Estimating RPI-adjusted equity market returns

Prepared for
Heathrow Airport Limited (HAL)

2 August 2019

www.oxera.com

Contents

1	**Introduction**	**1**
1.1	**Background**	**1**
1.2	**Structure of report**	**2**
2	**How does RPI fit into price regulation?**	**3**
2.1	**The role of inflation in regulation**	**3**
2.2	**Allowed return and inflation**	**4**
3	**The statistical history of the RPI**	**10**
3.1	**The history of the RPI**	**10**
3.2	**Assessing the impact of methodological changes on inflation measurement**	**16**
3.3	**What major changes have there been to the RPI calculation?**	**17**
4	**The RPI today**	**21**
4.1	**Elementary aggregates**	**22**
4.2	**Indices**	**23**
4.3	**Weights**	**23**
5	**What are the desirable properties of a good inflation measure for regulation?**	**25**
5.1	**Desirable properties of indices from a statistical perspective**	**25**
5.2	**Desirable properties from the perspective of economic regulation**	**26**
6	**Potential alternative inflation measures**	**28**
6.1	**Differences between different data sources**	**28**
6.2	**Adjusting the RPI**	**29**
6.3	**RPIJ and CPIH**	**33**
7	**Summary and recommendations**	**36**
A1	**Elementary aggregate formula**	**38**
A1.1	**Dutot**	**38**
A1.2	**Carli**	**38**
A1.3	**Jevons**	**39**

A1.4 A numerical example 39
A1.5 Relationship between elementary aggregates 40

A2 Step-indicator saturation 41

A3 Structural break tests 43

A3.1 Structural breaks of an extended price series 43
A3.2 Logs 43
A3.3 Breakpoints 44
A3.4 Controlling for the persistence in inflation 44
A3.5 Structural breaks analysis for Method 2 adjustment 46

Boxes, figures and tables

Box 3.1 Differences between the CPI and the RPI 13

Figure 3.1 Summary of the early history of the RPI 11
Figure 3.2 Summary of the methodological improvements in RPI from 1956–1987 14
Figure 3.3 Summary of the methodological improvements in RPI from 1993–present 15
Figure 3.4 Long-run evolution of the RPI series 17
Figure 3.5 Structural breaks in the RPI series 18
Figure 3.6 Structural breaks in the RPI series, in logarithms 19
Figure 4.1 Stylised representation of the RPI calculation 21
Figure 4.2 Contribution of the formula effect to the wedge between the RPI and the CPI 23
Figure 6.1 Evolution of data sources in different inflation measures 28
Figure 6.2 Difference between the RPI published by the ONS and other inflation measures 29
Figure 6.3 Backcasting of the weights on foreign holidays 30
Figure 6.4 Backcasting the price on foreign holidays 31
Figure 6.5 Stylised example on the use of structural breaks test in backcasting an adjusted RPI 32
Figure 6.7 RPI vs RPIJ 34

Figure A2.1 Oxera’s hypothetical time-series variable 41
Figure A3.1 Structural breaks of an extended RPI series 43
Figure A3.2 Partial autocorrelation function (PACF) of the RPI series 44
Figure A3.3 Structural breaks analysis with auto-regressive term, in levels 45
Figure A3.4 Structural breaks analysis with auto-regressive term, in logs 45

Table 2.1 Real historical returns for the UK equity markets, UKRN analysis (%) 7
Table 2.2 Real historical returns for the UK equity markets (arithmetic nominal average) 8

Table 3.1	Effect of identified breaks on RPI series	20
Table 4.1	Formulae used in RPI and CPI (2012 weights)	22
Table 4.2	Changes in the RPI basket over time, parts per 1,000	24
Table 5.1	Summary of features of a ‘good’ price index	26
Table 6.1	Inflation rates 1899–2016	31
Table 6.2	Adjustment of RPI by Method 2 from 1899 to 2016	32
Table 6.3	Adjusted real historical returns for the UK equity markets (%)	33
Table A1.1	Effect of price dispersion on when calculating indices using different formulae	39
Table A3.1	Sensitivity, structural breaks in log RPI series	43
Table A3.2	Structural breaks in RPI series from 1900 (in RPI points)	46

1 Introduction

Heathrow Airport Limited ('HAL') has commissioned Oxera to consider the properties of the Retail Prices Index ('RPI'), how it has evolved over time, and how its use could be improved for economic regulation.

1.1 Background

When setting price controls for regulated entities, regulators have to control for inflation in several parts of the calculation. Different regulators take different approaches and the approach chosen can have a material impact on both consumers and shareholders.

In the price control for HAL the regulatory asset base ('RAB') is indexed to the RPI. Therefore, the cost of capital assumption should exclude the level of RPI inflation that is **expected** over the period of the next price control. Similarly, the assumption for the expected total equity market return should also be expressed net of expected RPI inflation.

The historical return on the UK equity market is an important input into the calculation of the cost of equity when setting a price control. The 2018 Global Investment Returns Yearbook ('2018 GIRY') reports an arithmetic average annual return for the UK equity market of 11.2% (nominal) for the period 1899–2017.[1] The nominal return in the 2018 GIRY can be expressed as a real return by deflating the times series using a measure of inflation. Until recently, the GIRY deflated returns using the historical RPI index. However, the 2018 GIRY deflates returns using a hybrid index that combines historical data on both the RPI and the Consumer Price Index ('CPI'), illustrating the tension around the selection of an appropriate inflation measure.

In parallel, the methodology used to estimate inflation has evolved over time. In 2010, for instance, the Office for National Statistics ('ONS') updated its methodology for calculating RPI, which increased the divergence from the CPI measurement.[2] The ONS also developed a refined version of the CPI, the CPIH,[3] which became the lead inflation index in the UK.[4] In January 2015, the UK Statistics Authority published an independent review of consumer price statistics, led by Paul Johnson, Director of the Institute for Fiscal Studies.[5] The review supported previous findings that the approach used to calculate the RPI was failing to meet international standards.[6]

All these changes affect the estimation of real annual equity returns. Indeed, a recent study for the UK Regulators Network ('the UKRN study') stated that:

> Changes to the underlying methodology mean that the RPI is not comparable over time, whereas historical CPI estimates try to match current methodology. Historic equity returns deflated by RPI will therefore have limited informational content about future equity returns deflated by RPI.[7]

[1] Dimson, E., Marsh, P. and Staunton, M. (2019), 'Credit Suisse Global Investment Returns Yearbook 2018', February.
[2] See, for example, Office for National Statistics (2017), 'Measuring inflation – what's changed and why?', March, https://blog.ons.gov.uk/2017/03/20/measuring-inflation-whats-changed-and-why/, accessed 19 July.
[3] CPIH = 'Consumer Prices Index including Owner Occupiers' Housing Costs'.
[4] Consumer Price Inflation (includes all three indices: CPIH, CPI and RPI).
[5] Johnson, P. (2015), 'UK Consumer Price Statistics: A Review', January.
[6] Office for National Statistics (2012), 'National Statistician's consultation on options for improving the Retail Prices Index', 14 November.
[7] UK Regulators Network (2018), 'Estimating the cost of capital for implementation of price controls by UK Regulators', March, D-109.

The UKRN study reports that the geometric average annual equity return over the period 1899–2016 was 5.23% if deflated by the series labelled as CPI in the Bank of England's Millennium dataset,[8] and 5.07% if deflated by RPI.[9] The UKRN study applies an uplift to the geometric average, concluding on a range of 6–7% for the equity market return assumption expressed relative to CPI inflation. However, the historical CPI series in the Bank of England's Millennium dataset is a 'backcast' series as there is no contemporaneous data for CPI before 1989.

By assuming that the historical CPI is known and the historical average CPI-deflated equity market return was 6–7%, the UKRN study implies that investors will require a nominal return that is based on the assumed 6–7% real market return plus CPI inflation. As the HAL price control is indexed to RPI, the Civil Aviation Authority ('CAA') has proposed to deduct the difference between forecast RPI and CPI inflation (the RPI being approximately 100 basis points ('bps') greater) when setting the allowed cost of equity relative to the 6–7% real market return. The reliability of this approach rests on the accuracy of the 'backcast' CPI series and the premise that the average rate of inflation in the historical sample is the same for the RPI and the CPI.

The historical time series for RPI is longer, with actual data published since 1947 and estimates for the period 1870–1947, based on the 1947 definition of the RPI.[10] Therefore, an alternative approach is to consider what adjustments would be required to the historical RPI data to make the series consistent with the way the RPI is calculated today. In this context, HAL is reflecting on what would be an appropriate way to determine the expected equity market return in real terms relative to RPI.

1.2 Structure of report

The remainder of this report is structured as follows.

- Section 2 discusses the role of inflation measures in price regulation with a focus on the cost of equity. In particular, we explain the role of inflation assumptions in estimating the future cost of equity.
- Section 3 reviews the history of the RPI and the main methodological changes that affected it. We also conduct some statistical analysis to understand the major changes that affected the measurement of the RPI.
- Section 4 examines the way the RPI is calculated today and some methodological issues.
- Section 5 sets out the desirable properties of inflation indices in the context of price regulation.
- Section 6 discusses some potential alternative inflation measures.
- Section 7 summarises and presents recommendations.

[8] Bank of England (2017), 'A millennium of macroeconomic data for the UK', April.
[9] UK Regulators Network (2018), op. cit., Table D4.
[10] Johnson (2015), op. cit., section 3.2.

2 How does RPI fit into price regulation?

In this section, we discuss how inflation measures are used in price regulation. In particular, we explain the role of inflation forecasts in estimating the forward-looking cost of equity.

2.1 The role of inflation in regulation

HAL is subject to economic regulation that determines the maximum charges it is allowed to levy on its customers. Specifically, the maximum charges are determined using the RAB–weighted average cost of capital ('WACC') building block model, which is designed to allow the company to recover efficiently incurred operating expenditure ('OPEX') and capital expenditure ('CAPEX'). The RAB–WACC building block model determines the revenue that Heathrow Airport is allowed to recover from its customers.

The revenue allowance in relation to CAPEX is determined with reference to the RPI-indexed RAB. The RAB captures the current value of investment and increases over time with net new CAPEX (i.e. CAPEX less depreciation) and RPI inflation. The RAB is used to determine the depreciation allowance (or return *of* the RAB) and return allowance (or return *on* the RAB), where the latter is determined by multiplying the value of the RAB by the appropriate estimate of the real cost of capital. Once these two fundamental blocks are determined, the regulator adds a revenue allowance for OPEX to determine the total revenue allowance for Heathrow Airport.[11]

As a result of the current regulatory regime, HAL's investors receive revenues that are linked to RPI. This link may be valued by institutional investors in the context of their overall investment portfolios, particularly as a hedge against liabilities that are linked to RPI. For example, UK pension funds typically have liabilities that are linked to inflation, and often specifically to RPI.[12]

Overall, the choice of inflation index affects HAL through the total allowed revenues derived in the RAB–WACC building block model. We distinguish two potential effects of the inflation on the RAB–WACC building block model for HAL: (i) the choice of different inflation indices; and (ii) the difference between expected and actual inflation for a given inflation index.

2.1.1 Choice of inflation index

Conceptually, the choice of an inflation index (e.g. RPI or CPI) should not affect the expected financial value of Heathrow Airport's cash flows as investors are remunerated for inflation though RAB indexation and the use of an appropriate real cost of equity. This means that the choice of inflation index is expected to be NPV-neutral for the company, i.e. the present value of future cash flows would be the same for investors regardless of the choice of inflation index.

For example, if an inflation index with a lower expected value was used, the value of the RAB would be expected to grow at a lower rate—this would be an NPV-negative change for Heathrow Airport. However, this would be offset by a

[11] The CAA has adopted a single-till approach to airport regulation. Under this approach, the revenue generated from non-regulated activities is subtracted from the total revenue allowance determined using the building block model. The remaining revenue is recovered through regulated airport charges. Civil Aviation Authority (2017), 'Consultation on core elements of the regulatory framework to support capacity expansion at Heathrow', June, para. 2.29.
[12] For example, in the context of HAL, the Universities Superannuation Scheme owns 10% of the airport. See Heathrow Airport's website, https://www.heathrow.com/company/company-news-and-information/company-information, accessed on 15 July 2019.

higher real cost of equity, which would incorporate the lower inflation index—an NPV-positive change for Heathrow Airport. On average, these two effects should cancel each other out such that the net effect of the choice of inflation index is neutral.

However, the choice of inflation index would have an effect on the path of the allowed revenues and charges over time. Specifically, a lower inflation index would imply a higher real cost of capital. In the short term, the higher real cost of capital would lead to higher allowed revenues and higher charges. Over time, the increase in allowed revenues and higher charges would be moderated by a lower growth in the RAB. As outlined above, the net effect of these two drivers should be NPV-neutral for the company.

In practice, some investors may have a preference for a specific measure of inflation. For example, institutional investors, such as pension funds, may have RPI-linked liabilities and, therefore, may value assets that provide a return that also depends on the level of RPI inflation.

The CAA's current position is to continue using the RPI inflation index for RAB indexation in the next price control for HAL, hence we do not focus our analysis on the choice of inflation index but rather how to best apply RPI in this context.[13]

2.1.2 Differences between expected and actual inflation

In the price control for Heathrow Airport, the RAB is indexed to RPI. Therefore, the cost of capital assumption should exclude the level of RPI inflation that is **expected** over the period of the next price control.

It is important to note that there could be a mismatch between the level of **expected** inflation used to derive the real cost of capital and the **actual** outturn RPI inflation used to index the RAB.

In particular, the expected inflation should reflect an unbiased expectation of the actual RPI inflation to ensure a fair compensation to investors. If expected inflation used in deriving the cost of capital were mis-calibrated and were, for instance, higher than the 'true' level of expected inflation, then the allowance for the real cost of capital would be lower than the 'true' real cost of capital. As a result, investors would not be fully compensated for providing capital to the business. This argument is symmetric—if the expected inflation used in deriving the cost of capital were lower than the 'true' level of expected inflation, then investors would be overcompensated for providing capital to the business.

Section 2.2 outlines the current debate on the approaches for deriving an estimate of the real cost of capital that is consistent with expected inflation.

2.2 Allowed return and inflation

2.2.1 Introduction

As discussed in section 2.1, in the price control for HAL, the RAB is indexed to the RPI. Therefore, the cost of capital assumption should exclude the level of RPI inflation that is expected over the period of the next price control.

In general, the cost of capital is estimated as a weighted average of the cost of equity and the cost of debt, with the weights determined by the relative proportions of equity and debt in the capital structure of the company. While

[13] Civil Aviation Authority (2017),' Economic regulation of capacity expansion at Heathrow: policy update and consultation', CAP 1610, December, para. 3.30.

both the cost of equity and the cost of debt should reflect a consistent measure of inflation, in this section we focus on the impact of inflation on the estimate of the expected total equity market return, as this parameter is not directly observed and therefore has to be estimated.

The historical total market return ('TMR') on the UK equity market is an important input into the calculation of the cost of equity when setting a price control for HAL. In the context of HAL, the cost of equity has to be expressed net of expected RPI inflation.

