Occasional Paper Series

Work stream on inflation measurement

Inflation measurement and its assessment in the ECB's monetary policy strategy review

ECB
STRATEGY
REVIEW

Workstream leads

Christiane Nickel
European Central Bank
email: christiane.nickel@ecb.europa.eu

Annette Fröhling
Deutsche Bundesbank
email: annette.froehling@bundesbank.de

Luis J. Álvarez
Banco de España
email: ljalv@bde.es

Caroline Willeke
European Central Bank
email: caroline.willeke@ecb.europa.eu

Giordano Zevi
Banca d'Italia
email: giordano.zevi@bancaditalia.it

Sub-group leads

Luis J. Álvarez
Banco de España
email: ljalv@bde.es

Chiara Osbat
European Central Bank
email: Chiara.Osbat@ecb.europa.eu

Annette Fröhling
Deutsche Bundesbank
email: annette.froehling@bundesbank.de

Giordano Zevi
Banca d'Italia
email: giordano.zevi@bancaditalia.it

Ioannis Ganoulis
European Central Bank
email: ioannis.ganoulis@ecb.europa.eu

Coordinators

Gerrit Koester
European Central Bank
email: gerrit.koester@ecb.europa.eu

Romana Peronaci
European Central Bank
email: romana.peronaci@ecb.europa.eu

Eliza Lis
European Central Bank
email: eliza.lis@ecb.europa.eu

Workstream contributors

Luis J. Álvarez
Banco de España

Elke Hahn
European Central Bank

Jan Beka
National Bank of Slovakia

Lukas Henkel
European Central Bank

Cristina Conflitti
Banca d'Italia

Julika Herzberg
Bundesbank

James Costain
Banco de España

Marco Hoeberichts
De Nederlandsche Bank

Martin Eiglsperger
European Central Bank

Jana Jonckheere
National Bank of Belgium

Annette Fröhling
Deutsche Bundesbank

Yannick Kalantzis
Banque de France

Ioannis Ganoulis
European Central Bank

Demetris Kapatais
Central Bank of Cyprus

Erwan Gautier
Banque de France

Thomas Knetsch
Deutsche Bundesbank

Bernhard Goldhammer
European Central Bank

Gerrit Koester
European Central Bank

Evripides Kondelis
Bank of Greece

Omiros Kouvavas
European Central Bank

Olegs Krasnopjorovs
Bank of Latvia

Eliza Lis
European Central Bank

Patrick Lünnemann
Central Bank of Luxembourg

Fernando Martins
Banco de Portugal

Teresa Messner
OeNB

Laura Mociunaite
Bank of Lithuania

Christiane Nickel
European Central Bank

Derry O'Brien
European Central Bank

Chiara Osbat
European Central Bank

Romana Peronaci
European Central Bank

Dilyana Popova
Bundesbank

Mario Porqueddu
European Central Bank

Moreno Roma
European Central Bank

Fabio Rumler
OeNB

Patrick Schwind
Deutsche Bundesbank

Nika Sosič
Bank of Slovenia

Georg Strasser
European Central Bank

Riccardo Trezzi
European Central Bank

Fons van Overbeek
De Nederlandsche Bank

Lauri Vilmi
Bank of Finland

Aurelian Vlad
European Central Bank

Joris Wauters
National Bank of Belgium

Sebastian Weinand
Deutsche Bundesbank

Thomas Westermann
European Central Bank

Caroline Willeke
European Central Bank

Ladislav Wintr
Central Bank of Luxembourg

Zivile Zekaite
Central Bank of Ireland

Giordano Zevi
Banca d'Italia

This report is part of a set of papers within the ECB's Occasional Paper Series, related to the ECB's Strategy review 2020-21. This set includes the following papers:

Set of Occasional Papers related to the ECB's Strategy review 2020-21

No 263, "The implications of globalisation for the ECB monetary policy strategy".
No 264, "Inflation expectations and their role in Eurosystem forecasting".
No 265, "Inflation measurement and its assessment in the ECB's monetary policy strategy review".
No 266, "Digitalisation: channels, impacts and implications for monetary policy in the euro area".
No 267, "Review of macroeconomic modelling in the Eurosystem: current practices and scope for improvement".
No 268, "Key factors behind productivity trends in EU countries".
No 269, "The ECB's price stability framework: past experience, and current and future challenges".
No 270, "Non-bank financial intermediation in the euro area: implications for monetary policy transmission and key vulnerabilities".
No 271, "Climate change and monetary policy in the euro area".
No 272, "The role of financial stability considerations in monetary policy and the interaction with macroprudential policy in the euro area".
No 273, "Monetary-fiscal policy interactions in the euro area".
No 274, "Clear, consistent and engaging: ECB monetary policy communication in a changing world".
No 275, "Employment and the conduct of monetary policy in the euro area".
No 276, "The mandate of the ECB: Legal considerations in the ECB's monetary policy strategy review".
No 277, "Evolution of the ECB's analytical framework".
No 278, "Assessing the efficacy, efficiency and potential side effects of the ECB's monetary policy instruments since 2014".
No 279, "The need for an inflation buffer in the ECB's price stability objective – the role of nominal rigidities and inflation differentials".
No 280, "Understanding low inflation in the euro area from 2013 to 2019: cyclical and structural drivers".

Contents

Abstract

This paper – which takes into consideration overall experience with the Harmonised Index of Consumer Prices (HICP) as well as the improvements made to this measure of inflation since 2003 – finds that the HICP continues to fulfil the prerequisites for the index underlying the ECB's definition of price stability. Nonetheless, there is scope for enhancing the HICP, especially by including owner-occupied housing (OOH) using the net acquisitions approach. Filling this long-standing gap is of utmost importance to increase the coverage and cross-country comparability of the HICP. In addition to integrating OOH into the HICP, further improvements would be welcome in harmonisation, especially regarding the treatment of product replacement and quality adjustment. Such measures may also help reduce the measurement bias that still exists in the HICP. Overall, a knowledge gap concerning the exact size of the measurement bias of the HICP remains, which calls for further research. More generally, the paper also finds that auxiliary inflation measures can play an important role in the ECB's economic and monetary analyses. This applies not only to analytical series including OOH, but also to measures of underlying inflation or a cost of living index.

Executive summary

The choice of a suitable framework for measuring inflation is a key element of any price stability-oriented monetary policy strategy.

Since it was first established in 1998, the European Central Bank (ECB) has used the Harmonised Index of Consumer Prices (HICP) to formulate its quantitative definition of price stability and to assess whether the outlook for inflation (in the medium term) is in line with this definition.

As part of the current review of the ECB's monetary policy strategy, the Monetary Policy Committee (MPC) Work Stream on Inflation Measurement was tasked with re-assessing the appropriateness of the HICP for measuring price stability and highlighting any need for improvement.

Since the last review of monetary policy strategy in 2003, the quality of the HICP as a measure of price stability in the euro area has been continuously improved by Eurostat and the statistical offices of European Union (EU) Member States.

This report – which takes into consideration overall experience with the HICP as well as the improvements made to it since 2003 – finds that the HICP continues to fulfil the prerequisites for the index underlying the ECB's definition of price stability.

