seen in Figure 8. The largest contributions come from newer cars, with cars under 4 years of age accounting for over 50% of the weight of the petrol second-hand cars index. It can be seen that all ages of cars typically contribute to the change in petrol second-hand car inflation in the same direction over the period analysed, suggesting that market dynamics are similar independently of the age of car purchased. The differing magnitude of the contributions is primarily driven by the difference in weights for each age strata.

Figure 8: Contributions (ppts) to the change in the annual growth rate for petrol second-hand cars, by age group



The differences between the indices we see using our new data and methods were largely explained in our previously published article, but some additional improvements have been made since this article was published.

Firstly, as part of the hierarchy restructuring, we reselected the most representative car makes for each fuel type. All car makes with over 1% market share within the Auto Trader data are now selected as considered representative of each fuel type, resulting in approximately 90% of the approximated expenditure being captured across all included age strata (1-10 years).

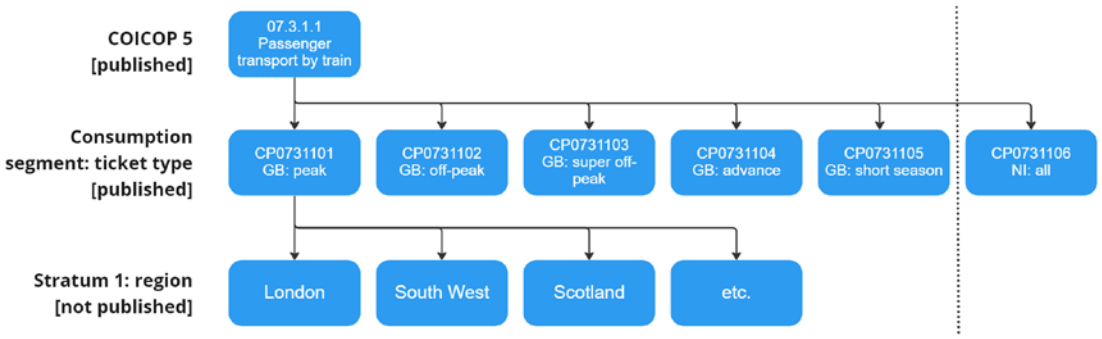
We have also improved our method of calculating weights for second-hand car indices, from using quantities of cars present in the data, to using approximates of expenditure based on the quantities of car types assumed sold and their average prices. While this change had a small impact on the resulting indices, using expenditure to aggregate indices aligns with international best practice.

Lastly, we have removed cars that are listed as belonging to write-off categories. This change is important in making sure only comparable cars are priced over time, however given the small proportion of the data that belong to these write-off categories (~1.5%), the impact of this improvement has been negligible.

Rail fares

For rail fares we will begin to publish six new indices based primarily on ticket-type, instead of the previous single item index. The new, detailed hierarchy for rail fares is provided in Figure 9. The new data sources will feed into the GB consumption segments (CP0731101 to CP0731105), the index for Northern Ireland (CP0731106) will continue to be produced using data from our current provider.

Figure 9: New, detailed hierarchy for rail fares



*Note that, while we have data for long season tickets, more research is needed as to how we calculate the consumption price of a longer season ticket. If we are able to calculate these effectively in future, they will get added as part of a future basket update as a new consumption segment.

Consumption segment indices for rail fares are shown in comparison to the published index in Figure 10, with annual growth rates for these indices shown in Figure 11. The growth rate is lower for most regulated tickets but offset by a higher growth rate for advance tickets.

Figure 10: Consumption segment rail fares indices and the currently published rail fares index, UK, Jan 2019 = 100