One of the most widely used sources for estimating the TMR is the GIRY report, which captures historical equity market returns around the world. In particular, the GIRY report includes the historical nominal TMR for the UK equity markets.[14] As HAL's regulatory regime requires a real cost of equity, the nominal return in the GIRY report should be expressed in real RPI-based terms.

In this section, we outline potential methodological considerations in expressing nominal returns in real terms, taking account of the recent UKRN study on this topic.[15]

- **The direct vs indirect approach for deriving an RPI-based TMR**. The UKRN study implies that investors will require a nominal return that is based on the assumed CPI-based real market return of 6–7%. An *indirect* approach to convert this CPI-based real TMR to RPI terms would be to deduct the difference between forecast RPI and CPI inflation. The CAA and Ofgem followed this approach in expressing the UKRN-recommended TMR in RPI terms.[16] As explained below, this approach rests on the premise that the backcast CPI series is sufficiently robust to implement this approach.
- **An alternative *direct* approach would derive the RPI-based TMR directly using the historical nominal TMR for the UK and a consistent measure of RPI inflation**. Conceptually, a direct approach would appear to be a good alternative as long-run estimates of the CPI series are not available. However, as there have been several changes to the RPI methodology over time, the historical RPI inflation should be adjusted to reflect the current RPI methodology in order to ensure that the real TMR accurately captures investors' expectations of the future equity market returns relative to RPI. We present an estimate of RPI inflation adjusted for consistency in section 6.
- **An RPI-based total market return could also be calculated by deducting forecast RPI inflation from the average nominal historical equity market return.** Although it is conventional to restate historical returns in real terms using historical inflation, given the uncertainty over the historical inflation data the robustness of the analysis may be improved by considering nominal historical returns as well as real historical returns.
- **Nominal arithmetic or geometric average TMR**. The real TMR is typically expressed in terms of the arithmetic rather than the geometric average return for capital budgeting and investment appraisal purposes. The arithmetic TMR could be derived by either (i) starting from a nominal

[14] Dimson, E., Marsh, P. and Staunton, M. (2019), 'Credit Suisse Global Investment Returns Yearbook 2018', February.
[15] UK Regulators Network (2018), op. cit., Appendix D.
[16] Civil Aviation Authority (2019), 'Appendices to Draft UK Reference Period 3 Performance Plan proposals', February, D-19; and Ofgem (2018) 'RIIO-2 Sector Specific Methodology Annex: Finance', 18 December, Table 13.

geometric average TMR, which is then converted to an arithmetic average using an assumed geometric to arithmetic average uplift or (ii) starting with the nominal *arithmetic* average TMR directly. The former approach is used in the UKRN study. Specifically, the study converts the nominal geometric TMR estimate into real terms and then applies an assumed geometric to arithmetic average uplift of 0.8–1.8% to present a range of real arithmetic TMR. On the other hand, should the latter approach be used instead, then the real arithmetic TMR could be derived directly by converting the nominal arithmetic TMR into real terms. This approach would imply a point estimate of the real TMR towards the top end of UKRN's range. We consider that a direct approach of starting with the nominal arithmetic TMR is potentially more robust as it does not require an additional estimation of the geometric to arithmetic average uplift.

We explore these methodological considerations in more detail below.

2.2.2 The UKRN study

The issue of expressing nominal returns in real terms for the purposes of setting the allowed cost of equity was analysed in the UKRN study.[17] Until recently, the GIRY report deflated returns using the historical RPI index. However, the latest editions of the GIRY report deflate returns using a hybrid index that combines historical data on both RPI and CPI. [18]

In particular, the UKRN study showed how the estimates of the real return based on historical averages would change depending on the measure of inflation—namely RPI, CPI and the GIRY report hybrid inflation index. The analysis is summarised in Table 2.1.

[17] UK Regulators Network (2018), op. cit., Appendix D.
[18] See, for example, Dimson, E., Marsh, P. and Staunton, M. (2019), 'Credit Suisse Global Investment Returns Yearbook 2018', February.

Table 2.1 Real historical returns for the UK equity markets, UKRN analysis (%)

	GIRY hybrid index	Historical RPI inflation[1]	CPI inflation
Nominal TMR (geometric average)	9.4	9.4	9.4
Less: inflation	3.7	4.2	4.0
Add: arithmetic uplift	0.8–1.8	0.8–1.8	0.8–1.8
Inflation-adjusted TMR	**6.25–7.25**	**5.8–6.8**	**6.0–7.0**

Note: Based on the data for the period 1899–2016. The following inflation measures are used in the analysis.

The **GIRY hybrid index** relies on CPI from 1988 onwards, RPI from 1962 to 1988, and the index of retail prices before 1962.

The **Historical RPI inflation** series relies on the Bank of England's Millennium dataset. The dataset is based on actual data since 1947 and estimates for the earlier periods.

The **CPI inflation** series relies on the Bank of England's Millennium dataset, which is largely a 'backcast' series as there is no contemporaneous data for CPI before 1989.

[1] We could not reconcile the RPI-based geometric real equity returns as presented by the UKRN. Specifically, our replication of UKRN analysis suggests that a lower real equity return of 5.01% rather that 5.07% should be used in the analysis. The change does not affect the implied real RPI-based arithmetic TMR range expressed to 1 decimal point (5.8–6.8%).

Source: Oxera analysis, based on UK Regulators Network (2018), 'Estimating the cost of capital for implementation of price controls by UK Regulators', March, D–121, E–125; Dimson, E., Marsh, P. and Staunton, M. (2017), 'Credit Suisse Global Investment Returns Yearbook 2017', February.

The 2018 UKRN study recommends a real arithmetic TMR range of 6–7% based on historical CPI inflation.

The range is 0.25% lower relative to the real TMR derived using the GIRY hybrid inflation index. Notably, if the RPI measure of inflation is used instead, the real RPI-based TMR would imply a further lowering of the range to 5.8–6.8%. As a result, the choice of historical inflation index would directly affect the allowed real cost of equity for HAL.

2.2.3 RPI-deflated TMR: direct vs indirect approaches

The UKRN study implies that investors will require a nominal return that is based on the assumed CPI-based real market return of 6–7%. As the HAL price control is indexed to RPI, and RPI inflation is forecast to be higher than CPI inflation, a CPI-based real market return of 6–7% combined with RPI-linked prices and asset values will deliver a higher return than with CPI-linked prices and asset values. For consistency, the real market return should be expressed on an RPI basis as opposed to CPI. There are two potential options for expressing the TMR on an RPI-deflated basis.[19]

- **An indirect approach (i.e. CPI inflation adjusted to RPI)**. Deduct the difference between forecast RPI and CPI inflation (currently approximately 100 bps)[20] from the UKRN-recommended CPI-based TMR of 6–7%. This approach would imply a real RPI-based TMR range of 5–6%. We note that the CAA and Ofgem followed this approach in expressing the UKRN-

[19] An RPI-based total market return could also be calculated by deducting forecast RPI inflation from the average nominal historical equity market return, although this is not the focus of the current paper.
[20] The OBR's March 2019 forecast implies an RPI–CPI wedge of around 104bps (based on RPI and CPI forecasts for 2023). Office for Budget Responsibility (2019), 'Historical official forecasts database', March.

recommended TMR in RPI terms.[21] However, it is not clear that the backcast CPI series is sufficiently robust to implement this approach.

- **A direct approach**. Derive the RPI-based TMR directly using the historical nominal TMR for the UK and a consistent measure of RPI inflation. We note that as there have been many changes to the RPI methodology over time, the historical RPI inflation should be adjusted to reflect the current RPI methodology in order to ensure that the real TMR accurately captures investors' expectations of the future equity market returns relative to RPI.

In addition, the real TMR is typically expressed in terms of the arithmetic rather than the geometric average return for capital budgeting and investment appraisal purposes.[22]

We note that the real CPI-based TMR range of 6–7% TMR recommended by the UKRN starts with the nominal *geometric* average TMR, which is then converted to an arithmetic average using an assumed geometric to arithmetic average uplift. An alternative approach would be to start directly with the nominal *arithmetic* average TMR and then derive a real TMR using the relevant inflation measure. We note that the top end of the CPI-based TMR range presented by the UKRN (i.e. 7%) is close to the historical arithmetic average returns. We consider that a more direct approach of starting with the nominal arithmetic TMR is potentially more robust as it does not require an additional estimation of the geometric to arithmetic average uplift.

Table 2.2 outlines two potential RPI-based ranges derived using (i) both direct and indirect approaches, and (ii) the direct approach of the nominal arithmetic average TMR.

Table 2.2 Real historical returns for the UK equity markets (arithmetic nominal average)

	Indirect approach: CPI inflation adjusted to RPI	**Direct approach: RPI inflation[1]**
Nominal TMR (arithmetic average)	11.2	11.2
Less: inflation	5.0	4.2
Inflation-adjusted TMR (RPI-based)	**6.0**	**6.7**

Note: Based on the data for the period 1899–2016. Real arithmetic TMR is derived by converting nominal arithmetic average TMR using the Fisher equation and geometric average inflation (as presented in Table 2.1). An alternative approach would be to (i) convert the nominal *annual* equity market returns into the real terms using the Fisher equation and annual inflation and then (ii) take an arithmetic average of the resulting real annual equity market returns. The difference in the real arithmetic TMRs under the two approaches is less than 0.01%.

[1] Based on the historical RPI dataset published by the Office for National Statistics. The real arithmetic TMR presented in this table (6.7%) is slightly different from the top end of the range for the real arithmetic TMR presented in Table 2.1 (6.8%) due to rounding. Specifically, the arithmetic TMR presented in this table is directly based on arithmetic average returns, while the

[21] Civil Aviation Authority (2019), 'Appendices to Draft UK Reference Period 3 Performance Plan proposals', February, D-19; and Ofgem (2018) 'RIIO-2 Sector Specific Methodology Annex: Finance', 18 December, Table 13.
[22] While there is debate about which is the more appropriate averaging method in any given context, the academic literature is broadly supportive of placing more weight on the arithmetic averages for estimating the equity risk premium to use when computing required equity returns. See, for example, Dimson, E., Marsh, P. and Staunton, M. (2015), 'Credit Suisse Investment Returns Sourcebook 2015', February, p. 34; and Jacquier, E., Kane, A. and Marcus, A.J. (2003), 'Geometric or Arithmetic Mean: A Reconsideration', *Financial Analysts Journal*, **59**:6**,** November/December, pp. 46–53.

arithmetic TMR presented in Table 2.1 is based on an assumed geometric to arithmetic average uplift. The actual difference in real arithmetic TMRs is less than 0.03%.

Source: Oxera analysis, based on UK Regulators Network (2018), 'Estimating the cost of capital for implementation of price controls by UK Regulators', March; Dimson, E., Marsh, P. and Staunton, M. (2017), 'Credit Suisse Global Investment Returns Yearbook 2017'; and the RPI data from the Office for National Statistics.

Overall, the indirect approach would imply an RPI-based arithmetic TMR of 6% (assuming a 1% RPI–CPI wedge), while the direct approach would imply an RPI-based arithmetic TMR of 6.7%.

As noted by the CAA, long-run CPI data, necessary to adjust TMR, are not available.[23] This means that all CPI measurements before 1988 need to be estimated. The UKRN study relies on an estimate of CPI provided by the Bank of England. The Millennium dataset inflation measure is published by the Bank of England as part of a research dataset. We understand that it does not have national statistics status. The CAA considers that for 1949 to 1988, the CPI series is estimated based on the removal of the 'formula effect' difference between RPI and CPI.[24] The CAA notes that before 1949, the CPI and RPI series were essentially the same. In other words, the UKRN choice of CPI rests on the premise that it is possible to find a reliable estimate of the 'formula effect' between 1949 and 1988.

Conceptually, we consider that a direct approach would be a good alternative, as long-run estimates of the CPI series are not available. However, there have been significant changes to the RPI methodology over time, meaning that that measure itself is not necessarily consistent. In order to provide a consistent estimate of inflation on an RPI basis, and to ensure that the real TMR accurately captures investors' expectations of the future equity market returns, the reported RPI inflation should be adjusted to reflect the current RPI methodology.

2.2.4 Conclusion

Using different inflation measures yields different real TMR estimates. This highlights the challenges when selecting an inflation measure. This is made more challenging because inflation is not directly observable and needs to be estimated. Statistical agencies rely on a sub-sample of prices and mathematical formulae to build different estimates depending on their purpose. This means that all inflation measures need to be assessed depending on the coverage of the prices they include (both from a geographic, historical and product range perspective) as well as their mathematical properties.

Conceptually, we consider that a direct approach for deriving an RPI-based TMR would be a good alternative to an indirect approach that starts with the CPI-based TMR, as a long-run CPI series is not available. However, there have been significant changes to the RPI methodology over time, meaning that that measure itself is not necessarily consistent.

In the remaining sections of this report, we evaluate the changes in the RPI methodology over time in order to derive RPI-based real equity returns using a consistent RPI-based measure of inflation.

[23] Civil Aviation Authority (2019), 'Draft UK Reference Period 3 Performance Plan proposals', sections D18-D24.
[24] Civil Aviation Authority (2019), 'Draft UK Reference Period 3 Performance Plan proposals', sections D18-D21.

3 The statistical history of the RPI

This section describes the history of the RPI, the main methodological improvements to its computation, and the magnitude and direction of those changes.

The price of goods and services in the economy tend to rise over time. This is known as inflation. An inflation index attempts to measure the overall rate of growth in prices across the economy. The RPI is one example of a ‘consumer price index’, a measurement of the rate of price growth experienced by consumers in the economy. Different households purchase different goods and services and therefore have different costs; this leads to a variety of different consumer price indices that were devised for different purposes.

The RPI measure of inflation arose in the late 19th century from a need for governments to improve economic policy and better understand the lives of their citizens. Its early history was marked by the two World Wars and an economy in which prices were controlled for long periods.

More recent history has seen the birth of a more scientific approach to the estimation of inflation. The RPI has been affected by the 2010 ‘formula effect’, which saw the difference between RPI and CPI increase substantially overnight after methodological changes. These methodological changes had a significant impact on the reported measure of RPI, which we assess using statistical techniques.

3.1 The history of the RPI

The following subsections chart the history of the RPI from its inception to the present day.

3.1.1 Early history: 1880–1947

The first retail price indices were produced in 1903, using prices collected in the late 19th century in Britain. These indices covered only the prices of food, clothing and rent.[25] Reports were published in 1904, 1908 and 1913, over which period fuel prices were included and the geographic coverage of sampling was progressively increased.[26]

In 1914, the cost of living index (‘CLI’) was published for the first time.[27] Its objective was to measure the cost of maintaining a basic standard of living for working class households. It was a key element in the decision-making of the British government, which was trying to protect workers from the economic consequences of the First World War.[28]

Prices were captured on a monthly basis, but weights were not changed despite changing consumption patterns. The CLI measure of inflation was also criticised on the grounds that it made judgements on the way working class households ought to live their lives (e.g. beer was excluded and the weight given to tobacco was disproportionately low).[29]

[25] O’Neill, R., Ralph, J. and Smith, P.A. (2017), *Inflation: History and Measurement*, Springer, section 5.3.
[26] Ibid., sections 5.4–5.6.
[27] A 1984 Advisory Committee states that the RPI is an index of price changes and not a cost of living index (Ibid., p. 152). The main difference between a cost of living and a cost of goods index is that the former is more flexible as it allows consumers to adjust their consumption, which is not allowed by the 1914 CLI (ibid., p. 267).
[28] Johnson (2015), op. cit., section 3.2.
[29] Ibid.