Notwithstanding past improvements, the work stream sees scope for enhancing the HICP – especially by including owner-occupied housing (OOH). There appears to be a valid case for including OOH using the net acquisitions (NA) approach. A request to this effect will be submitted to the European Commission. In line with the allocation of responsibilities in the area of European statistics, implementing an official "HICP-H" index (i.e. an index which combines the basket of the HICP with expenditure on OOH) falls within the remit of Eurostat and the national statistical institutes (NSIs) and would be based on an enhanced European Parliament and European Council legal act. Implementation would inevitably be on a step-by-step basis and might take some time to complete.

In addition to integrating OOH into the HICP, further improvements are needed in harmonisation, especially regarding the treatment of product replacement, quality adjustment and sampling issues. Such measures may also help reduce the measurement bias that still exists in the HICP. There is no clear evidence that this bias has substantially declined since the last strategy review; the effects of improvements to the HICP have been outweighed to a certain extent by counter-directional effects related to economic developments such as e-commerce. Overall, a knowledge gap concerning the exact size of the measurement bias of the HICP remains, which calls for further research.

More generally, the report also finds that auxiliary inflation measures can play an important role in the ECB's economic and monetary analyses. This applies not only

to analytical series of OOH-augmented HICPs, but also to measures of underlying inflation or a cost of living index (COLI).

Assigning an important role to auxiliary inflation measures and announcing new measures such as an HICP-H has major implications for the ECB's communications. Careful communication will be required to avoid confusion about the ECB's quantitative measure of price stability, especially in the long implementation period during which OOH is integrated into the HICP.

1   Introduction

The choice of a suitable framework for measuring inflation is a key element of any price stability-oriented monetary policy strategy. This is particularly important when a quantitative definition of price stability is set against which the public can hold the central bank accountable. Central criteria for the suitability of an inflation measure for monetary policy purposes[1] are credibility[2], reliability, comparability over time and timeliness. In the case of the ECB and the Economic and Monetary Union cross-country comparability is a further key requirement, not least because price stability is also a convergence criterion for accession to the latter.

The HICP was selected in 1998, and confirmed in the 2003 monetary policy strategy review, as the best available measure to quantify the price stability objective for the euro area.[3] The ECB's definition of price stability was explicitly formulated in terms of the overall HICP. Other concepts, such as underlying inflation measures that exclude certain volatile price components, were given a role in monitoring medium-term price developments. However, they were explicitly ruled out as a suitable yardstick for measuring price stability in the euro area as they would exclude important items of households' consumption baskets.

The 2003 review of the ECB's monetary strategy identified some areas for improvement.[4] For example, the coverage of OOH was seen as an important element to improve the representativity of the HICP[5] and its cross-country comparability. A positive measurement bias in the HICP was also noted.

In the context of the current review of the ECB's monetary policy strategy, the MPC Work Stream on Inflation Measurement re-assessed whether the reasons for choosing the HICP still hold and to what extent the HICP continues to meet the criteria required of a benchmark for the assessment of price stability. A key issue in the work stream's agenda is the incomplete role of OOH in the inflation measure. Since the creation of the HICP, the ECB has repeatedly underlined the importance of including the cost of OOH in the HICP, while also acknowledging the practical and conceptual challenges that would entail. The ECB has continuously stressed that the inclusion of OOH should not compromise the quality criteria established for the HICP.[6] Recently, and not least in the context of a number of "ECB Listens"[7] events, it has become more apparent that the public seems to perceive the lack of a significant housing cost component in the headline price index

---

[1]   As also applied in the 2003 strategy review; see Issing (2003), pp. 12-13.

[2]   Credibility refers here to the perception by the general public that the choice of the measure is suited to providing full and effective protection against losses in the purchasing power of money.

[3]   Box 1 of Chapter 2 sets out a number of inflation concepts that play a role in the assessment that the HICP remains the preferred measure for defining price stability.

[4]   See Issing (2003), pp. 35-38.

[5]   See the letter of 14 July 2005 from ECB President Jean-Claude Trichet to the European Commissioner for Economic and Monetary Affairs, Joaquín Almunia, on the statistical features of the HICP.

[6]   See, for example, the letter of 12 June 2018 from ECB President Mario Draghi to MEP Sander Loones.

[7]   See the ECB Listens Portal.

as impairing the representativeness of the HICP. The work stream also assessed the methodological advances that have been made in the European Statistical System (ESS) over time to address various sources of potential measurement bias. Finally, the work stream revisited the role of measures of underlying inflation and the possible usefulness of a COLI for monetary policy analysis. For all of these elements, the work stream developed evaluations as well as proposals for improvements and communication. This also holds for recent experience with the impact of the coronavirus (COVID-19) pandemic on changes in consumption habits and what this implies for price measurement in the euro area.

The MPC report was prepared by five subgroups, each contributing one chapter. The ECB's price stability concept is discussed in Chapter 2. This is followed in Chapter 3 by an assessment of measurement bias in the HICP that also reviews measurement uncertainty. Chapter 4 assesses the role of OOH and options to improve the HICP in order to better reflect the development of price developments for housing. Chapter 5 focuses on measuring the cost of living, while Chapter 6 discusses the role of underlying inflation. Chapter 7 sets out the conclusions.

2 The ECB's price stability concept

At the start of Economic and Monetary Union, the HICP was chosen as the ECB's measure of price stability in view of its favourable characteristics. Since then, the HICP has been continuously improved to address new and existing challenges.

This notwithstanding, several HICP measurement challenges remain, some of them already highlighted in the last strategy review in 2003. Since that review, various dimensions of the HICP have been further improved, leading in particular to increased harmonisation between countries and a reduction of the substitution bias. However, and most importantly, the integration of OOH into the HICP remains outstanding. There is also room for further harmonisation in sampling and quality adjustment methods, while quantifying the overall measurement bias in the euro area remains a difficult issue.

New challenges have also emerged, including, for example, the availability of new data such as scanner and web-scraped data, as a result of ongoing digitalisation, and new shocks and sudden shifts that have affected the consumption basket, such as those related to the COVID-19 pandemic. These challenges emphasise the need for more timely information about consumption structures for the annual update of HICP weights. The need for more transparency about the inclusion of new data and new methods in the HICP is another ongoing issue. Greater flexibility is also needed in chain-linking, where applicable. This chapter explains how these new challenges can affect inflation analysis and projection and how increased transparency in the ESS might help reduce the uncertainty of price measurement, thereby supporting a credible monetary policy. Finally, this chapter briefly explains the link between measurement issues and the optimal definition of the price stability objective.

2.1 The HICP: characteristics, previous assessment and credibility issues

According to the mandate given by the Treaty on the Functioning of the European Union,[8] **"the primary objective of the European System of Central Banks (…) shall be to maintain price stability.** Without prejudice to the objective of price stability, it shall support the general economic policies in the Union with a view to contributing to the achievement of the objectives of the Union."[9]

---

[8] Consolidated Version of the Treaty on the Functioning of the European Union (OJ C 326, 26.10.2012, p. 47). See in particular Article 127.