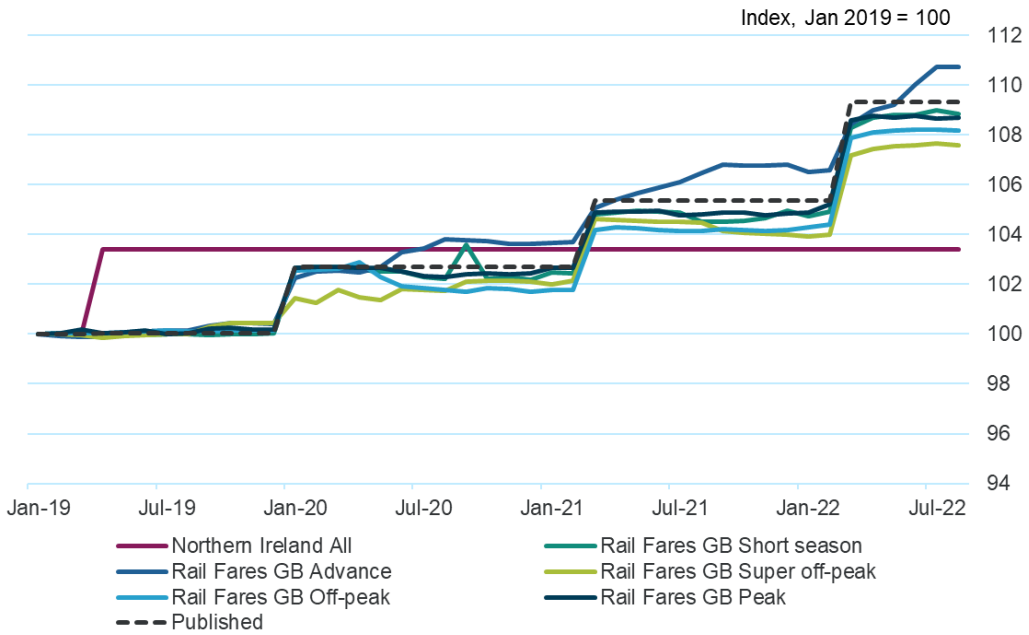
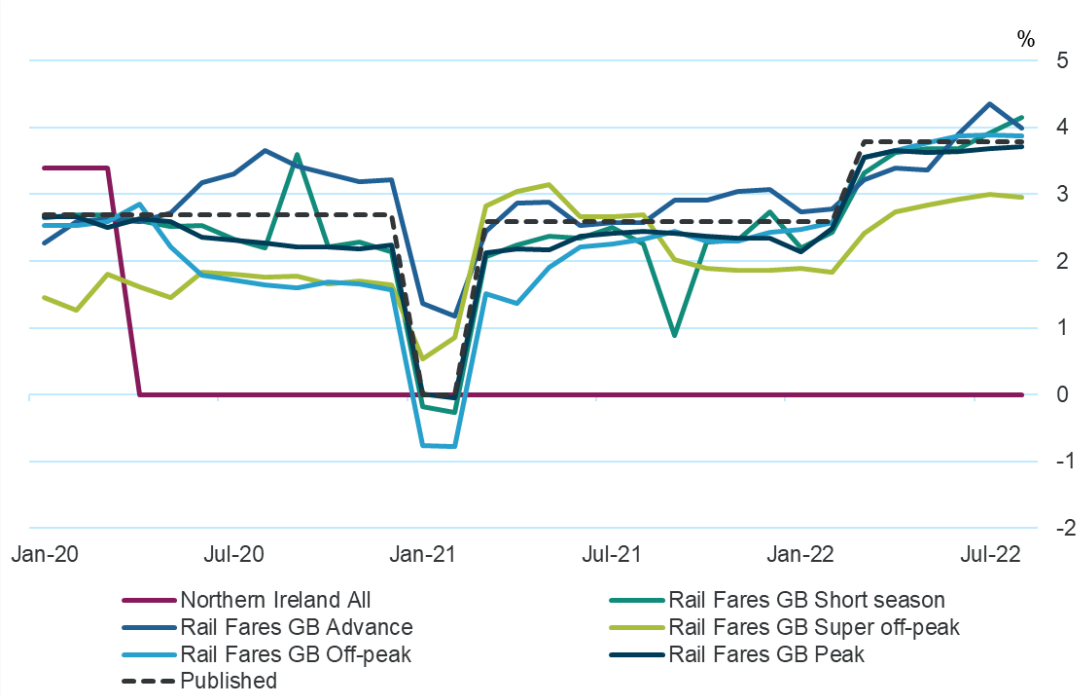


Figure 11: Consumption segment rail fares indices and the currently published rail fares index, UK, Annual % change