Given the methodological limitations of the CLI, a new expenditure survey was undertaken in the late 1930s to produce a new index. However, the outbreak of the Second World War disrupted this work.[30]

Figure 3.1 Summary of the early history of the RPI

1903
- The first official set of price indices for the years 1877–1901
- Based on very small sample sizes and limited geographic coverage

1904
- Expenditure survey carried out to address the limited expenditure share information in earlier price series
- While this represented a major improvement over the available expenditure shares, issues remained, including non-random sampling, failure to include a comprehensive range of goods and services, and limited sample size

1908
- Extended the geographic range of rent, food, and clothing prices

1914
- 'Cost of living index' developed. Used during the First World War to protect workers from the economic consequences of the war. Weights unchanged despite changing consumption patterns

1937
- Expenditure survey with major methodological improvements over the 1904 one, including the use of random sampling and a 5x increase in the sample size. These weights would only be used for 1947–55
- Nonetheless, the Second World War and the ensuing lack of stability in expenditure weights and prices meant that expenditure shares were not compiled again until 1956

1947
- While this set of 'Interim' indices used pre-war expenditure shares, the collection of prices moved significantly towards modern practice, with the inclusion of alcohol, significantly more items (5x the number of food items), and improvements in the accounting for quality changes in clothing

O'Neill, R., Ralph, J. and Smith, P.A. (2017), *Inflation: History and Measurement*, Springer; and Johnson, P. (2015), 'UK Consumer Price Statistics: A Review', January.

3.1.2 Introduction and development of the RPI: 1947–present

In the aftermath of the Second World War, the Cost of Living Advisory Committee ('CLAC') was set up. Its role was to consider the future of the CLI measure of inflation.[31]

In 1947, the Interim Index of Retail Prices was introduced, using the weights data collected in the late 1930s.[32]

[30] O'Neill, Ralph and Smith (2017), op. cit., pp. 125–126.
[31] O'Neill, Ralph and Smith (2017), op. cit., p. 134.
[32] O'Neill, Ralph and Smith (2017), op. cit., p. 136.

The year 1956 was a turning point in the history of retail price indices. It saw the birth of the modern RPI (then called the Index of Retail Prices). This new index included a range of important methodological improvements. In particular, all wage-earning households were included—not only the working class.[33] The index took its weights from the more recent 1953 expenditure survey, rather than the pre-war late-1930s survey.[34] Owner-occupier housing costs were also included for the first time.[35]

The RPI methodology was progressively improved during the 1960s–1980s. From 1962, expenditure weights were updated on an annual basis.[36] In 1968, prices of food and drink purchased in restaurants were introduced. In 1975, mortgage interest payments were introduced to represent owner-occupiers' housing costs.[37] This was prompted by concerns that equivalent rents did not measure housing costs for owner-occupiers well, especially given recent rises in interest rates and the growth of owner occupation (meaning that more than half of all households fell into this category).[38]

Further methodological changes were implemented in 1986. In particular, it was decided to exclude the top 4% of households, based on their household income (before this, households earning more than a certain amount were excluded). In the following years, holidays started being included as well.[39]

The 1990s saw the birth of competing inflation indices. Notably, the European Union introduced the Harmonised Index of Consumer prices ('HICP'), which provided a unified and comparable inflation index across the EU. The CPI is the UK version of the HICP.[40] Box 3.1 explains the differences between the CPI and the RPI.

[33] Johnson (2015), op. cit., p. 118.
[34] Expenditure weights calculated from the Family Expenditure Survey (FES) were monitored for 'significant changes'. The FES was about a quarter of the sample size of the 1953 Budget enquiry.
[35] Using the concept of equivalent rents. See O'Neill, Ralph and Smith (2010), op. cit., pp. 141–143.
[36] This was done by taking a rolling average over three years.
[37] O'Donoghue, J., McDonnell, C. and Placek, M. (2006), 'Consumer price inflation, 1947–2004', *Economic Trend*, **626**, January.
[38] O'Neill, Ralph and Smith (2017), op. cit., p. 150.
[39] Johnson (2015), op. cit., p. 49.
[40] The HICP was renamed to the CPI in December 2003. See O'Neill, Ralph and Smith (2017), op. cit., p. 230.

Box 3.1 Differences between the CPI and the RPI

Population base

The CPI and the CPIH (i.e. CPI, but including owner-occupiers' housing costs) cover a broader population than the RPI. The RPI covers only private households and excludes the top 4% of households by income, as well as excluding pensioner households that receive at least 75% of their income from benefits. The CPI and CPIH cover the expenditure of all private households, institutional households and visitors to the UK.[1]

Item coverage

The coverage of the CPIH and the CPI is currently identical, except for the inclusion of a measure of owner-occupiers' housing costs in the CPIH. The RPI includes certain items relating to housing costs (such as mortgage interest payments) that are not included in the CPIH or the CPI. Conversely, there are also some services covered by the CPIH and the CPI—such as charges for financial services—that are not covered in the RPI.

Index methodology—formula

The CPIH and CPI methodologies mostly use the geometric mean (with some use of the arithmetic mean), whereas the RPI methodology uses the arithmetic mean to combine prices at the first stage of aggregation.

Item coding

The CPIH and CPI structure follows a standard international classification system, whereas the RPI has its own unique system.

Note: [1] Office for National Statistics (2017), 'Consumer Price Inflation (includes all three indices — CPIH, CPI and RPI) QMI', 20 December, section 5.

Source: Office for National Statistics (2017), 'Consumer price indices, a brief guide', section 13: 'Differences between the CPIH, CPI and RPI', 3 November, https://www.ons.gov.uk/economy/inflationandpriceindices/articles/consumerpriceindicesabriefguide/2016, accessed 28 June 2019.

In 2003, the UK government changed the measure used for inflation targeting from the RPI to the CPI. Aligned with this decision, in 2010, it was announced that the CPI would become the index of reference for the indexation of pensions and other benefits. In the same year, the 'formula effect' controversy arose.[41] In this year, the difference between the CPI and the RPI due to the formula effect increased from approximately 54bps to approximately 86bps following minor changes to the way that clothing prices were collected.[42] This arose because of the property of a mathematical formula used to compute the RPI—the 'Carli index' (see section 4.1).

In 2013, the RPI lost its status as an official statistic following an inquiry by the ONS into the formula effect. It was found that the UK was the only country to use a price index based on the Carli index.[43]

The CPIH gained the status of official statistic in 2017.

[41] The formula effect measures the impact of the use of different mathematical formulae to construct the CPI and RPI.
[42] Office for National Statistics (2011), 'CPI and RPI: increased impact of the formula effect in 2010'.
[43] Johnson (2015), op. cit., p. 7.

Figure 3.2 Summary of the methodological improvements in RPI from 1956–1987

1956

- Expanded to included all wage-earning households, not just the working class
- First attempt to measure owner-occupier housing costs with the introduction of equivalent rents, whereby owner-occupier costs were represented by estimates of the rent they would pay if their dwelling was rented, rather than owned
- Motor vehicle purchase costs were included, adding 0.8% to the weight of the transport and vehicles group

1962

- Household expenditure re-evaluated every year

1968

- Eating out introduced to the index as it was becoming an important item of expenditure

1975

- Mortgage interest payments were introduced to represent owner-occupiers' housing costs. This was prompted by concerns that equivalent rents did not measure housing costs for owner-occupiers well, especially given recent rises in interest rates and the growth of owner occupation (meaning that more than half of all households fell into this category)
- The index using the mortgage interest method was higher than using the rental equivalence method, and was closer to the movement in house prices. This change in method also changed their respective weights[1]

1986

- Changing from excluding households whose head earned more than a certain amount to excluding the top 4% of households based on their total household income[2]

1987

- Overhaul of the structure of the RPI, with durable households, miscellaneous goods and services groups being restructured into five new groups

Note: [1] O'Neill, R., Ralph, J. and Smith, P.A. (2017), *Inflation: History and Measurement*, Springer, p. 150. [2] Johnson, P. (2015), 'UK Consumer Price Statistics: A Review', January, p. 49.

Source: O'Donoghue, J., McDonnell, C. and Placek, M. (2006), 'Consumer price inflation, 1947–2004', *Economic Trend*, **626**, January, unless otherwise cited.

Figure 3.3 Summary of the methodological improvements in RPI from 1993–present

1993

- Inclusion of foreign holidays, which added 3.0% to leisure services' weight that year
- Council Tax included in RPI, replacing the community charge

1994

- Inclusion of UK holidays, adding 0.8% to leisure services' weight that year
- Exclusion of the investment portion of owner-occupied housing costs

1995

- Inclusion of housing depreciation, adding 3.0% to housing costs' weight that year
- Before 1995, the locations from which prices were collected largely reflected the location and availability of civil servants in Unemployment Benefit Offices around the country who were then carrying out the price collection on behalf of the ONS. Therefore, the choice of locations had not changed for many years. In 1995, this changed to a more systematic process, ensuring that the country was fully represented and that the number of locations selected reflected the proportion of total consumer expenditure for each Government Office Region[1]

2010

- Methodological improvements to the measurement of clothing, which began to be implemented in January 2010
- Change in method from calculating the mortgage interest rate using the standard variable rate to a weighted method

2012

- The Statistics Authority decided to make no further improvements to the RPI measure of inflation[2]

2013

- The RPI's status as a national statistic was revoked and statistical authorities began to advocate actively against its use

2017

- The CPIH re-designated as a national statistic[3]

Note: [1] Office for National Statistics (2014), 'Consumer Price Indices Technical Manual 2014 Edition', January, p. 21. [2] House of Lords (2019), 'Measuring inflation', 17 January, p. 4. [3] Office for National Statistics (2017), 'Consumer Prices Index including Owner Occupiers' Housing Costs (CPIH) re-designation', 31 July.

Source O'Donoghue, J., McDonnell, C. and Placek, M. (2006), 'Consumer price inflation, 1947–2004', *Economic Trend*, **626**, January, unless otherwise cited.

3.2 Assessing the impact of methodological changes on inflation measurement

As explained above, inflation forecasts and backcasts are both used in the context of price regulation. The methodological changes to the RPI raise the question of what inflation in the past would have been had the methodology been consistent through time and whether past values need to be updated.

Below, we propose a system of categorisation for assessing the impact of methodological changes on RPI. For a given measure, they can be divided in three broad categories:

1. **weight changes**, where the weights associated with a given item or group of items evolve;
2. **inclusion of new products or classes of products**, where goods and services that were previously not accounted for are added;
3. **changes to sampling methodology**, such as an increase in the geographic coverage or population base.

These categories are not mutually exclusive as these methodological changes are intertwined. For instance, the introduction of a new product in the sample means that the weights of all the others products are diluted.

The effect of weight changes on the RPI depends on whether the product that is given more importance has a relatively high inflation compared with other products, including those that are given less importance. For instance, if greater weight had been given to petrol during the oil shock in 1973, one could have expected the inflation measure to increase.

Similarly, the inclusion of a new product will positively affect inflation measures if it is a relatively high inflation product compared with other products, and whether this comes from a decrease in importance of a relatively low inflation product. For instance, foreign holidays, introduced in 1993, roughly match the rate of inflation of the rest of leisure services, and are therefore not likely to affect inflation substantially (assuming that their introduction was not done by removing a particularly high-inflation or low-inflation product).[44]

The inclusion of new products and weights reflects the evolution of consumption patterns. However, some also reflect general improvement in the methodology or the fact that it was not possible to account for all items in the past. In other words, some changes would have been made earlier, had it been possible. As it is possible to assess the effect of weight changes and the inclusion of new products on the RPI, it is also possible to build an adjusted RPI series where important changes would have been made earlier. We will discuss this possibility further in section 6.2.

The effect of changes to the sampling methodology on inflation measures is harder to assess. They might increase inflation if they lead to an increase in price dispersion, for instance by sampling prices in areas where inflation increases faster. However, these are hard to evaluate as we do not have access to such disaggregate information.

[44] From 1994 to 2019 (the period when foreign holidays were included), foreign holiday prices averaged 3.92%, while leisure services averaged 3.93%. This is based on Oxera analysis of ONS RPI data, series SVGL and CZEN.

3.3 What major changes have there been to the RPI calculation?

There have been several important changes to the RPI methodology. The second half of the 20th century and early 21st century have also seen major shocks and changes to the global economy, ranging from reconstruction after the Second World War to ‘stagflation’ in the 1970s and the subprime crisis in the late 2000s.

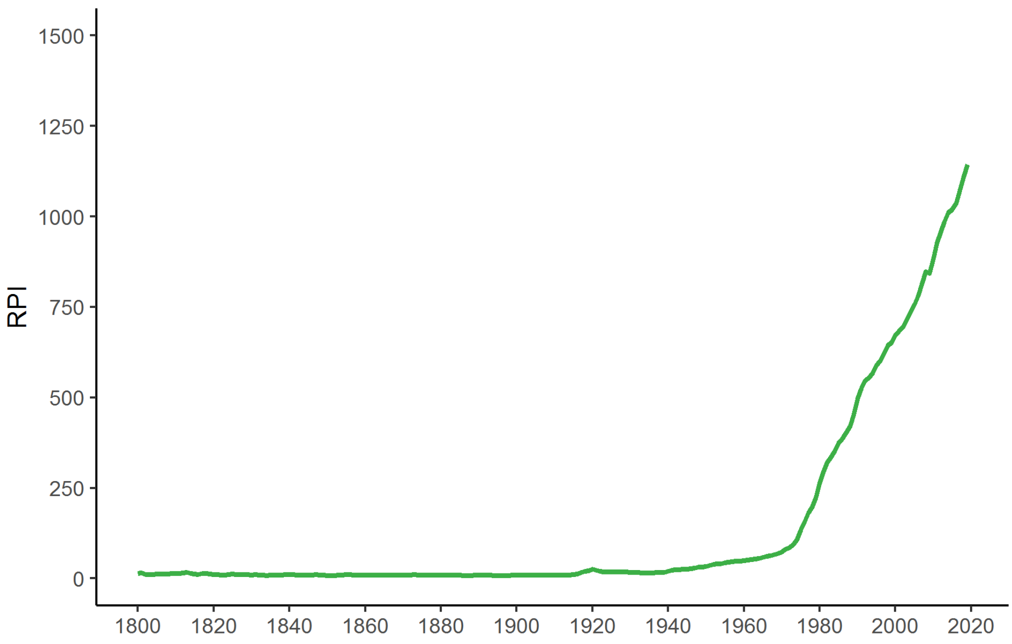
While all changes in the methodology and the UK economy affect inflation, some have deeper consequences than others. In this section, we try to identify the major changes in the RPI series by relying on a statistical approach.

The methodology we use is called ‘structural breaks testing’. Structural breaks represent immediate changes in a time series. There are two main ways to identify them: the first relies on the judgement of the modeller to make assumptions about the dates of each structural break; the second does not make any assumption about the date of the break and ‘lets the data speak’. More precisely, the second approach sees the use of an algorithm to identify the dates where it thinks a shock changes the pattern of the series.