[9] These objectives are set out in Article 3 of the Consolidated Version of the Treaty on European Union (OJ C 326, 26.10.2012, p. 13). They include promoting the well-being of its peoples, providing an area of freedom, security and justice, and aiming at sustainable development, full employment, social progress and a high level of protection and improvement of the quality of the environment.

Against this background, the ECB's Governing Council decided in 1998 to measure price stability in the euro area in terms of the HICP and to define price stability in quantitative terms as "a year-on-year increase in the HICP for the euro area of below 2%".[10] It also decided that price stability, according to this definition, "is to be maintained over the medium term". One argument for aiming for a positive inflation rate (instead of zero inflation) was the existence of a presumably positive measurement bias of the HICP[11] (see Box 7 and Chapter 3). Moreover, the medium-term orientation also accounted for measurement uncertainties. In 2003 – as an outcome of the last strategy review – the Governing Council further clarified that, in the pursuit of price stability, it aims to maintain inflation rates below, but close to, 2% over the medium term, thereby introducing an inflation buffer with respect also to the assumed measurement bias.[12]

The choice of HICP as a measure of price stability in view of its favourable characteristics was confirmed in 2003 (see also Box 2) and the assessment of its suitability remains valid. The HICP measures price developments for a basket of goods and services representing average household consumption structures from the previous year.[13] It is a cost of goods index (COGI) that emphasises consumer welfare by focusing on household consumption rather than measuring all prices in the economy.[14] The COGI concept was a natural choice, as the HICP is derived from national consumer price indices (CPIs) in the euro area, which also follow a COGI approach. This use of the same concept should support public understanding of the HICP. Public understanding is also facilitated by the fact that the HICP is compiled according to the "acquisition approach", i.e. it includes only items whose purchases involve prices based on actual monetary transactions between the household sector and other sectors in the economy, therefore excluding non-market goods and services. Imputed prices and transfers in kind, asset prices (see the box 2 in Chapter 4) and interest rate developments are also excluded.[15] Moreover, the HICP is a timely and reliable index. It is published monthly and is in principle not revised.[16] The HICP is comparable across euro area countries in view of its harmonisation. It is also used as a convergence criterion. Overall, the design of the HICP fulfils the requirements that consumers actually view it as an expression of their purchasing power and believe that monetary policy can steer it. The 2003

---

[10] See European Central Bank (1999), in particular pp. 45-47. See also Box 1, p. 39, "The Eurosystem's mandate: Key excerpts from the Treaty".

[11] See European Central Bank (1999), in particular pp. 46-47.

[12] See European Central Bank (2003). Deflation risks and inflation differentials between euro area countries were also mentioned as reasons for the introduction of an inflation buffer. The Price Stability Objective Strategy Work Stream also reviewed these reasons.

[13] The HICP is constructed as a Laspeyres index with annually updated weights that is chain-linked over December t-1. HICP weights in year t shall represent consumption structures of t-1. They are based on annual average expenditure shares of t-1 which in practice are mostly approximated by the penultimate year. The weights are obtained by price-updating these expenditure shares to December t-1. See, for example, Eurostat (2018).

[14] However, it cannot be derived from welfare theory, as it does not refer to the utility (see Box 1 and Chapter 5 for the difference between a COGI and a COLI).

[15] The main exception is the treatment of seasonal items, for which out-of-season prices may be imputed based on prices observed for other products, except in special circumstances such as the COVID-19 pandemic (see Box 5).

[16] Only in exceptional circumstances, e.g. when new methods are introduced.

strategy review confirmed that the HICP continues to meet the requirements of an index to be used in attaining the ECB's price stability objective.[17]

The view that the HICP meets the requirements to act as a credible target corresponds with the empirical finding that consumers' inflation perceptions are broadly in line with HICP inflation. The favourable characteristics of the HICP should support its credibility. Indeed, this seems to be reflected in practice, as there is on average a strong co-movement of perceived and measured inflation (see Chart 1 and Box 3).[18] Still, there is a persistently positive gap between the two.[19] One explanation for this gap is that consumers are not experts and may lack awareness and knowledge of inflation and related concepts (e.g. percentage changes or quality adjustment).[20] Also, they often have in mind specific price developments, for example of frequently purchased goods, or different consumption baskets, that reflect their personal experience (see, for example, Box 10, on OOH and perceptions, in Chapter 4). Furthermore, the co-movement of actual and perceived inflation is time-varying. Periods during which this relationship weakens appear to coincide with periods of shocks or structural changes (e.g. euro cash changeover or the COVID-19 crisis). In such times of increased uncertainty, it becomes clear that consumers might not fully grasp the nature of the data generally captured in the HICP. It becomes equally clear that the HICP reflects a representative consumer basket which might not fully reflect the criteria to which individual households link their purchasing power. This illustrates the need for continuous improvements in communication and transparency and, more generally, for investing in improving consumers' economic and financial literacy. To monitor periods of weakening credibility, more direct insights into consumers' opinions about inflation and its measurement would also be desirable. Surveys providing such insights include the ECB's Consumer Expectations Survey (CES)[21]; the Oesterreichische Nationalbank (OeNB) Barometer survey from 2013, which posed multiple questions on consumers' inflation opinions; the new Deutsche Bundesbank survey on household expectations; and the Special Survey of Italian Households (SSIH) by the Banca d'Italia.[22]

---

[17]   See Camba-Mendez (2003) and Issing (2003).

[18]   See also European Central Bank (2005) and European Central Bank (2014).

[19]   In the European Commission's Business and Consumer Survey, respondents are asked, for example, to state their beliefs about how consumer prices have changed in the past twelve months. Replies are either of a qualitative nature, when respondents can choose between categories such as "has risen a lot" or "stayed about the same" or of a quantitative nature. The questions in this survey are not specific to the CPI or HICP. Other surveys querying perceptions might ask more specific questions or give more guidance on the development of prices (i.e. ranges).

[20]   In this context, the ECB has provided an interactive publication on inflation on its website which is intended to deepen the general public's understanding of HICP concepts and methods and how they are used by the ECB. The publication also includes a personal inflation calculator.

[21]   See Consumer Expectations Survey.

[22]   The second quarter of 2013 edition of the OeNB Barometer, a quarterly (now biannual) representative survey of 2,000 Austrian households on central bank-specific topics, canvassed consumers' quantitative inflation perceptions and expectations. It also included direct questions on their opinions about the reliability of price indexes such as the CPI and HICP. See Fluch et al. (2013) (German only) or Fritzer and Rumler (2015) for more information on the survey. For the German Panel on Household Finances, see, for example, Deutsche Bundesbank (2019c). See the Special Survey of Italian Households.