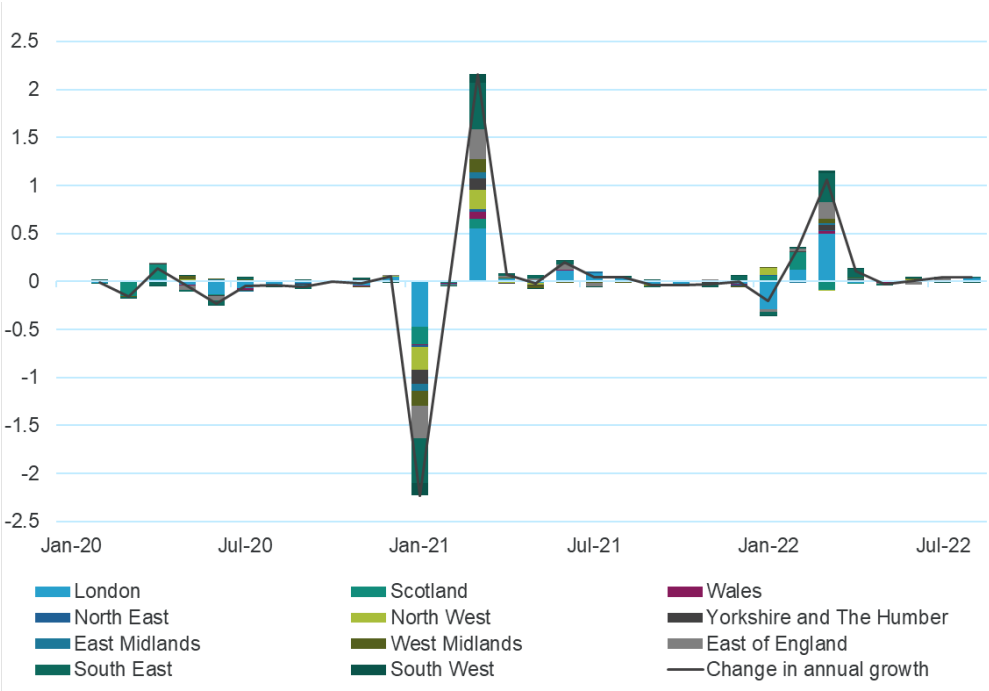


All indices (except Northern Ireland who have not had a price change since April 2019), see a decline in annual growth in January 2021 that quickly reversed in March, this was due to the change in timing of the annual fares uplift from January to March. Season tickets see a spike in September 2020 and the base effect of this spike the following year. This is due to an increase in prices on a single day in September 2020 in London, we are still investigating whether this increase is genuine or an error and how these sorts of observations should be handled – this is discussed further in our readiness assessment.

The GB rail indices will be based on regional aggregates, so we will also be able to provide users with detail (through the detailed briefing note) as to what regions are contributing to the change in inflation for each fare product group, although the regional indices themselves will not be published. These regional indices will also help us to produce regional CPIH/CPI more readily in future.

Figure 12 shows the contribution from each region to the change in annual growth (%) for GB rail: peak tickets. In most months, there are few changes coming from any region. It can be seen that the change in the application of the annual uplift from January 2021 to March 2021 was applied by all regions (the difference in regional contributions is primarily driven by the difference in the respective weights for each region). While most regions continued with a slightly larger uplift in March 2022, this was marginally offset by Scotland (and the North West region to an even lesser extent).

Figure 12: Contribution from each region to the change in annual growth (%) for GB rail: peak tickets, UK, January 2020 – August 2022



Again, while many of the differences between the published and new measures are discussed in our previous methods publication, we have made some further improvements to our calculation of the indices since.

Our previously published indices showed a substantial fall of 0.6% in the rail fares index in October 2021 that was attributed to the London region, showing a drop of 2.2% in peak tickets and a 4.0% drop in off-peak tickets. This was found to be due to reclassification of a particular ticket type between a contactless pay ticket and a pay as you go ticket, rather than being a genuine price decrease. We were able to identify these reclassified tickets based on their issuing business and have now combined these ticket types, eliminating the previously seen fall in the index in this month.

Through this analysis, we also discovered that we were capturing “un-started” and unfinished tickets; where a passenger would have started a journey by swiping through a ticket gate but not swiped out, or vice versa. In these cases, the passenger is charged a max fare for the journey. As we consider this journey cap a distinct product from the rail fare, we have excluded any rows where the origin and destination stations are the same, since we cannot confirm that a journey has taken place.

We have also completed further research into how to treat refunded rail fares. When a transaction is refunded, in the data a new separate "refund" transaction is created that either partially or fully offsets the original value. However, linking the refund transactions to their original counterparts is complex; based upon the information given to us by the data supplier we were either finding too many matches (duplicating the transactions three times over) or too few matches (a match rate of less than 0.02%). Subsequently, the supplier has conveyed to us the difficulty with refund linking and believes that we are unlikely to achieve a successful match rate.

Since we do not expect that the refunded transactions have incorrect prices, just that the item has not been consumed, we believe that refunds would have a small impact on the rail fares indices. Therefore, we would expect the main impact of refunds to be on the weights used to aggregate indices for different ticket types and regions together, especially if there's a greater prevalence of refunds in a certain region or ticket type. The refund transactions account for 1.9% of the total row count and 4.8% of expenditure, so the impact of these transactions on either the price indices or the weights is expected to be small. We have therefore not pursued linking refunds to their original transactions when producing the indices.