We follow the latter approach in this section. Indeed, given the complexity of the economic activity and the number of methodological changes in the calculation of the RPI series, we do not want to prejudge which changes made a material difference. We use a machine learning method called ‘indicator saturation’ (described in Appendix A2) to identify periods for which RPI varies unexpectedly.

The evolution of the RPI between 1900 and 2018 is shown in Figure 3.4. One major change is visible, with inflation increasing substantially from the 1970s onwards.

Figure 3.4 Long-run evolution of the RPI series

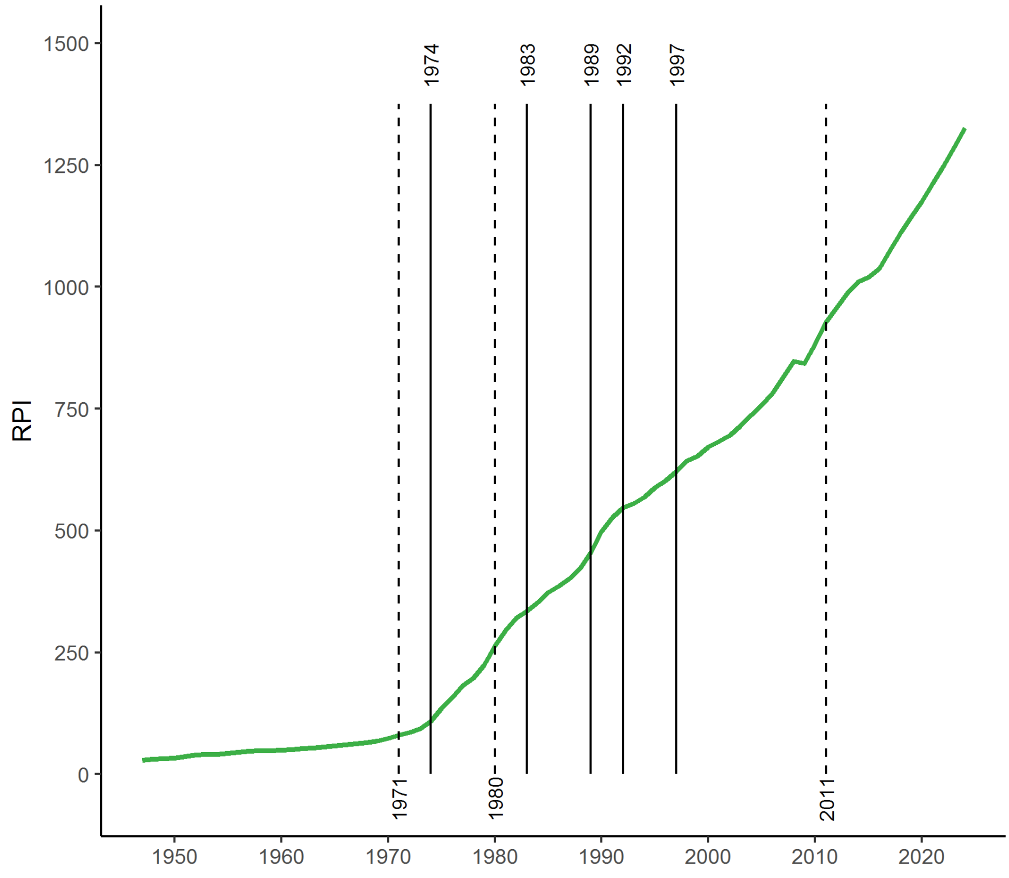


Source: Office for National Statistics, ‘Retail Prices Index: Long run series: 1947 to 2019’ (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

Since the modern RPI series started in 1956, the structural breaks analysis presented here uses only post-1956 data. We conducted a similar analysis on a longer series including pre-1956 data; that analysis gave broadly similar results.

Figure 3.5 shows the results of our structural breaks analysis. We identify two types of break. The first type is 'level' shifts, represented by dashed lines. Level shifts are when the RPI measure shifts between two years then continues its growth at the same pace. The second type we refer to is 'trend' shifts, represented by solid lines. Trend shifts indicate that the rate at which the RPI series grows has changed. We also present the results of these tests with the natural logarithms of the RPI series in Figure 3.6. Since logarithms tend to reduce variations in a statistical series, fewer breaks are found, although they are consistent with the results presented in Figure 3.5.

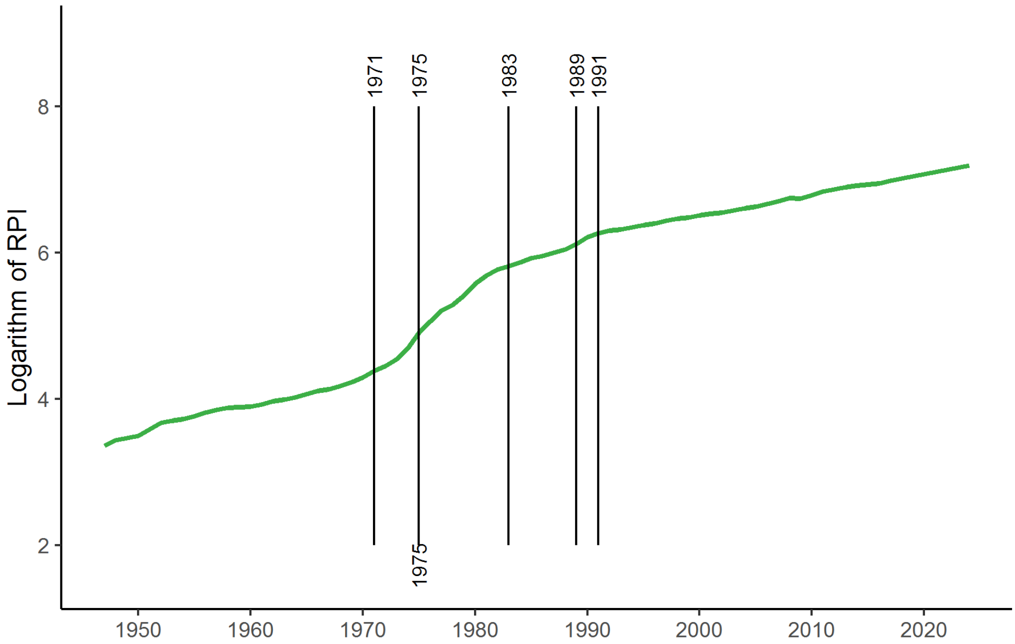
Figure 3.5 Structural breaks in the RPI series



Note: Solid lines represent 'trend' shifts while dashed lines represent 'level' shifts.

Source: Oxera analysis; and Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

Figure 3.6 Structural breaks in the RPI series, in logarithms



Note: A shift in level and a shift in trends are identified in 1975.

Source: Oxera analysis; and Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

In Table 3.1, we explore potential explanations for these breaks. These may be based on some methodological changes in RPI measurement and/or major economic events. As both happen at the same time, it is impossible to clearly disentangle them.

For instance, our method identifies an increase in inflation in the mid-1970s, which corresponds to the stagflation period when inflation became high and growth low. We also identify a positive shift in the RPI series in the early 2010s, which may be due to the 'formula' effect or recovery from the Great Recession. To assess the robustness of our results, we also tested other specifications. These are detailed in Appendix A3 and are broadly consistent with those described here.

Table 3.1 Effect of identified breaks on RPI series

Year	Effect	Potential explanation	Nature
1971	Level increase	• macroeconomic policies leading to stagflation	Economic
1974	Trend increase	• 1973 oil crisis • macroeconomic policies leading to stagflation • switch from equivalent rents to mortgage interest payments as a measure of owner-occupiers' housing costs	Economic or methodological
1980	Level increase	• macroeconomic policies leading to stagflation	Economic
1983	Trend decrease	• Margaret Thatcher's policy to promote low inflation	Economic
1989	Trend increase	• Lawson boom	Economic
1992	Trend decrease	• recession arising from US savings and loan crisis • introduction of domestic and foreign holidays	Economic or methodological
1997	Trend increase	• introduction of housing depreciation in 1995	Methodological
2011	Level increase[1]	• changes in clothing price methodology • recovery from the Great Recession	Economic or methodological

Note: These results are based on the level RPI series (not logged). [1] The structural break test did not identify the 2011 break as a trend increase, which indicates that in addition to the change in the way clothing prices were collected, other factors also changed at this point in time.

Source: Oxera analysis; and Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

We interpret these results as indicative of which methodological changes may be significant enough to produce a material change in the RPI series. These include:

- the switch from measuring owner-occupied housing using the method of equivalent rents to using mortgage interest payments in 1975;
- the inclusion of foreign and domestic holidays in 1993 and 1994 respectively;
- the inclusion of housing depreciation in 1995;
- the change in the clothing price methodology in 2011.

4 The RPI today

In this section, we present the way the RPI is computed today. The objective is to gain a better understanding of the methodological challenges associated with its calculation.

The RPI comprises three main elements:

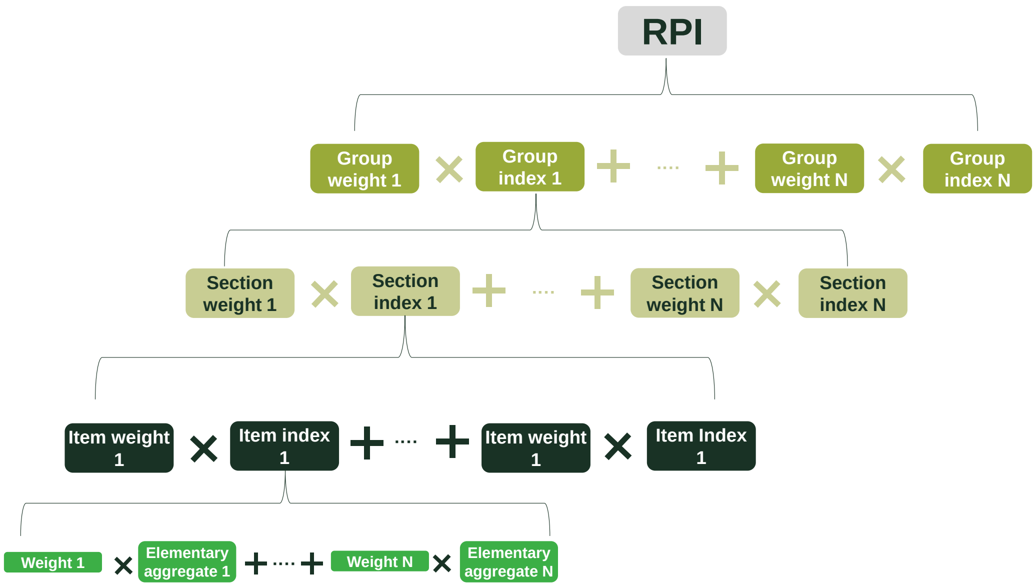
1. **elementary aggregates**, combining individual quotes of a given product or service collected around the UK (e.g. food, motoring expenditure, and clothing and footwear);
2. **indices**, used to summarise the price information of multiple products;
3. **weights**, giving the right importance to different products to represent their share in household budgets.

Combining these elements, at the highest level, the RPI can be described as a weighted average of indices and the elementary aggregates that compose them.

In a first stage, elementary aggregates for a given product are combined into ‘item indices’ using ‘stratum weights’. For instance, an elementary aggregate would be ‘electricity’ in a region of the UK.[45] The corresponding item index would be ‘electricity’ throughout the UK. Items are themselves aggregated into section indices such as ‘fuel and light’. In a final stage, ‘higher level’ or ‘group’ indices represent the broadest categories. In the electricity example, it would correspond to ‘housing and household expenditure’.[46]

This aggregation process is illustrated in Figure 4.1.

Figure 4.1 Stylised representation of the RPI calculation



Note: ‘N’ represents a hypothetical number of goods.

Source: Oxera analysis.

[45] The elementary aggregates of ‘electricity’ can potentially be composed of different products that would correspond to different electricity providers or plans in the UK.
[46] For more details see Office for National Statistics (2014), ‘Consumer Price Indices Technical Manual’, p. 40 and Appendix 5.

4.1 Elementary aggregates

There are three main formulae used to calculate elementary aggregates:[47]

- the Carli formula—the arithmetic mean of the ratio of price relatives;
- the Dutot formula—the ratio of the arithmetic mean of prices;
- the Jevons formula—the geometric mean of the ratio of price relatives.

Choosing which elementary aggregates formulae to use is critical and is a key difference between the RPI and the CPI. As shown in Table 4.1, the RPI relies heavily on Carli and Dutot formulae;[48] by contrast, the CPI relies mostly on the Jevons formula. The use of the Carli formula in RPI but not CPI is the source of the 2010 formula effect mentioned above.

Table 4.1 Formulae used in RPI and CPI (2012 weights)

	RPI	CPI
Carli	27%	0%
Dutot	29%	5%
Jevons	0%	63%
Other/weighted formula	43%	33%

Source: UK Statistics Authority (2016), 'Elementary aggregate formula', https://www.statisticsauthority.gov.uk/wp-content/uploads/2016/04/Elementary-aggregate-formula-description-action-2.pdf, last accessed 2 June 2019.

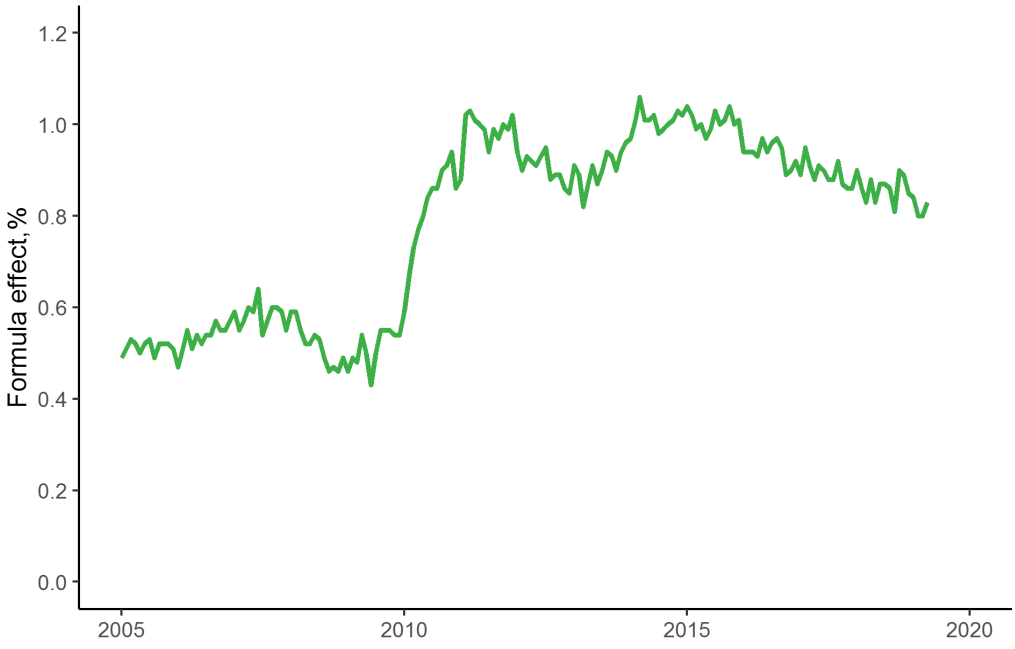
Indeed, because of its mathematical properties, the Carli formula consistently produces higher results for measured inflation than either the Jevons formula or the Dutot formula.[49] In contrast, the Dutot formula may be either greater or less than the Jevons formula, allowing for the possibility that these effects will 'cancel out'.[50] There is some evidence that, in practice, the difference between the Jevons index and the Dutot index is not large.[51]

The formula effect increased further in 2010, when the ONS implemented methodological changes to the data collection process for clothing and footwear.[52] These modifications led to an increase in the dispersion of the clothing prices that are collected, resulting in the 32bps increase in difference between the CPI and the RPI inflation statistics as estimated by the ONS.[53]

The ONS publishes detailed statistics on the source of the difference between the RPI and the CPI. For instance, these statistics allow us to identify whether they differ because they use different weights or because of the mathematical properties of the underlying indices. Figure 4.2 illustrates the difference between the RPI and the CPI solely due to the use of different elementary aggregates. We note that this has been positive in recent periods, reflecting the greater use of the Carli index in the RPI methodology, and the increased formula effect in 2010.