Chart 1

Relationship between HICP inflation and inflation perceptions in the euro area

(y-axis: index of qualitative perceptions, x-axis: annual percentage change in percentage terms)
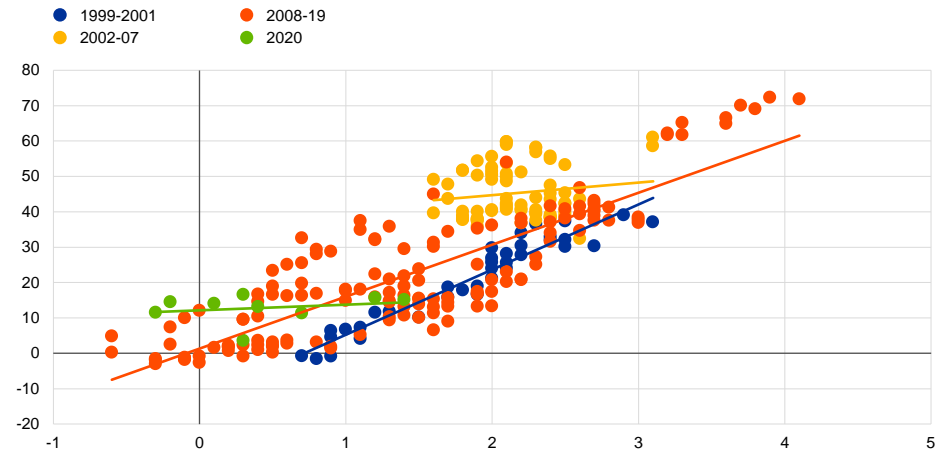
● 1999-2001   ● 2008-19
● 2002-07   ● 2020

Sources: European Commission, Eurostat and authors' calculations.
Notes: The index value on the y-axis refers to a balance statistic. The Commission survey provides several response categories for price changes: PP "risen a lot", P "risen moderately", E "risen slightly", M "stayed about the same" and MM "fallen". From these categories, a balance statistic is calculated as follows: B = PP + P*0.5 – M*0.5 – MM.

The ECB's strategy review in 2003 identified areas for further improvements in the HICP, notably the inclusion of OOH in the HICP basket. OOH represents a very large share of household expenditure in the euro area. That said, the share of expenditure represented by, and the price concepts applied to, OOH in each euro area country's CPI differ greatly. Therefore, its full inclusion in the harmonised index would improve the representativeness and comparability of national HICPs (see Chapter 4). At the same time, the 2003 strategy review recommended further methodological harmonisation of HICP measurement across countries with regard to the treatment of product replacement, quality adjustment and sampling issues (see Chapter 3). More timely updating of annual weights was also suggested, for example using data from the previous year rather than the year before that (see Chapters 2 (Section 2), 3 and 5).

The 2003 review also concluded that it was not feasible to quantify the HICP measurement bias in the euro area precisely. The bias could result from factors such as a delayed updating of basket weights or a missing quality adjustment (see Box 2 and Chapter 3).[23] Yet the positive bias was one of the reasons for setting the price stability objective as a positive inflation rate (instead of zero inflation) as early as 1998 and was an important reason for introducing an inflation buffer in 2003.[24] HICP measurement uncertainties (arising for example from statistical factors such as sampling errors) are to some extent also accounted for by the medium-term orientation of monetary policy. Moreover, signals from the HICP are usually cross-checked with auxiliary indicators when obtaining information about future inflation and the underlying inflation trend. This also helps to look through possible measurement issues. Auxiliary indicators and their relationship with the

---

[23] The main reference here is still The Boskin Commission Report, which was published 1996 for the United States.

[24] European Central Bank (2003). However, this was not the only reason: deflation risk and inflation differentials also played a role.

HICP are discussed in Chapters 5 and 6. Taking the above factors into account, the best way to increase the credibility of the HICP is to reduce possible measurement issues by means of ongoing improvements to its quality by the ESS and a transparent communication strategy by both the ESS and the ESCB in their respective fields of competence.

2.2 Developments in the HICP since the 2003 review: improvements and new challenges

In general, the quality of the HICP has been improved in several ways since 2003, mostly with respect to increased accuracy and further harmonisation, as recommended in the strategy review of that year (see Figure 1).[25] First, HICPs have become more representative as prices, which can fluctuate greatly within a reporting month, for example for fuel and package holidays, are now collected on several days within a reporting month (and not only, say, mid-month). In addition, seasonal items like summer and winter clothing, seasonal food and holidays are more prominently represented in HICP baskets.[26] The treatment of out-of-season items has been harmonised and carry-forward approaches have been discontinued. It is now compulsory for euro area Member States to provide flash estimates, which are published at the end of each reporting month. Euro area flash estimates are also provided for the main HICP aggregates. Common principles have been drawn up for sampling and product replacement. Greater detail has been introduced to the HICP's European Classification of Individual Consumption according to Purpose (ECOICOP). Annual updates of HICP weights have become compulsory and national accounts data must be used as the primary source to obtain HICP expenditure shares.[27]

---

[25] Mandatory improvements in HICP methods are set out in legal provisions. The legal texts of relevance to the HICP are listed by Eurostat. In 2020, specific legal requirements were set out in an HICP implementing regulation.

[26] See European Central Bank (2011a) and European Central Bank (2016a).

[27] See European Central Bank (2012).

Figure 1

Improvements in HICP quality since 2003

Extended **price collection** within the reporting month
(2008)

Treatment of **seasonal products**
(2011)

Lower level of index aggregation - **ECOICOP-5**
(2016)

(2009)
Guidance drawn up on **quality adjustment and sampling**

(2012)
Annual update of **expenditure weights**

(2020)
New Implementing Act: National Accounts as basic source of **weights**

Flash estimates publication for all EA countries

Source: ECB.

Despite these improvements to the HICP's statistical quality over the years, the integration of OOH has not been addressed. The treatment of OOH was the main issue raised in 2003 (and even before then) in terms of coverage. The ESS started the development of OOH price indices (OOHPIs) based on the net acquisition approach in the early 2000s and has provided official OOHPIs for all euro area countries apart from Greece since 2016. However, since they include an asset price element and are published only quarterly and with a significant delay, OOHPIs are not yet considered suitable for integration into all-items HICPs.[28] The integration of OOH is discussed in detail in Chapter 4 of this report.

Nor has it been possible to quantify the HICP measurement bias in the euro area so far, or to assess its exact development over time. Approximating the measurement bias is potentially important in terms of possible adjustments to the price stability objective. Possible measurement biases and ways to assess them, according to recent works of the Eurosystem for instance, will be examined more closely in Chapter 3. The integration of OOH in particular, as well as advances in HICP measurement (for example with regard to quality adjustments), remain a high priority for the improvement of the HICP.

Given the new measurement challenges, further, ongoing advances will be required, mainly referring to transparency around the inclusion of new methods and new data in the HICP. Challenges arise, first, from the growing importance of new sources of price data, such as scanner and web-scraped data, which are increasingly used to compile the HICP. Of course, the inclusion of new data sources has several advantages as it could, for example, increase the HICP's representativeness, and should therefore in principle be welcomed. Still, the introduction of new sources could potentially create statistical breaks. It could also introduce a degree of heterogeneity to price indices across countries, arising not only

---

[28] See European Commission (2018).

from differences between the data sources, but also from differences in incorporating these data and in the compilation methods applied. Badly designed compilation methods could even lead to new biases in the index caused by chain drift, for example.[29] Box 4 sets out in more detail the main advantages and difficulties of using these price data to compile the HICP. Further challenges stem from methodological changes incorporated in previous years, such as the redefinition of special aggregates due to the introduction of a more detailed level of harmonised product breakdown for the HICP (the ECOICOP, as mentioned earlier).

Another challenge that has become apparent more recently relates to chain-linking over December, which might amplify statistical breaks and distort the measurement of annual rates of inflation. In general, the incorporation of new data and new methods might produce statistical breaks. For example, a different treatment of seasonal products resulted in the profile of the HICP subindex for accommodation services in Germany changing considerably in 2013 (see Chart 2).[30] Chain-linking over a seasonal peak in December amplified this statistical break and annual rates were strongly distorted for one year. A similar break occurred when a new seasonal profile of the package holiday price index was introduced in the HICP for Germany.[31] The resulting breaks complicate the computing, analysis, comparison and forecasting of inflation rates. They highlight the need for high transparency and good communication with respect to the computation of price indices by Eurostat and the NSIs, especially as regards the incorporation of new data and new methods. One suggestion could be an annual consolidated assessment by Eurostat of how methodological changes affect price index computation. Selected countries already publish such assessments in line with recommendations in the HICP methodological manual,[32] sometimes focusing on their national CPI.[33]

---

[29] Chain drift is a systematic bias, normally downward, in a price index caused by failing to compare the prices of outgoing and incoming products in replacement situations. If the price of the outgoing product is low (for example, on sale) and that of the incoming product high, and this price difference is not recorded in the index, chain drift occurs and the index becomes biased.

[30] See Deutsche Bundesbank (2013).

[31] See Eiglsperger (2019b) and Deutsche Bundesbank (2019a). Other examples of chain-linking issues, also in other countries, can be found in Dietrich et al. (2020).

[32] The HICP manual includes recommendations in Chapter 10, in particular on p. 213, on publication, with explicit reference to transparency for users.

[33] To give just one example, the National Institute of Statistics and Economic Studies (INSEE) regularly publishes documents about methodological changes, the integration of scanner data and quantitative evaluations of their impact. See, for example, National Institute of Statistics and Economic Studies (2021) or (2020).

Chart 2

Incorporation of new methodologies and HICP chain-linking

HICP accommodation services in Germany

(index: 2015=100)
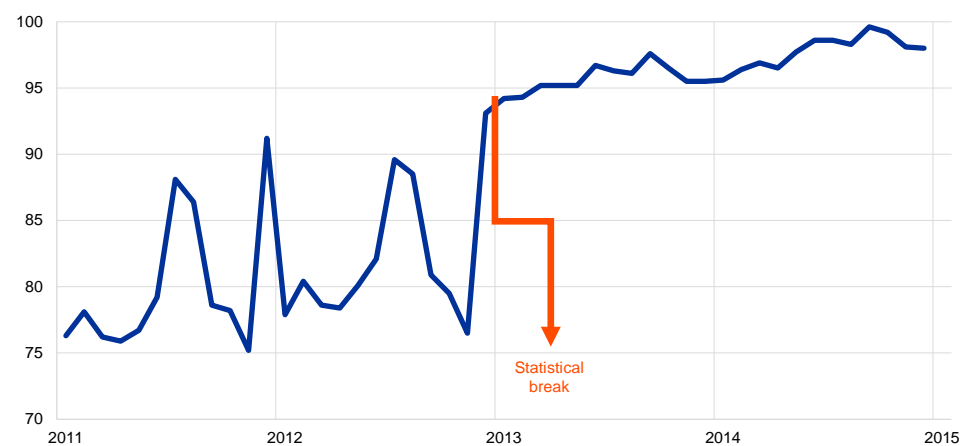
Source: Eurostat.
Notes: Monthly data and index values, with a statistical break in December 2012/January 2013. Index 2015 = 100 refers to 2015 on average. The latest observation is for December 2014.

The need for more transparency in price measurement became even more apparent during the COVID-19 pandemic. This is discussed in detail in Box 5. Pandemic-related radical shifts in consumption patterns were already prompting extensive discussion of HICP representativity in 2020 itself, most prominently in the paper on COVID inflation rates by Cavallo (2020). Those experimental indices sought to achieve better representativity using more timely weights. However, in most cases they only covered parts of average consumption.[34] While experimental indices can inform inflation analysis, if properly constructed, the utmost care needs to be taken in communication so that the risk of undermining the HICP's credibility is kept to a minimum. In this respect, the availability of more timely representative information about consumption structures would have helped assess the measurement biases of both the HICP – with respect to annual weighting – and experimental indices for which monthly weighting could be applied. Forward-looking information about the impact of the weight change in 2021 would have been even very important. HICP weights for 2021 represent consumption patterns for 2020, which were strongly affected by the pandemic. The ESS has developed a harmonised approach to take into account data sources other than the national accounts, based on Eurostat guidelines drawn for the application of the weights for 2021.[35] Nevertheless, there was an element of surprise in inflation numbers for January 2021 as the change in weight was unusually large. Furthermore, as the weights are usually only released at the beginning of the year, it is not possible to quantify the impact in advance with sufficient precision, for example by including new weights in historical short-term inflation projections (see Chart 3).[36] This hampered inflation analysis and forecast updating, not least because the weight effect could not

---

[34] Calculations by the National Institute of Statistics and Economic Studies (2020) and the Office for National Statistics (2020) were based on representative weights.

[35] See Eurostat (2020a). See also Box 6 in Chapter 5 and Gonçalves et al. (2021).

[36] See Box 5.

be separated from certain contemporaneous shocks (e.g. rising energy prices, a VAT increase and the climate package in Germany).[37] The need for timely representative information about consumption structures will most likely arise again: in 2022, when weights will potentially change again as a result of the fading impact of COVID-19 and, more generally, in times of strong consumption shifts. Therefore, more timely representative information about consumption structures for use in updating HICP weights is desirable. In this context, communication and collaboration around the treatment of imputations could be used as a blueprint for the handling of any new and future measurement challenges. The treatment of missing price quotes in the months of the lockdown, a major impediment to HICP measurement, was communicated very well by the ESS; this in turn supported the analysis of inflation developments.