[47] A further explanation of each of these formulae is provided in Appendix A1.
[48] An alternative version of the RPI, the RPIJ, relies mostly on the Jevons formula.
[49] Johnson (2015), op. cit., p. 51.
[50] Levell, P. (2012), 'A winning formula? Elementary indices in the Retail Price Index', November, p. 7.
[51] Silver. M. and Heravi. S. (2006), 'Why Elementary Price Index Number Formulas Differ: Price Dispersion and Product Heterogeneity', July, p. 7.
[52] For an explanation of what these improvements were, see Office for National Statistics (2011), 'CPI and RPI: increased impact of the formula effect in 2010', pp. 2–3.
[53] This formula effect is estimated as there are other differences between RPI and CPI. For instance, Ofgem has estimated a formula effect of 40bp. Ofgem (2014) 'Decision on our methodology for assessing the equity market return for the purpose of setting RIIO-ED1 price controls', 17 February, para 1.17.

Figure 4.2 Contribution of the formula effect to the wedge between the RPI and the CPI



Source: Oxera analysis based on Office for National Statistics (2019), 'Difference between CPI and RPI due to formula effect 2015=100', 19 June. https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/dra9/mm23, last accessed 18 June 2019.

4.2 Indices

There are multiple ways to aggregate information.[54] Both the CPI and the RPI are Laspeyres-type indices. This means that they measure the average change in prices between a price reference period and a comparison period using expenditure shares from a period prior to the reference period.[55] In other words, the goods basket is fixed and inflation is measured with reference to it. In practice, it is the most recent basket available in the last 12 months, and the price of this basket is compared between the current period and the last 12 months.[56]

4.3 Weights

The weights are used to ensure that the inflation measure correctly summarises the consumption patterns of households. As a consequence, the weights attached to each good vary over time and need to be re-adjusted to account for changes in consumption habits and revenues.[57]

These weights are based on the Living Costs and Food Survey ('LCF'),[58] which estimates household income and expenditure via a survey of households. This is used to identify potential new items and review existing items.[59] The LCF is a voluntary sample survey of private households. A random sample of households is selected and approached to take part in the survey. For those

[54] See Johnson (2015), op. cit., Annex C for a technical presentation.
[55] See Eurostat (2019), 'What is a Laspeyres-type index', https://ec.europa.eu/eurostat/web/hicp/faq, last accessed 13 June 2019.
[56] See Office for National Statistics (2017), 'Consumer/Retail Prices Indices Microdata', 10 April, p. 28.
[57] For instance, when households get richer they tend to consume different goods, such as those they were unable to afford in the past.
[58] The modern version of the Family Expenditure Survey, which was introduced in the mid-1950s.
[59] O'Neill, Ralph and Smith (2017), op. cit., p. 166.

who agree to participate,[60] there are two parts to the data collection. The first is an interview and the second is a diary, in which the household records all their items of expenditure for two weeks. It is documented that there are several categories where there is consistent under-reporting of expenditure, including alcohol and tobacco. In this case, adjustments are made to estimate their true expenditure weights.[61]

The information from the LCF is used to construct the section weights for RPI. The main exceptions are for some housing sections including mortgage interest payments and depreciation, where other sources are used.

As Table 4.2 shows, there have been major changes in the RPI basket since the creation of the RPI in 1956. Some of these reflect evolving tastes and consumer habits, such as the decline in expenditure share on alcohol and tobacco, while others may be due to methodological changes. For example, the addition of foreign holidays raised the expenditure on leisure services from 32 parts per thousand in 1992 to 62 parts per thousand in 1993.

Table 4.2 Changes in the RPI basket over time, parts per 1,000

	1956	**2018**
Alcoholic drink	71	56
Clothing and footwear	106	43
Food	350	102
Fuel and light	55	37
Housing	87	257
Leisure services	23	85
Tobacco	80	23
Others[1]	228	397
All items	**1,000**	**1,000**

Note: [1] Other items include transport and vehicles, services, and household and other goods.

Source: Oxera analysis.

[60] To reduce the effect of non-response, various adjustments based on the Census, where response is compulsory, are used. See O'Neill, Ralph and Smith (2017), op. cit., p. 207.
[61] While the LCF has come a long way from the first expenditure surveys, there are still several issues. For example, the LCF yields different results when compared with market research sources. Declining response rates are also increasingly an issue. For a more detailed description, see O'Neill, Ralph and Smith (2017), op. cit., section 9.5; and Office for National Statistics (2017), 'Living costs and food survey: technical report for the survey year 2015/16', February.

5 What are the desirable properties of a good inflation measure for regulation?

This section describes the desirable properties of an inflation measure in the context of economic regulation. Since inflation indices summarise thousands of sources of information in one single figure, they need to satisfy some mathematical properties to be mathematically consistent. We present these properties below, and highlight why the use of the RPI for regulation increases the difficulty of finding a good inflation measure.

5.1 Desirable properties of indices from a statistical perspective

As inflation is not directly observable for the whole economy, statistical agencies use indices as mathematical constructs to estimate the overall change in the level of end prices in the economy. These indices need to be able to aggregate prices for a sample of products purchased by a representative sample of the population. Over time, the ONS has attempted to improve its sampling procedure to produce a more representative sample of products for a more representative group of households.

However, even if the prices of all goods and services for all households could be collected, they would still need to be aggregated in a single measure that summarises all the available information. As explained in section 4, the way that prices are aggregated also matters for the construction of price indices. In other words, there is no correct way to aggregate prices because they are artificial constructs. However, there are better ways to design indices depending on what the statistician is trying to achieve. In the case of inflation indices, the objective is to track the aggregate behaviour of all the prices and quantities through time.[62]

Whether one index is better than another is assessed by the properties it exhibits. In the academic literature, price indices are often considered as 'good' when they satisfy a set of 10 features.[63]

- **Feature 1 (monotonicity in prices)**: if any of the current period prices are increased (decreased), then the resulting price index also increases (decreases).
- **Feature 2 (proportionality, in prices)**: if all current prices in the current or base period are multiplied by the same factor, then the price index is multiplied by the same factor. For instance, if all prices in the current period double, then the index doubles.
- **Feature 3 (identity)**: if the prices in the base and current periods are the same, the index shows no change.
- **Feature 4 (homogeneity)**: multiplication of all comparison and base period prices by the same factor does not change the price index number.
- **Feature 5 (change in units of measurement)**: changing the units of measurement of each of the items should not change the result of the price index. The Dutot index fails this test.

[62] Balk, B.M. (2012), *Price and quantity Index Numbers: Models for Measuring Aggregate Change and Difference*, Cambridge University Press, p. 56.
[63] For a technical presentation, see Balk (2012), op. cit., section 3; and Johnson (2015), op. cit., Annex D. Of these 10 features, some are described as 'axioms' and others can be 'tests'. Axioms are 'stronger' than tests and should not be violated. Features 1–6 are axioms. Two others tests are often added (product test and factor reversal test). See Balk (2012), op. cit., section 3.4; and Johnson (2015), op. cit., Annex D.

- **Feature 6 (mean value test)**: the price index lies between the minimum and maximum of the price relatives.
- **Feature 7 (symmetric treatment of prices)**: changing the order of prices in the current or base periods (but retaining the same pairing of prices) does not change the price index.
- **Feature 8 (price bouncing)**: changing the order of prices in the current or base period (and allowing the pairings to be different) does not change the price index. The Carli index fails this test.
- **Feature 9 (time reversal)**: if the data for the current and base periods is swapped, then the resulting index is the inverse of the original price index. The Carli index fails this test.
- **Feature 10 (circularity)**: the price index from going from the base period to the current period directly is the same as the price index for the base period to an intermediary period multiplied by the price index for the intermediary to the intermediary period. The Carli index fails this test.

The results are summarised in Table 5.1.

Table 5.1 Summary of features of a 'good' price index

Feature	Carli	Dutot	Jevons
1. Monotonicity	✓	✓	✓
2. Proportionality	✓	✓	✓
3. Identity	✓	✓	✓
4. Homogeneity	✓	✓	✓
5. Change in units of measurement	✓	×	✓
6. Mean value test	✓	✓	✓
7. Symmetric treatment of prices	✓	✓	✓
8. Price bouncing	×	✓	✓
9. Time reversal	×	✓	✓
10. Circularity	×	✓	✓

Source: Oxera.

5.2 Desirable properties from the perspective of economic regulation

The statistical properties of a desirable index are important in ensuring that the index does not lead to counterintuitive results in practice. However, there are other considerations when choosing an appropriate index for the purposes of economic regulation.

In 2016, Oxera considered possible criteria for evaluating what constitutes a suitable index.[64] In this section, we evaluate how the RPI performs against these criteria and use this insight to inform potential improvements to the RPI.

- **Availability—**a much longer time series is available for RPI than for CPI or CPIH.[65] This may be important in a regulatory context for capturing inflation

[64] Oxera (2016), 'Indexation of future price controls in the water sector', 31 March. https://www.ofwat.gov.uk/publication/indexation-of-future-price-controls-in-the-water-sector/, last accessed 18 June 2019.
[65] The CPI and CPIH time series are only available from 1989, while the RPI time series is available since 1947.

rates over a longer historical period, which are needed to assess long-run returns on equity.

- **Consistency with public authorities**—given the move towards CPI/CPIH in the longer term, one should consider how RPI could be made more consistent with other national statistics, if RPI continues to be used.
- **Volatility**—the RPI is more volatile than the CPI and the CPIH because of its exposure to mortgage interest payments. Index volatility is undesirable because it makes forecasting more difficult. Some investors have RPI-linked liabilities, which makes volatility less of a concern.
- **Predictability**—the extent to which a given measure of inflation can be accurately forecast may affect its suitability for use within the price-setting framework.

While there are disadvantages to using the RPI, it has the key advantage of being available over a long historical period. We suggest a balanced approach that takes advantage of the long data series, while accommodating adjustments that mitigate the shortcomings of the RPI.

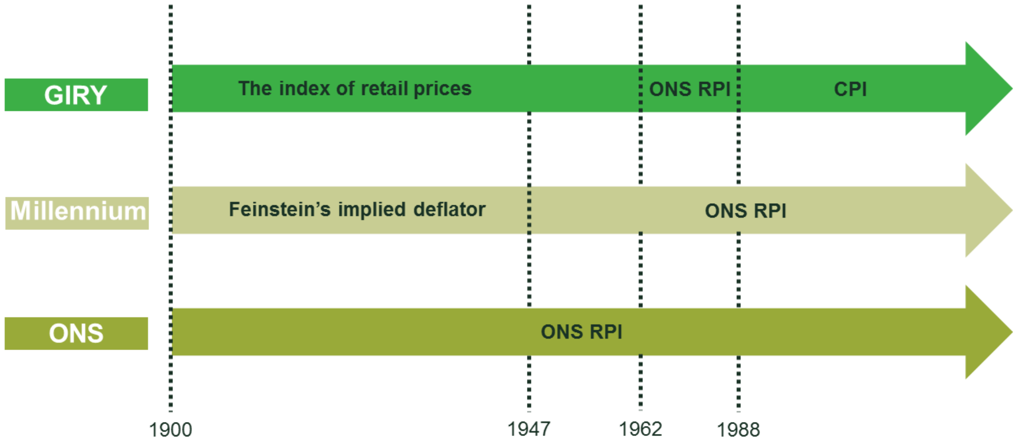
6 Potential alternative inflation measures

In this section, we consider potential adjustments to the RPI in order to address its shortcomings, as discussed in the earlier sections of this report. While there are disadvantages to using the RPI, it has the key advantage of being available over a long historical period. We suggest a balanced approach that takes advantage of the long data series and makes adjustments to restate the historical RPI in a consistent manner.

6.1 Differences between different data sources

The RPI series we use in this report is from the ONS and was last accessed in July 2019. The index produced by the ONS is agreed between the ONS, the Bank of England and the House of Commons Library.[66] The GIRY hybrid index is partially based on RPI (between 1962 and 1988) and CPI from 1988 onwards. Before 1962, the GIRY index is based on the index of retail prices.[67] The Bank of England's Millennium dataset RPI measure is based on work by O'Donoghue, Goulding and Allen (2004).[68] It uses 'Feinstein's implied deflator' before 1947 and the ONS RPI thereafter.[69] Figure 6.1 summarises the data sources used in each inflation measure over time.

Figure 6.1 Evolution of data sources in different inflation measures

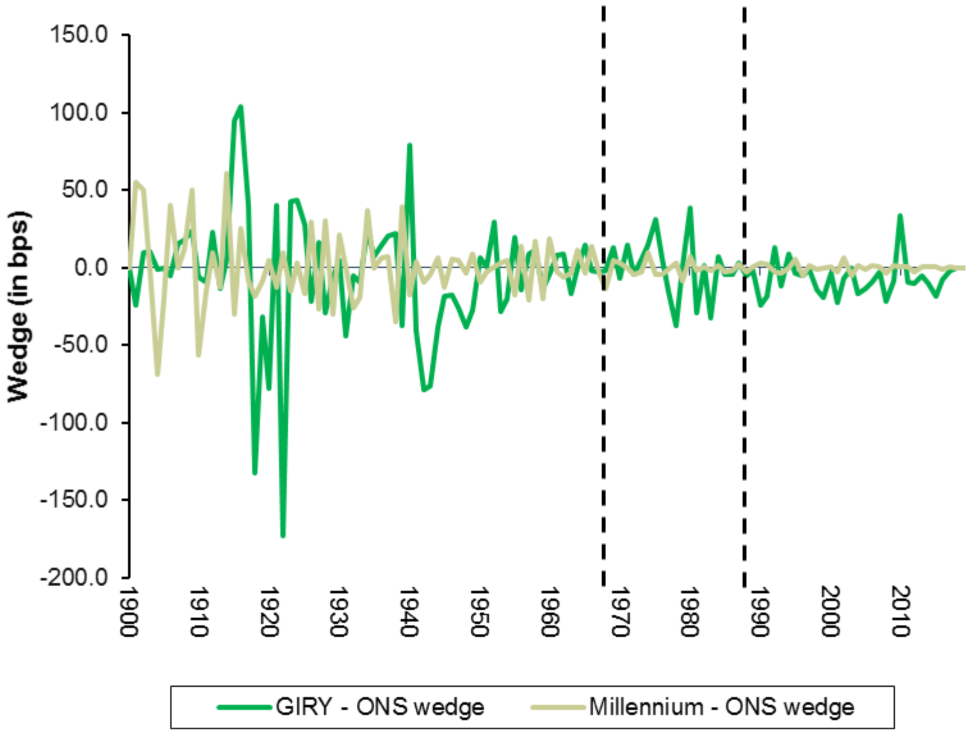


Source: Dimson, E., Marsh, P. and Staunton, M. (2017), 'Credit Suisse Global Investment Returns Yearbook 2017'; Bank of England's Millennium dataset; and Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

In Figure 6.2, we show the difference between the RPI published by the ONS and the GIRY hybrid index. We note that although these series are supposed to originate from similar sources for some periods, for instance between 1962 and 1988, they do not necessarily match. It not clear why these series differ from those published by the ONS. A potential explanation could be a statistical adjustment made to the GIRY and Millennium series for consistency with other inflation indicators used.

[66] House of Commons Library (2012), 'Inflation: the value of the pound 1750–2011', 29 May.
[67] We note that the index of retail prices was produced from 1956. It is not clear how the series prior to this date is estimated in GIRY.
[68] O'Donoghue, J., Goulding, L. and Allen, G. (2004), 'Consumer Price Inflation since 1750', ONS Economic Trends 604, March.
[69] Feinstein's implied factor is derived from estimates of consumers' expenditure valued at current and constant prices. See Feinstein, C.H. (1972), 'National income, expenditure and output of the United Kingdom, 1855-1965. Studies in the national income and expenditure of the United Kingdom', Cambridge University Press.

Figure 6.2 Difference between the RPI published by the ONS and other inflation measures



Note: The vertical dashed lines represent the period when these inflation measures are supposed to be identical.

Source: Dimson, E., Marsh, P. and Staunton, M. (2017), 'Credit Suisse Global Investment Returns Yearbook 2017'; Bank of England's Millennium dataset; and Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

6.2 Adjusting the RPI

An important feature of the RPI is that it is upward-biased because it uses the Carli rather than the Jevons index. The size of the wedge between the Carli and Jevons indices depends on the dispersion of relative prices. The greater the price dispersion, the greater the wedge between the Carli and Jevons formulae.

One alternative would therefore be to adjust the RPI in order to reduce the amount of price dispersion in the series. For example, one could:

- exclude price series that are very volatile; or
- mitigate the volatility of certain price series by excluding outlying price variations.

However, these methods would require having access to a very disaggregate level of information, which is not easily available outside the ONS.

In the context of backcasting inflation measures for price regulation purposes, one could adopt the methodology described in section 3.2. It is possible to build an adjusted RPI series that accounts for all methodological improvements that should have been made earlier but have not. For instance, holidays taken in the UK were only included in RPI from 1994 onwards; it is not the case that UK citizens did not take holidays in the UK in the years before 1994, but rather that this data was not collected. If the rate of inflation for UK holidays was lower than the reported RPI inflation, then failure to include holidays means that the RPI inflation was overstated before their inclusion.

We could therefore build an adjusted RPI by adding important elements that were not included in the past because of methodological limitations. We have examined two possible adjustment methods that do not rely on disaggregate data and present preliminary results on their application. More work would be needed to assess the validity of these approaches.

6.2.1 Method 1

The first method we propose involves estimating what weights and prices would have been associated with certain items before they were introduced. Once prices and weights have been backcasted, the adjusted RPI could be estimated by including the items that should have been part of the RPI. In the example of the addition of foreign holidays, the adjusted RPI would be:

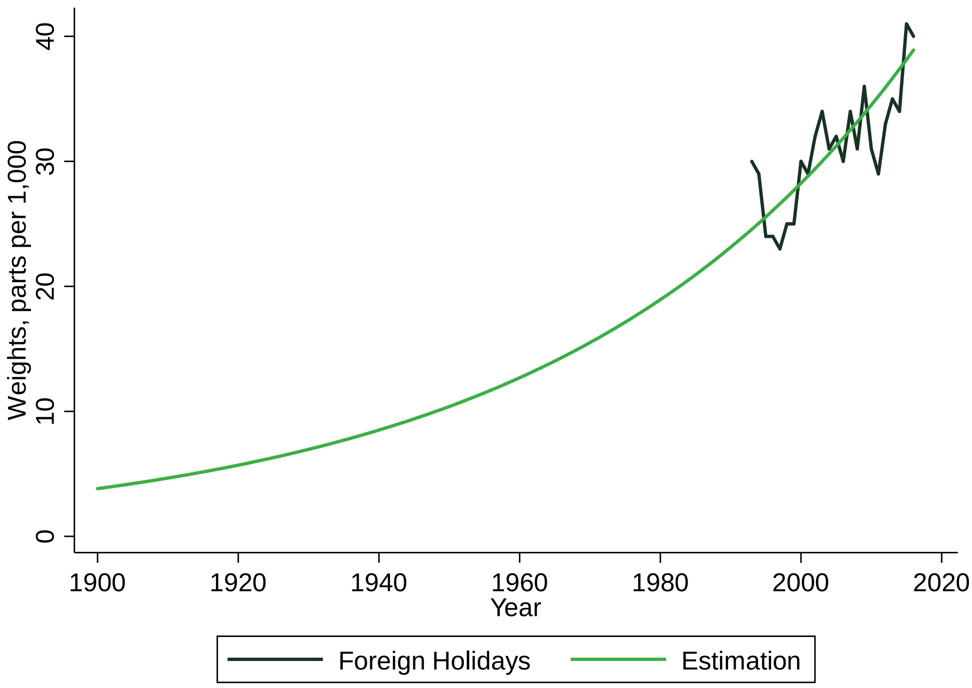
$$\begin{aligned} adjusted\ RPI &= (1 - backcasted\ foreign\ holidays\ weights) \times RPI\ index \\ &+ backcasted\ foreign\ holidays\ weights\ \times\ backcasted\ foreign\ holidays\ index \end{aligned}$$

To illustrate this approach, we undertake a simple method of backcasting the weights and prices of foreign holidays. We estimate a log model using annual RPI data over the period 1993–2018,[70] where

$$\log(foreign\ holiday\ weights) = \beta_0 + \beta_1 year + \epsilon_t$$

The predicted values are then exponentiated to obtain the backcasted weights. This is presented in Figure 6.1.

Figure 6.3 Backcasting of the weights on foreign holidays



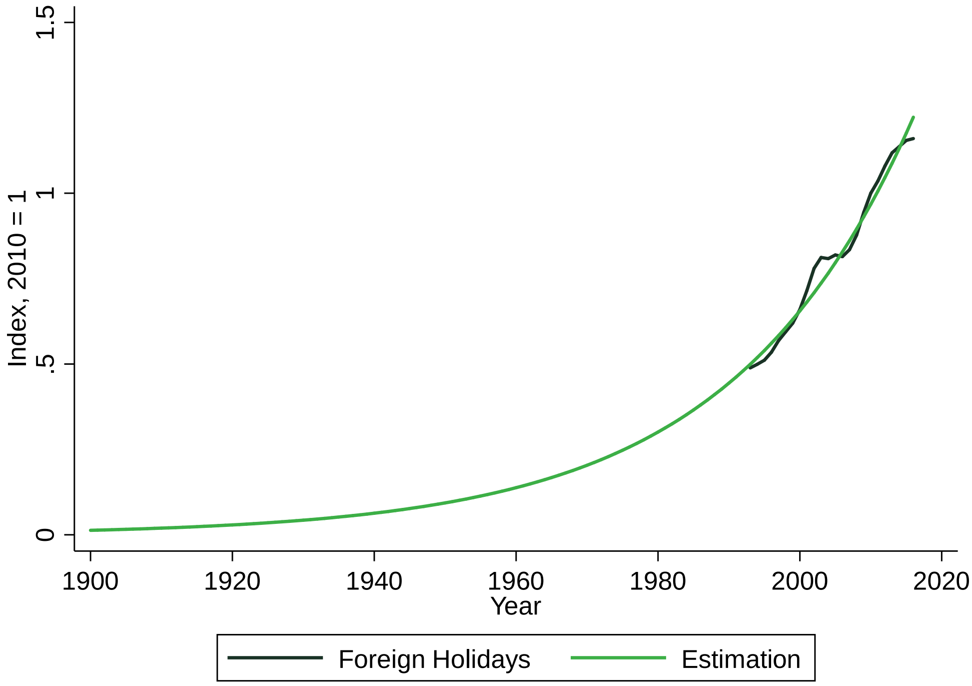
Source: Oxera analysis.

[70] Foreign holidays were introduced in 1993.

An implication is that the weight on foreign holidays in 1899 is small at just 3.74 parts per thousand, compared with 47 parts per thousand in 2018. This is plausible given lower real incomes and less connectivity in the early 20th century.

We adopt a similar methodology in forecasting the price of foreign holidays (see Figure 6.2).

Figure 6.4 Backcasting the price on foreign holidays



Source: Oxera analysis.

With these two components, we can estimate what RPI in 1899 would have been had foreign holidays been included in the basket of goods. Since the weights on foreign holidays in 1899 would have been very small, the effect that it has on the overall inflation rate is also small. The results are shown in Table 6.1.

Table 6.1 Inflation rates 1899–2016

	Index	Inflation rate
(1)	RPI All Items (ONS)	4.17%
(2)	RPI All Items, adjusted to include foreign holidays	4.16%

Source: Oxera analysis.

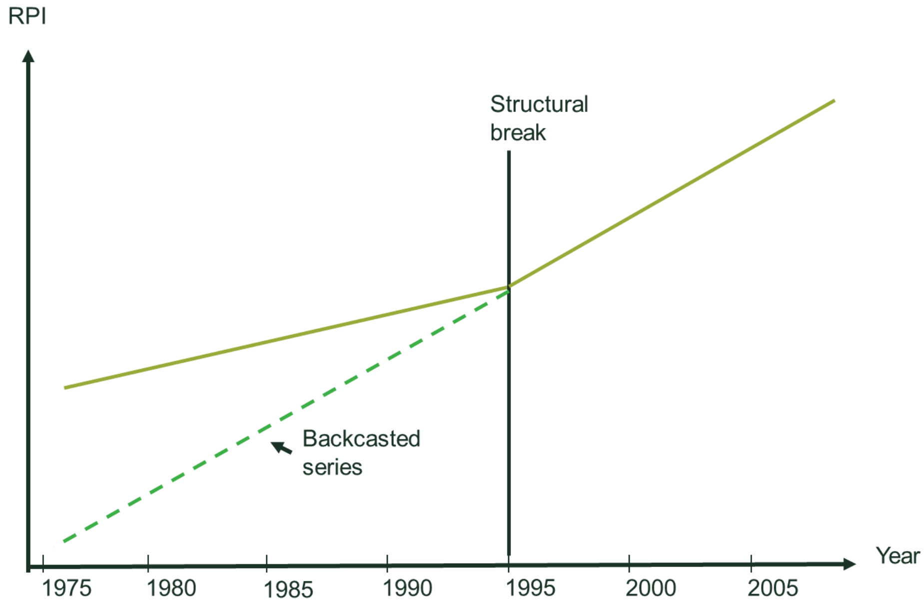
These results also align with our understanding that items are introduced into the RPI basket only when they form a significant proportion of expenditure. Given that these items were not included in the basket before, it is likely that they were not an important part of consumer spending, and so their contribution to RPI inflation was relatively small.

For completeness, this approach could be taken for a wider range of items than those that have been included in the RPI. That exercise is beyond the scope of this report.

6.2.2 Method 2

Another approach would be to use the results from the structural breaks test to build a counterfactual series. When a structural break is found by the algorithm described above, it is possible to estimate its magnitude. The coefficient associated with it can then be used to approximate what the RPI series would have been. This method is illustrated in Figure 6.5.

Figure 6.5 Stylised example on the use of structural breaks test in backcasting an adjusted RPI



Source: Oxera analysis.

We applied this methodology by backcasting the series, including structural breaks we identified as methodological. This means that some of the breaks will have positive effects, while others will have negative effects on backcasted inflation, as shown in Table A3.1. The impact of this estimation on inflation is presented in Table 6.2. The two sensitivities presented on Method 2 are based on the upper and lower bound of the structural breaks estimate.

Table 6.2 Adjustment of RPI by Method 2 from 1899 to 2016

	Index	Inflation rate
(1)	RPI All Items (ONS)	4.17%
(2)	Adjusted RPI—Method 2	4.33%
(3)	Adjusted RPI—Method 2, sensitivity 1	4.47%
(4)	Adjusted RPI—Method 2, sensitivity 2	4.20%

Source: Oxera analysis.

The limitation with this method is that it relies on a clear identification of whether the break is methodological or due to the economic activity. For example, it would not make sense to estimate an RPI series assuming that the oil shock in the 1970s would have happened earlier. Further work would be needed to identify the causes of each break we identified.

6.2.3 Updated estimates of historical real TMR

Based on the two preliminary methods developed above, we are able to compute alternative real TMR measures, as shown in Table 6.3.

Table 6.3 Adjusted real historical returns for the UK equity markets (%)

	Adjusted RPI method 1	Adjusted RPI method 2	Adjusted RPI method 2—sensitivity 1	Adjusted RPI method 2—sensitivity 2
Nominal TMR (arithmetic average)	11.20	11.20	11.20	11.20
Less: inflation	4.16	4.33	4.47	4.20
Inflation-adjusted TMR	**6.8**	**6.6**	**6.4**	**6.7**

Source: Oxera analysis; UK Regulators Network (2018), 'Estimating the cost of capital for implementation of price controls by UK Regulators', March, D–121, E–125; Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

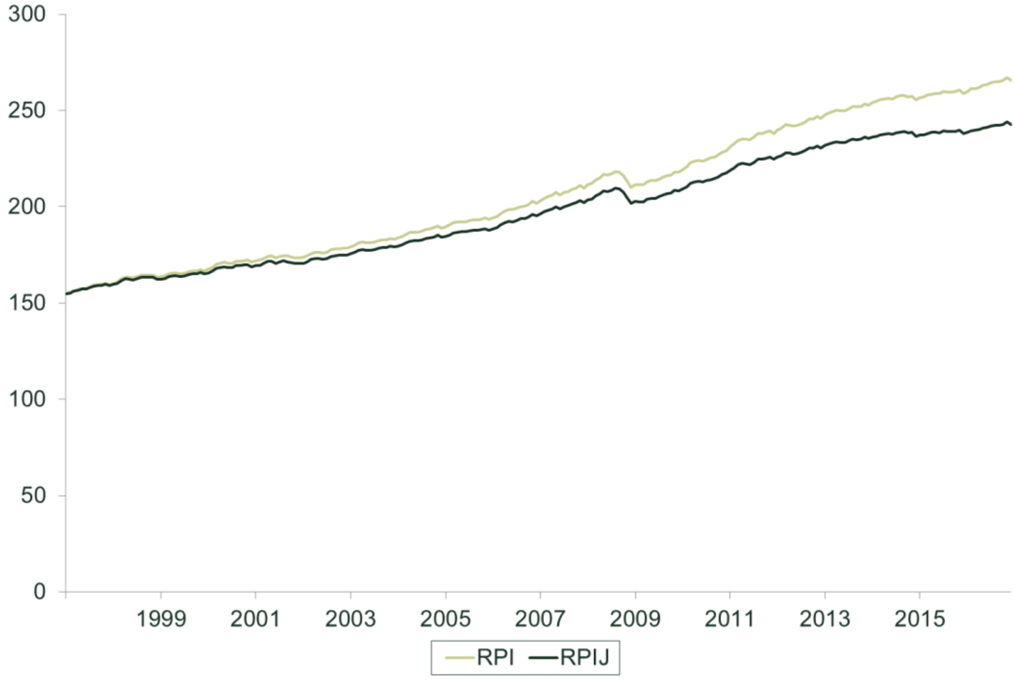
6.3 RPIJ and CPIH

A priori, the RPIJ and the CPI would be good candidates to replace or supplement the RPI.