Chart 3

Impact of pandemic-driven change in 2021 HICP weights

Euro area HICPX

(annual percentage change)
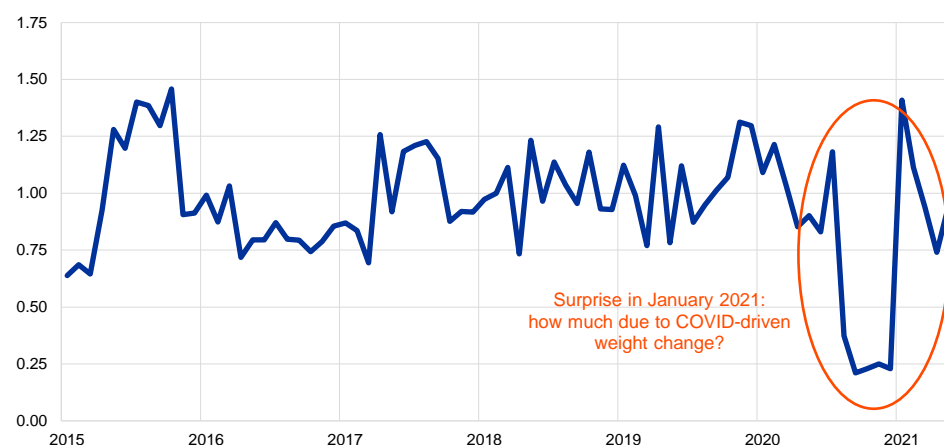
Source: Eurostat.
Notes: HICPX: HICP excluding energy and food. The latest observation for HICPX is the Eurostat flash estimate for May 2021.

In general, HICP users benefit from clear communication in terms of data, concepts and methods. Comprehensive all-items HICPs and their breakdowns – for the euro area, the EU as a whole and its Member States – are provided on Eurostat's website; regular data transmission to the ESCB allows the use of HICP data without significant delays. The HICP conceptual framework and its detailed application are set out in the HICP Methodological Manual.[38] More concretely, Eurostat informs the ECB (as one of the key users of the HICP) regularly of important compilation aspects on which, where necessary, the ECB is invited to share its views with the ESS. The ECB generally liaises with NCBs on HICP compilation issues via the established ESCB structures. The ECB and NCBs participate, as observers, in ESS I and task forces on HICP-related topics. These working arrangements in the HICP field between Eurostat and the ECB, the NCBs

---

[37]   See Deutsche Bundesbank (2021).

[38]   For more details, see HICP Methodological Manual.

and the NSIs remain important with respect to both future methodological advances and understanding the latest inflation developments.

2.3 Evaluation and recommendations: main areas for further improvement of HICP quality

Overall, despite continuous improvements in HICP quality, several HICP measurement challenges remain; some were already present at the last review and some are the result of new developments. New challenges from ongoing digitalisation have affected both consumption behaviour and price setting. New price data sources, which are increasingly used to compile the HICP, have gained in importance. Finally, changing consumption structures have increasingly been a focus for attention, notably during the COVID-19 pandemic. Against this background, first, the availability of more timely representative information about consumption structures (of harmonised quality) is desirable. Second, ongoing transparency around the integration of new data and new methods by the ESS is important. Given that NSIs largely pursue the implementation of these data and methods independently and in line with their capacities, further harmonisation efforts across countries will be needed. Third, and last, as recent experience with the pandemic has shown, changes in methodology and significant changes in weights could affect the overall HICP, including through chain-linking (see Box 6).[39] More flexible chain-linking, when needed, would therefore be a welcome advance. These new HICP measurement issues should be overcome by continuous improvements in HICP quality by the ESS.

Box 1
The HICP's conceptual basis: COGI

Price index theory commonly distinguishes between two classes of concept for measuring consumer price inflation: COGIs and COLIs. Each of these concepts is designed to serve a different purpose and to address different economic concerns.[40]

A COGI such as the HICP is designed to measure whether, and to what extent, the purchasing power of money is changing over time. Therefore, a COGI may measure the expenditure required to purchase a fixed basket of goods or, in order to compare like for like, it may instead measure the expenditure required to purchase a basket of goods of fixed quality. In this sense, the change in a COGI only captures pure price changes. A related, but distinct, economic question is the cost, to households, of achieving a certain level of well-being. An index that measures the cost of maintaining a fixed level of utility, under some class of utility functions, is called a COLI (see also Chapter 5). The change in a COLI measures the change in minimum expenditure required by a household to purchase a basket of goods and services that preserves a certain utility level or standard of living. Hence, constructing a COLI requires some assumptions about the preferences

---

[39] See, for example, Dietrich et al. (2021).

[40] For a thorough treatment of the theory of price indices, see von der Lippe (2007).

underlying households' decisions. Ideally, this also implies that a COLI should cover all goods and services that contribute to households' welfare.

Most CPIs in use today follow the COGI concept.[41] The HICP too is constructed as a COGI within each calendar year. Specifically, it is defined as an annually chain-linked Laspeyres-type index.[42] The scope of the HICP comprises goods and services with prices recorded in monetary transactions. The HICP focuses on "total household final monetary consumption expenditure", which links it to households' economic welfare.[43] A focus on monetary transactions has several advantages. First, it means that the HICP is calculated in terms of goods and services with prices that are observable in the market, by both consumers and statisticians. Since these prices can be directly observed by consumers when they shop, the HICP concept may gain in credibility in the eyes of the public. In principle, the prices do not need to be estimated or imputed; instead, they can be recorded, for example by statistical surveys of retail outlets. Lastly, the scope of the HICP provides a clear link to monetary policy. By highlighting the monetary cost of a typical basket of final consumption goods, the ECB's commitment to maintain price stability as measured by the HICP demonstrates to consumers that the money they hold will maintain a stable value in terms of consumption.

In contrast, the scope of an ideal COLI is not limited to goods with a market price, but is also designed to include non-market consumption items, including publicly provided services such as public health and education and, ideally, even public goods (e.g. national defence or clean air). The prices and consumption weights of these items must therefore be estimated or imputed. Moreover, an ideal COLI rapidly takes into account how consumption patterns change over time, ideally at real-time frequency (because, by definition, a COLI should take substitution into account). In practice, however, timely and sufficiently frequent information about consumption structures, and their changes over time, may not be available. In addition, individual levels of utility are not known; their aggregation across households requires certain assumptions. In practice, therefore, COLIs implement the minimum-expenditure concept indirectly by applying "superlative" indices, i.e. aggregation formulae that approximate the change in expenditure under an optimal consumption allocation by constructing averages between past and current consumption basket weights. The choice of the COLI's scope and the measurement of relevant non-market products also pose conceptual and statistical challenges. COLI-type indices often concentrate, therefore, on average consumption structures and exclude many or all non-market products, limiting the scope to those goods and services that can reasonably be measured ("conditional COLI").[44] Therefore, conditional COLIs may come to closely resemble COGIs, especially if statistical procedures such as lower-level

---

[41] The US Personal Consumption Expenditure (PCE) price index is designed as a COLI, with monthly updated consumption weights and a broad coverage of non-market products (see Chapter 5). The CPI for Sweden also follows the COLI concept, but with annual weights and only housing services as a non-market product.

[42] The HICP is constructed according to the Laspeyres concept, taking the previous year as the base period for defining expenditure weights. HICP time series are created by chain-linking monthly Laspeyres-type indices (monthly price indices, annual weights) over the December of the previous year.

[43] Commission Implementing Regulation (EU) 2020/1148 of 31 July 2020, Article 2(1). As Camba-Mendez (2003), p. 35, states: "The choice of consumption as the correct domain is justified on its own by the fact that consumption is the ultimate source of welfare."