The RPIJ overcomes some of the statistical issues with using RPI by replacing the use of the Carli formula with the Jevons formula. Besides this technical difference, it measures exactly the same thing as the RPI (with the same weights, same population, and same commodity coverage).[71] The RPIJ is a national statistic in the UK. As it is identical to the RPI apart from the use of the Jevons formula, both measures behave in a similar way, as depicted in Figure 6.6.

[71] Johnson (2015), op. cit., Figure 3.3.

Figure 6.6 RPI vs RPIJ



Source: Oxera analysis.

The CPI does not use the problematic Carli formula, but it differs from the RPI in its coverage.

However, both CPI and RPI are available for a limited amount of time, which means that there is limited historical information and fewer data points on which inflation could be forecasted. Both are key requirements for economic regulation. It means that simply replacing the RPI with the RPIJ or the CPI would not necessarily be the best option.

Nevertheless, the difference between the RPI and other indices could be used in the construction of a composite index or to backcast RPIJ in earlier years.

For example, a backcasting method has been designed to estimate pre-1989 CPI.[72] At a disaggregate level, ONS statisticians have matched the different CPI categories to the corresponding RPI categories during the period where both measures overlap. They have then adjusted the RPI categories to correct for difference in coverage, thereby allowing an estimation of the structural difference between RPI and CPI (the formula effect). This estimate is then used to infer what CPI would have been had it been computed before 1989.

This backcasting of the CPI introduces two additional sources of uncertainty, however. First, there can be significant differences between the definitions of different items in the RPI and the CPI. For example, the ONS acknowledges that there are considerable differences between the RPI and the CPI such as housing and fuel.[73] Second, because the formula effect is not observed but estimated, it adds a layer of uncertainty in the historical series that are backcasted.

A similar method could potentially be employed to estimate the difference between the RPIJ and RPI. However, to the best of our knowledge,

[72] Office for National Statistics (2013), 'Modelling a back series for the Consumer Price Index', September.
[73] O'Neill, Ralph and Smith (2017), op. cit., p. 4.

disaggregate information for the RPIJ is not available. A backcasting approach would therefore rely on more aggregate data.

7 Summary and recommendations

The RPI measure of inflation is used by the CAA to calculate allowed returns and prices for Heathrow Airport. The RPI series has undergone a series of methodology changes over time, and these have affected the basket of goods and the approach to calculating the index.

The revisions to the calculation of the RPI inflation statistic made by the ONS in 2010 created a structural increase in the RPI measure of inflation. All else equal, this would make the historical equity market returns deflated by historical RPI an upwardly biased estimate of the future TMR calculated relative to future RPI. However, there are likely to have been other revisions to the calculation of RPI during the history of the UK equity returns data set, some of which might have introduced a downward bias to average historical real equity market returns.

The CAA has proposed that to achieve consistency between forecast and historical measures of real equity market returns relative to RPI inflation, the forecast difference of 100bps between the RPI and CPI inflation metrics should be deducted from historical CPI-deflated equity returns.[74] The reliability of this approach rests on the accuracy of the 'backcast' CPI series (which was not published before 1989) and the premise that the average rate of inflation in the historical sample is approximately the same for the RPI and the CPI. The ONS has documented its reservations about the use of the backcast CPI series.[75]

There are a number of limitations with relying on a backcasted CPI series. The backcasted CPI is actually based on a restructuring of the RPI categories so that it matches the CPI categories as closely as possible. This is called the 'RPIA' (RPI-adjusted) series, which means that the backcasted CPI series is exposed to methodological changes in the RPI in the same way that the RPI is. Furthermore, some categories in the CPI and the RPI do not match closely,[76] resulting in further estimation errors. To arrive at the backcasted CPI, an estimate of the formula effect is subtracted from the RPIA, leading to another source of estimation error.

In the absence of a reliable estimate of the historical difference between RPI and CPI inflation, it is not robust to restate the historical RPI inflation series by adding the forecast difference of 100bps between RPI and CPI inflation to the historical RPI series.

The alternative to backcasting the CPI is to restate the historical RPI series by applying a methodology which is more consistent with how the RPI is calculated today. This report shows that in addition to the changes made by the ONS in 2010 there have been other revisions to the calculation of RPI during the history of the UK equity returns dataset, some of which may imply an upward bias to historical reported RPI inflation and hence a downward bias to average historical RPI-deflated equity market returns. While these changes in the RPI series might make RPI less attractive as a consistent measure of changes in consumer prices across the economy, RPI has an important advantage in that it is available for a much longer period of time than alternative measures, such as RPIJ or CPI. That additional period of time

[74] Ofgem (2018) 'RIIO-2 Sector Specific Methodology Annex: Finance', para. 3.62.
[75] The ONS has stated that its method of backcasting CPI 'provides only approximate results' and that 'there is no way to determine how accurate [the ONS'] method is as sufficient data to calculate the CPI do not exist prior to 1987.' The ONS emphasises that 'because of the assumptions made in [the backcasted CPI's] construction, these estimates are not National Statistics'. For further details, see Office for National Statistics (2013), 'Modelling a Back Series for the Consumer Price Index', 4 January, pp. 2–3.
[76] For example, fuel and light. See ibid., p. 4.

includes substantial structural changes to the UK economy, such as growth in international trade as well as significant shocks to the UK economy, such as two World Wars and recessions.

These advantages of the RPI—in addition to the disadvantages of backcasting other inflation measures, as highlighted above—mean that it is important to understand the impact of adjusting the historical RPI used to deflate TMR.

Recommendations

The ideal index would include a long time period to allow for a better estimate of TMR but also mitigate for the impacts of the challenges to the use of RPI identified in this report.

The overview of the history of the RPI and the structural break tests outlined in this report suggest that there are changes to the RPI series which, if appropriately adjusted for, could enable regulators to take advantage of the longer series, which includes low-probability but high-impact events, when estimating the TMR. We have investigated methodological changes and structural breaks in the series from both a statistical and methodological perspective. It allowed us to make some initial estimates of adjustments to the RPI series to create a series which is consistent over the longer period over which RPI is available.

Adjusting the reported RPI index will never result in a perfect index, but making reasonable adjustments based on testing the magnitude of the effect using data on both the level and the trend of RPI could help regulators improve the estimation of allowed returns and allowed prices for consumers in regulated markets. We have suggested two possible ways to achieve this, and have undertaken some initial analysis of the impact of these approaches, both of which result in adjustments to the average historical RPI inflation rate to make it comparable with the current definition of RPI. The adjustments result in changes between -1bp or +30bps, respectively, to the average historical ONS RPI inflation and equivalent opposite adjustments to the average historical equity market returns measured relative to RPI (i.e. +1bp or -30bps).

The historical annual arithmetic average equity market return for the period 1899–2016, deflated using these adjusted RPI series, is between 6.4% and 6.8%. This is preliminary analysis, and these results are only indicative of the size of adjustment that could be required. Nevertheless, the range based on the adjusted historical RPI series is 40–80bps higher than the 6% that the CAA presents as the arithmetic average equity market return based on backcasting the CPI.[77]

This is a significant difference between real equity returns calculated by directly adjusting the historical RPI compared with backcasting CPI and then deducting the forecast differential between RPI and CPI inflation. This implies that placing weight on the indirect approach to restating historical equity returns relative to RPI is likely to bias downward the estimated of the RPI-deflated expected total market return.

[77] Civil Aviation Authority (2019), ‘Appendices to Draft UK Reference Period 3 Performance Plan proposals’, February, D-33.

A1 Elementary aggregate formula

At the lowest level of aggregation in consumer price statistics (called the 'elementary aggregate' level), there is often a lack of reliable expenditure data to weight products together. In this appendix, we consider elementary aggregates as a description of a level of price index for which no quantity information is available. For example, we may know how much is spent on apples in the UK, but we do not know how much is spent on different types of apples such as Royal Gala, Braeburn and Golden Delicious. In these circumstances, elementary aggregate formulae are applied.[78]

There are three main averaging techniques. Below, we explain their properties and how they are constructed. These techniques are important for understanding how RPI is constructed and how it compares with alternatives such as CPI and RPIJ.

A1.1 Dutot

The Dutot formula takes the arithmetic average of prices in each period, then calculates the rate of change. It is the ratio of average prices.

Consider a set of N goods and services indexed $i \in \{1, \dots, N\}$, which has a price of $p_{i,t}$ in period t. Then, the Dutot index is given by:

$$I_{t,0}^{D} = \frac{\frac{1}{N}\sum_{i=1}^{N} p_{i,t}}{\frac{1}{N}\sum_{i=1}^{N} p_{i,0}}$$

Although there is no explicit weighting, the Dutot index can also be re-written as:

$$\sum_{i=1}^{N} \frac{p_{i,0}}{\sum_{j=1}^{N} p_{j,0}} \frac{p_{i,t}}{p_{i,0}}$$

We can see that there is an implicit weighting scheme to the Dutot index, which places the greatest weight on the items that were relatively more expensive in the first period covered by the price index.[79] The Dutot index is used for homogeneous items as the formula implicitly gives the greatest weight to the highest-priced product.[80]

A1.2 Carli

The Carli formula takes the rate of change in each price, then takes the arithmetic average of those changes. It is the average of price relatives. The Carli index is given by:

$$I_{t,0}^{C} = \frac{1}{N}\sum_{i=1}^{N} \frac{p_{i,t}}{p_{i,0}}$$

The Carli formula has some unintuitive properties. The Carli index exhibits chain drift. An index exhibits chain drift if the index of period 2 relative to period

[78] UK Statistics Authority (2016), 'Elementary aggregate formula', https://www.statisticsauthority.gov.uk/wp-content/uploads/2016/04/Elementary-aggregate-formula-description-action-2.pdf, last accessed 2 June 2019.
[79] O'Neill, Ralph and Smith (2017), op. cit., p. 294.
[80] UK Statistics Authority (2016), op. cit.

0 is not equal to the product of the index of period 1 relative to period 0 and the index of period 2 relative to period 1. That is, the chain drift is given by:

$$I_{2,0} - I_{1,0} * I_{2,1}$$

A1.3 Jevons

The Jevons formula uses the geometric mean of prices. The Jevons index is given by:

$$I_{t,0}^{D} = \left(\prod_{i=1}^{N} \frac{p_{i,t}}{p_{i,0}}\right)^{\frac{1}{N}}$$

A1.4 A numerical example

Elementary aggregates measure the evolution of prices of a good at the most disaggregate level. They are based on ‘price relatives’, which represent the price of a good relative to a base period. For instance, if the price of Cameo apples was £0.8 in 2009 and £1 in 2010, and if the base year is 2010, the price relative is 1.25 (=1/0.8).

This is illustrated in Table A1.1. In 2010, the prices of all the apples have increased by a small amount compared with 2009. Since the amount of price dispersion is still relatively low, the three indices give similar results. In 2011, the price differences between each of the items are substantially larger.[81] In this example, where price dispersion increases, the Carli index yields a higher measure of inflation than the Dutot or Jevons indices.

Table A1.1 Effect of price dispersion on when calculating indices using different formulae

Apples	2009	2010	2011
Prices			
Cameo	0.80	0.82	2.00
Braeburn	0.80	0.87	1.00
Bramley	0.80	0.90	1.00
Cox	0.80	0.92	0.50
Price relatives			
Cameo		1.03	2.44
Braeburn		1.09	1.15
Bramley		1.13	1.11
Cox		1.15	0.54
Price index			
Carli		1.10	1.31
Dutot		1.10	1.28
Jevons		1.10	1.14

Source: Oxera analysis.

[81] In mathematical terms, this creates variance in the price relatives.

A1.5 Relationship between elementary aggregates

If the Jevons and Carli indices for the price change between time 0 and time t are I_{Jevons}^{0t} and I_{Carli}^{0t}, the Jevons index can be written as:[82]

$$I_{Jevons}^{0t} \approx I_{Carli}^{0t} - \frac{\sigma_R^2}{2I_{Carli}^{0t}}$$

We can interpret the equation intuitively as follows: the difference between the Jevons and Carli indices is determined by the variance of price relatives. Therefore, increased price dispersion increases the difference between the two indices.

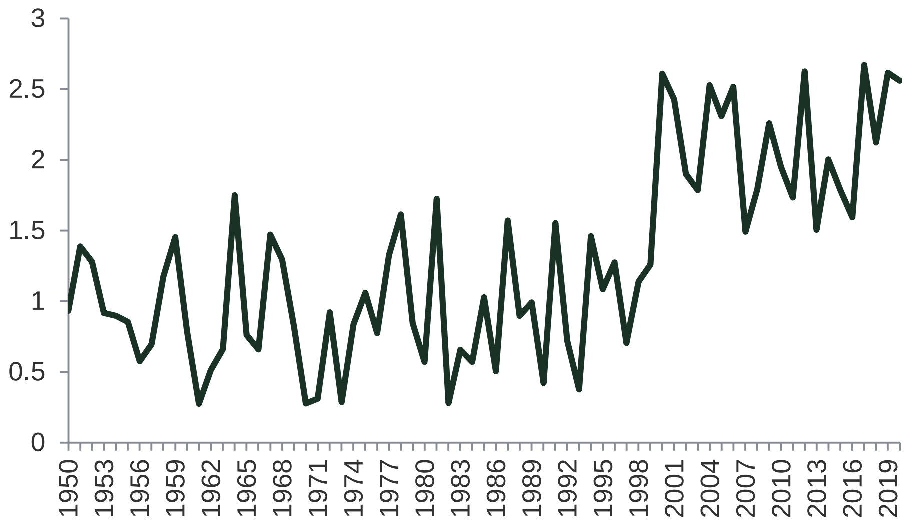
[82] The Taylor expansion of the Jevons index about the value of the Carli index is taken. See O'Neill, Ralph and Smith (2017), op. cit., p. 251.

A2 Step-indicator saturation

Time series data is a sequence of a single variable over time. Examples include GDP, oil prices and, in our case, the RPI. A statistical model of time series is essentially a mathematical description of the variable. For example, we may say that the variable takes, on average, a particular value, or grows at a constant rate.

Consider a hypothetical variable X_t, and suppose we have data from 1950–2020.

Figure A2.1 Oxera's hypothetical time-series variable



Source: Oxera analysis.

A simple time-series model of the data may be that the data, on average, takes the value μ over the entire period:

$$X_t = \mu + \epsilon_t$$

where ϵ_t is a noise term.[83] Various statistical and econometric techniques can be used to estimate the parameter μ.