[44] The coverage of the Swedish CPI differs from the HICP with respect to the inclusion of OOH; the CPI includes OOH according to the user cost approach.

aggregation, sampling and quality adjustment are treated in the same manner, which in practice is typically the case.[45]

When assessing the extent of any remaining biases in the HICP, it is important to clarify which reference concept is being considered. Chapter 3 evaluates potential HICP measurement biases in two ways. First, it assesses quality adjustment, the treatment of new products and outlets, and related technical issues relative to the legal framework defining the HICP. Second, it evaluates HICP relative to an optimally implemented COGI with truly representative consumption weights. That is, it measures representativity bias by calculating the difference between the HICP and a superlative index based on an average of current and past weights.[46] This use of superlative indices echoes the COLI concept, so that the difference between assessing a COGI relative to an optimally implemented COGI (in terms of representativity) or to a COLI of the same scope becomes very small, relating only to the frequency of weight updating.[47]

Box 2
The choice of price index: criteria and assessment in 1998 and 2003

In 1998 the Governing Council of the ECB adopted the HICP as the price index in its quantitative definition of price stability. The HICP was chosen because it fulfils the main criteria, and has the properties, that are considered necessary for a price index that is relevant to euro area monetary policy.[48] First, it is harmonised across euro area Member States. Second, it is subject to frequent statistical improvements, which enhance its accuracy, reliability and timeliness. Third, it is relevant and transparent such that it "most closely approximates the price of a representative basket of consumption goods and services purchased by euro area households". And fourth, the general public accepts that with such a target variable "monetary policy aims to provide full and effective protection against losses in the purchasing power of money".[49] The ESS had already taken these requirements into consideration when designing the characteristics of the HICP to be used for monetary policy purposes. It should be noted that the conceptual work related to the compilation of this price index is carried out by the European Commission (Eurostat) in close liaison with the NSIs. As a key user, the Eurosystem has been closely involved in this work.

---

[45] When classifying the HICP as a COGI rather than a COLI, it must be kept in mind that the aggregation formulae used at the upper level – across broad product classes – differ from those used at the lower level, between fine product types. Upper-level aggregation in the HICP takes the form of a Laspeyres index, but lower-level aggregation is performed with unweighted averages, due to the lack of adequate data on consumption shares across finely differentiated products. The procedure is usually the same for COGI and COLI at the lower level.

[46] Representativity bias refers to the difference between the HICP and the situation where weights aim to be fully representative of actual consumption. Diewert (2002) argues that the substitution bias should be considered part of the representativity bias, but that the latter also includes any other changes in buyer behaviour.

[47] That said, in considering the HICP the representativity of weights should be assessed on an annual basis. Annual weights data are a key element of the HICP concept, since more frequent reweighting would give rise to more technical difficulties related to seasonal products, seasonal discounts, and so forth.

[48] See, for example, Issing (2003), in particular pp. 11-30.

[49] See Issing (2003), p. 12.

In 2003, Camba-Mendez reviewed the choice of the HICP as the appropriate yardstick for judging price stability in the euro area and highlighted the following key conclusions:[50]

- "A price index should be credible and widely recognised as reflecting changes in all the prices paid by consumers."

- "At the present time, not enough is known to estimate with any degree of precision the magnitude of measurement biases in European consumer price indices."

- "The HICP has been designed to ensure the comparability of consumer price indices across EU countries. The harmonisation process is largely completed except for sampling practices, quality adjustment and the frequency at which weights are revised."

- "Household consumption expenditure is the correct domain of definition of a price measure. The inclusion of owner-occupied housing services in the HICP is desirable because it improves the comparability of the HICP across countries."

- "Sufficiently frequent re-basing of the consumption basket by all countries is important to ensure the accuracy and comparability of the HICP, and as a way of reducing the substitution bias."

- "There are both theoretical and empirical concerns against including asset prices in the index used for quantifying the price stability objective".

- "Underlying measures of inflation represent an appealing concept, but the large degree of uncertainty behind its computation is a deterring factor for its use in the ECB definition of price stability. Also, these measures may very well lack credibility with the public. However, they are useful indicators for monetary policy."

In line with Camba-Mendez's conclusions, the ECB's strategy review in 2003 confirmed that the HICP is the best available measure of inflation for the euro area for the ECB's quantitative definition of price stability. The HICP was positively assessed against practical properties such as credibility (as perceived by the general public), reliability (infrequent revisions), comparability (over time and across countries) and timeliness (e.g. the HICP is superior to the GDP deflator and the consumers' expenditure deflator, given its monthly frequency). High scores for all of these properties are the unique strength of the HICP.

Box 3
The relationship between HICP inflation and perceptions

The HICP, the Eurosystem's reference price index, is constructed to be a reliable and credible indicator measuring changes in costs of goods and services and should in practice reflect the changes in costs that consumers face. One approach to assess these qualities in practice is to analyse any gap and co-movement between measured HICP inflation and consumers' inflation perceptions. These perceptions are usually queried in consumer surveys such as the European Commission's Joint Harmonised EU Programme of Business and Consumer Surveys (hereinafter

---

[50] See Camba-Mendez (2003).

EU Commission survey).[51] Surveys directly asking whether or not the HICP can be considered a credible reference measure for price changes are virtually non-existent.

As several studies using such data have highlighted,[52] consumers' inflation perceptions have persistently been higher than, but have closely co-moved with, HICP inflation since the very beginning of the monetary union. The literature provides several explanations for this persistent perception gap. For example, survey respondents are not experts on this topic: they might lack knowledge of various concepts related to inflation[53] and are likely not aware of quality adjustments in the HICP. They may have difficulties with concepts such as growth rates or percentages, or simply may not follow price developments or associate them with past experiences.[54] The design of the consumer surveys may also in part explain the gap. With respect to explanations related to the way inflation is measured, some studies argue that consumers may have different consumption baskets in mind when being asked about consumer price changes (including, for example, housing prices: see Box 10 in Chapter 4).[55]

As can be seen in Chart A, simple scatter plots between qualitative inflation perceptions from the EU Commission survey and HICP inflation from January 1999 to September 2020 suggest a positive relationship overall.[56] The strength of this relationship changes over time, however, and varies across countries. The time-varying co-movement between inflation and perceptions suggests that in some periods, the HICP seems to be very much in line with perceptions. Periods during which this relationship weakens coincide with periods of greater uncertainty (euro cash changeover) and price measurement problems (COVID-19 crisis). Of course, periods of uncertainty can be noisy, and the strength of the relationship is therefore a result of other factors as well.

---

[51] In the EU Commission survey, respondents are asked to state their beliefs about how consumer prices have changed in the past twelve months. Replies are either of a qualitative nature, when respondents can choose between categories such as "has risen much" or "stayed about the same", or of a quantitative nature. The questions are not specific to the CPI or HICP but to consumer prices in general. The survey design can be viewed as relatively vague: it neither refers to a specific price index nor provides respondents with past values or ranges of inflation. Other surveys include the ECB's Consumer Expectations Survey and national surveys such as the Survey of Household Income and Wealth (SHIW) conducted in Italy by the Banca d'Italia.