Often, time-series data is not so straightforward. For example, in the RPI example, a change of economic regime, such as an oil price shock, or the introduction of a new and important item in the RPI may lead to the behaviour of RPI to be systematically different from its past behaviour. This is called a **structural break**.[84]

In the hypothetical example, it appears that before 2000, the data is consistently lower than the data after 2000, i.e. there is a structural break in the model in 2000. The data analyst may then formulate a different model:

$$X_t = \mu + \delta \cdot 1(year > 2000) + \epsilon_t$$

This model allows the mean of the data to be systematically different before and after the year 2000. A test of whether there is a structural break would be

[83] In the example, it is an identically and independently distributed uniform random variable that has support [-0.75, +0.75].
[84] This may involve a change in the mean, or any other parameters of the process that produce the series.

to test whether the estimated δ is statistically different from zero. This is known as a Chow test, and is the classical way to allow and test for structural breaks in data.

A disadvantage of this test is that it requires the data analyst to know the exact dates of the structural break. Often, we do not know *a priori* what the structural breaks are.

In the context of this report, while there are a number of methodological changes in the RPI, we do not know which are significant enough to cause a structural break in the RPI series.

Step-indicator saturation ('SIS') is a machine learning technique that can identify and estimate structural breaks without prior identification of the dates of the structural breaks. This technique works by considering a model that allows for a full set of structural breaks. That is, all possible breaks are added, thereby 'saturating' the model:

$$X_t = \mu + \sum_{j=1951}^{2020} \delta_j 1_{\{year \leq j\}} + \epsilon_t$$

This model allows for a structural break to occur in every year of the data. However, estimating this model will result in every data point being fitted exactly. Castle, Doornik, Hendry and Pretis (2015)[85] explore a split-half approach to SIS. This consists of the following algorithm:

1. Add the first half of the indicators $1_{\{year\leq 1951\}}, \dots, 1_{\{year\leq 1985\}}$ and select those which are statistically significant at the α significance level.
2. Add the second half of the indicators to the original model $1_{\{year\leq 1986\}}, \dots, 1_{\{\leq 2020\}}$ and select those which are statistically significant at the α significance level.
3. Finally, combine the statistically significant variables from the two stages.

The authors undertake simulation studies to explore the properties of SIS. They consider its propensity to retain *irrelevant* indicators (called the 'gauge') and its propensity to retain *relevant* indicators (called the 'potency').[86] The authors find that under various scenarios, SIS performs well. For further details, we refer the reader to the paper.

This method can also be extended beyond step-shifts to trend shifts, which we use in this report. The specification would then be:

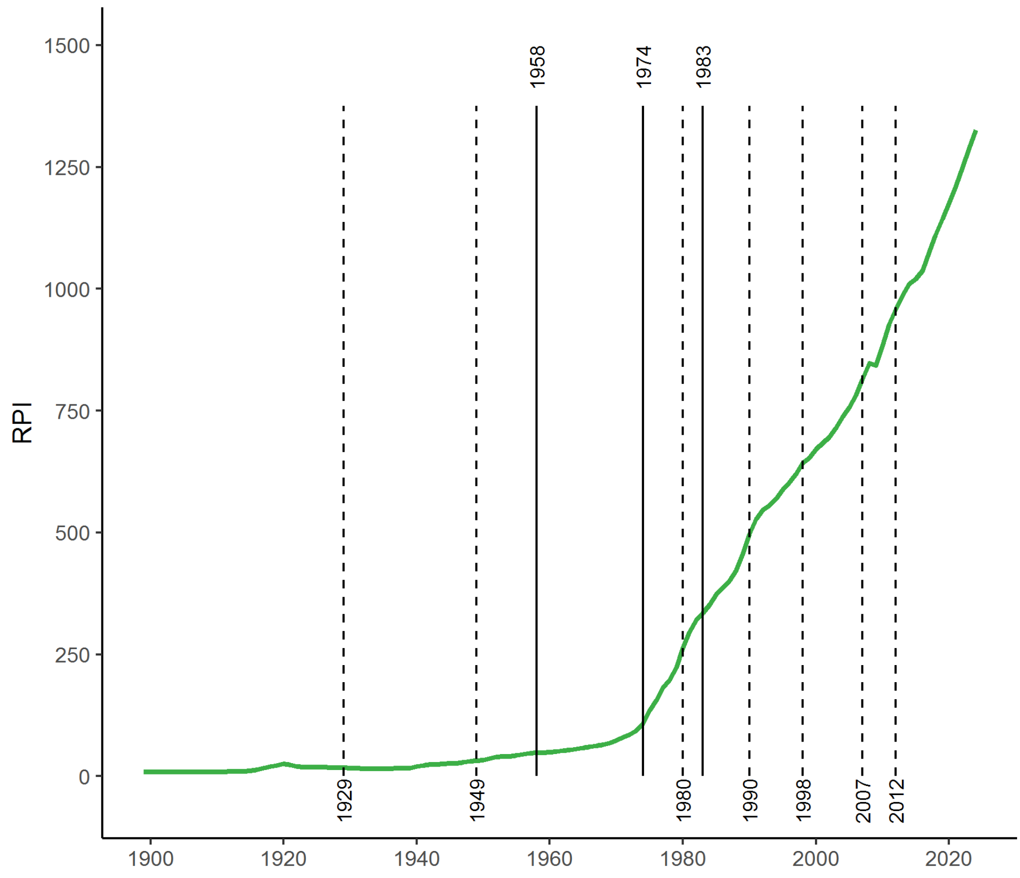
$$X_t = \mu + \sum_{j=1951}^{2020} \delta_j 1_{\{year \leq j\}}(2020 - j) + \epsilon_t$$

[85] Castle, J., Doornik, J., Hendry, D. and Pretis, F. (2015), 'Detecting location shifts during model selection by step-indicator saturation', *Econometrics*, **3**:2, pp. 240–264.
[86] For a further discussion of these ideas, see Castle, J.L., Doornik, J.A. and Hendry, D.F. (2011), 'Evaluating automatic model selection', *Journal of Time Series Econometrics*, **3**:1.

A3 Structural break tests

A3.1 Structural breaks of an extended price series

Figure A3.1 Structural breaks of an extended RPI series



Source: Oxera analysis; Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

A3.2 Logs

Table A3.1 Sensitivity, structural breaks in log RPI series

Year	Effect	Potential explanation
1970	Trend increase	• UK macroeconomic policies resulting in stagflation • introduction of dining out
1975	Level increase	• 1973 oil crisis • switch from equivalent rents to mortgage interest payments as a measure of owner-occupiers' housing costs
1976	Trend increase	• 1973 oil crisis
1980	Trend decrease; level increase	• Margaret Thatcher's policy to promote low inflation
1992	Trend decrease	• recession arising from US savings and loan crisis • introduction of domestic and foreign holidays

Source: Oxera analysis; and Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

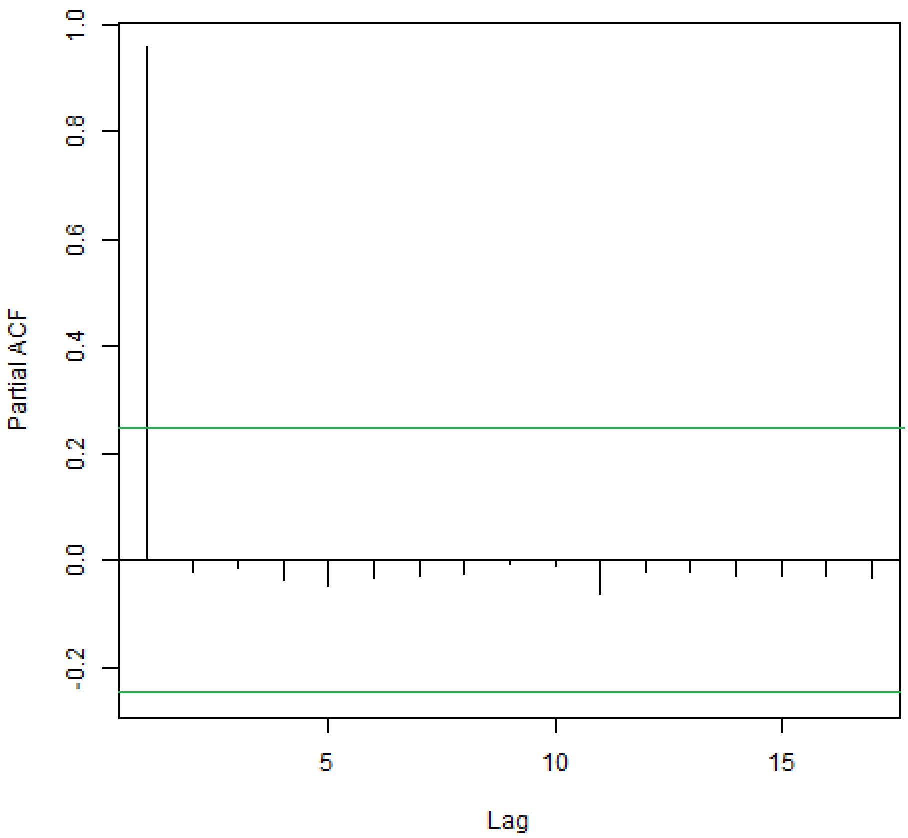
A3.3 Breakpoints

A3.4 Controlling for the persistence in inflation

The structural breaks analysis presented in section 3.3 allows for shifts in both the level and the trends of RPI. In the economic analysis of time series, it is common to further allow for 'persistence' over time.[87] We have not allowed for this in our base specifications because the fitted model is already complex with multiple structural breaks. Allowing for further complexity increases the risk of over-fitting the data rather than capturing true trends.

Nonetheless, there are good economic reasons to allow for persistence. For example, there may be menu costs or sticky wages, which prevent prices from adjusting quickly to their steady-state levels. Figure A3.2 shows the partial autocorrelation function of the RPI series. There is evidence to suggest that RPI has a strong dependence on the values in the preceding year (although this may be due to the time trend).

Figure A3.2 Partial autocorrelation function (PACF) of the RPI series

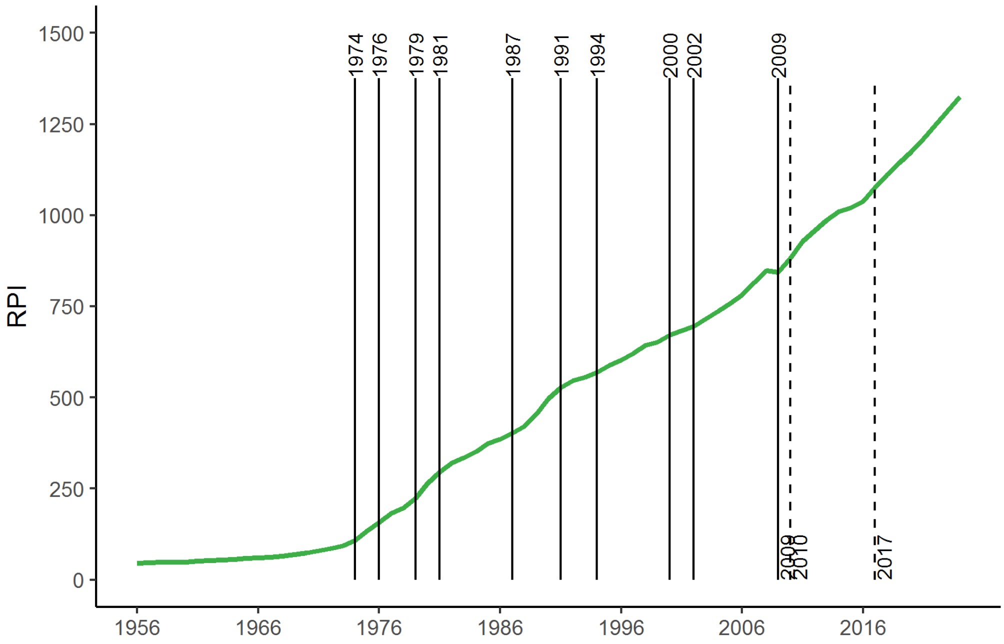


Source: Oxera analysis of ONS RPI data.

We therefore produce sensitivities for the structural breaks analysis including a first order autoregressive term. The results are broadly similar, although there are notably more trend breaks in the levels of RPI series.

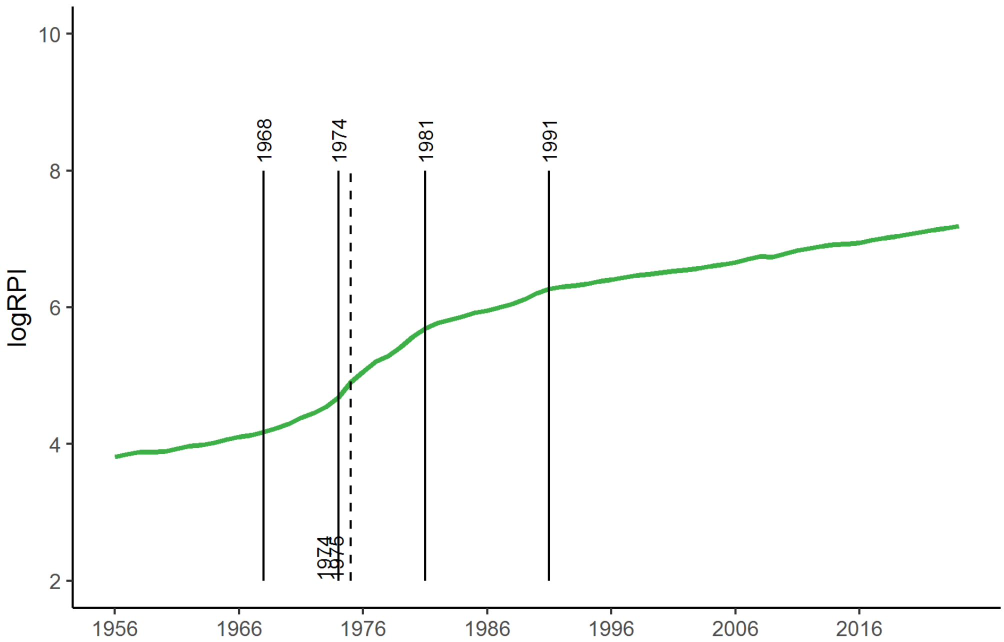
[87] Persistence essentially means that if the level of RPI is high in the period before, it may be high in the current period. This is econometrically implemented by including the lag of RPI in the regression function.

Figure A3.3 Structural breaks analysis with auto-regressive term, in levels



Source: Oxera analysis; Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

Figure A3.4 Structural breaks analysis with auto-regressive term, in logs



Source: Oxera analysis; Office for National Statistics, 'Retail Prices Index: Long run series: 1947 to 2019' (with annual series starting from 1800), retrieved from https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cdko/mm23.

A3.5 Structural breaks analysis for Method 2 adjustment

Table A3.2 Structural breaks in RPI series from 1900 (in RPI points)

Structural break	Effect	Magnitude
1974	Trend increase	+21.15
1992	Trend decrease	-25.20
1997	Trend increase	+8.23
2011	Level increase	+55.55

Note: This is based on the model presented in section 3.3. We only present the breaks that we classify as potentially methodological and used in Method 2.

www.oxera.com

oxera
compelling economics