[52] See, among others, Zekaite (2020b), Arioli et al. (2017), Aucremanne et al. (2007) and Döhring and Mordonu (2007) for Europe; Fritzer and Rumler (2015) for Austria; Del Giovane et al. (2009) for Italy; and Bruine de Bruin et al. (2009), Burke and Manz (2014), Binder (2015), D'Acunto et al. (2019) and D'Acunto et al. (2020) for the United States.

[53] See Rumler and Valderrama (2020) and D'Acunto et al. (2019).

[54] See Sims (2010) and Malmendier and Nagel (2016).

[55] See Zekaite (2020b). OeNB Barometer survey respondents stated that one of the main reasons for scepticism towards the CPI/HICP was that the price index did not reflect their consumption behaviour; see Fluch et al. (2013). Sceptical consumers also tend to have higher inflation perceptions/expectations; see Fritzer and Rumler (2015) and also D'Acunto et al. (2021).

[56] Corresponding, taking all country data together, to a correlation coefficient of 0.6.

Chart A

Relationship between HICP inflation and qualitative inflation perceptions

(y-axis: index of qualitative perceptions, x-axis: annual percentage change)
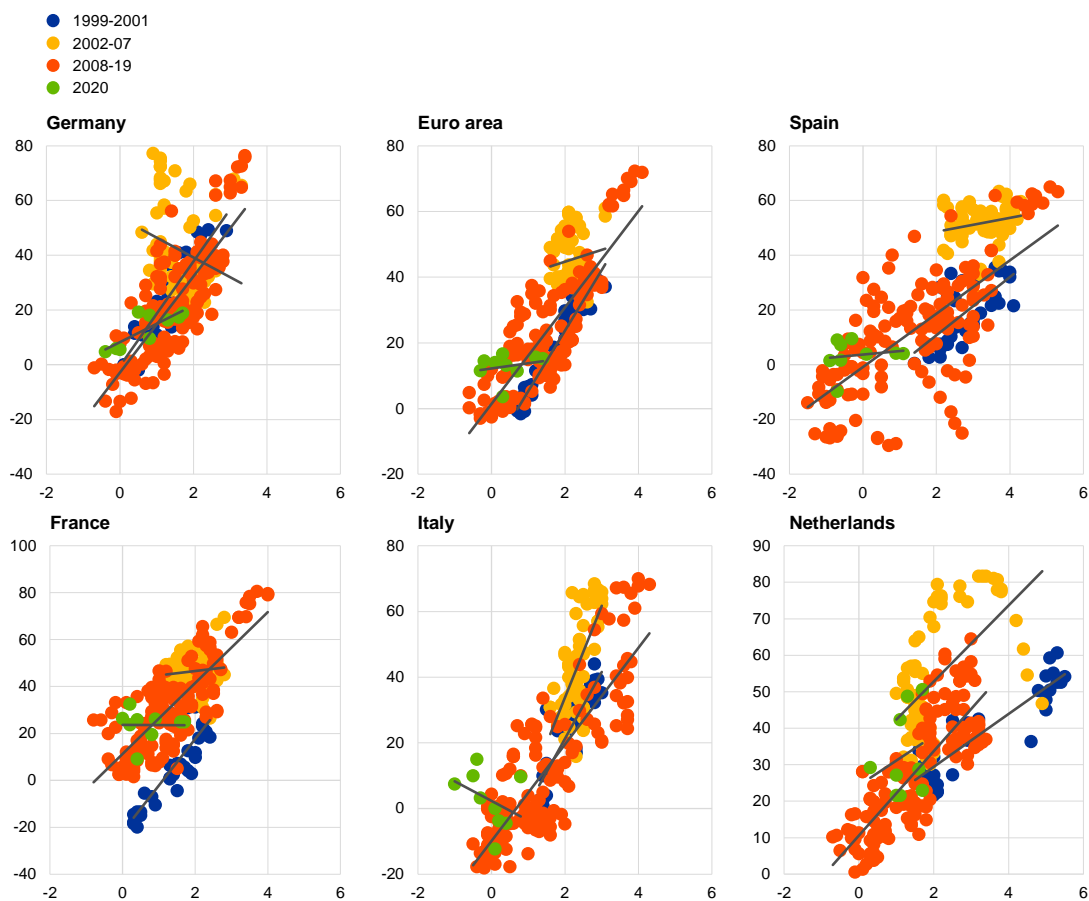
Sources: European Commission, Eurostat and authors' calculations.
Notes: The index value on the y-axis refers to a balance statistic. The Commission survey provides several response categories for price changes: PP "risen a lot", P "risen moderately", E "risen slightly", M "stayed about the same" and MM "fallen". From these categories, a balance statistic is calculated as follows: B = PP + P*0.5 – M*0.5 – MM.

From the perspective that for most of the time, the HICP follows the movement of inflation perceptions relatively reliably, it may be considered an adequate reference index reflecting consumers' perceptions. In the current COVID-19 pandemic, which is characterised by impediments to inflation measurement caused by large shifts in consumption patterns, lockdowns and the imputation of price data when computing the index, co-movement has decreased. This calls for thorough analysis, more transparency and more communication, particularly in times of increased uncertainty. As for the perception gap, some of its determinants could be addressed by improving non-experts' understanding and awareness of inflation and related concepts (therefore economic and financial literacy). To be able to monitor periods of weakening credibility of the HICP per se, more direct insights into consumers' opinions about inflation and its measurement are needed. Examples of how to achieve such insights include the CES or the OeNB Barometer survey from 2013, which posed multiple questions on consumers' inflation opinions, including whether or not they consider them reliable.

Box 4

New data and new methods: challenges posed by the introduction of new data sources to the HICP

Over the past few years, most NSIs have adopted, and even intensified their use of, alternative sources of price data for the computation of the HICP. These new sources include retail scanner data and web-scraped data. The main advantages of incorporating scanner and web-scraped data are an increase in coverage, product variability and the frequency of price collection. Importantly, the inclusion of new products is timelier. Product sampling can be made less subjective and price quotes from the new data sources more closely mirror the average price over a month.[57] A further advantage of scanner data is that these comprise actual transactions (e.g. turnover of products sold in one week or more weeks) and as such – if taken into account – sales-weighted data enter the index. While this increases representativity, incorporating sales-weighted prices requires more careful product sampling. Sales and discounts that affect turnover play a greater role and may result in more volatile price indices and biases.[58] Given the large amounts of data involved, statistical offices need to rely on automated data-handling methods, in particular when identifying and mapping different product varieties to groups, such as ECOICOP classes. Changing product characteristics, as well as relaunches and changes in product identifiers, can constitute a major challenge[59] for the creation of price indices. Procedures involving matching product attributes, text mining or machine learning become increasingly important.[60]

While in theory scanner and web-scraped data can be collected from many different types of retailer, much of the data used so far has come from a few large, uniform retail chains. With traditional price collection, several different outlet types, therefore including smaller grocery stores, corner shops and similar outlets, are usually covered. NSIs need to consider the composition of the outlets from which prices are collected, especially when outlets in the new datasets have different prices and price dynamics. Furthermore, scanner and web-scraped data are often only available for supermarket and fashion goods and for travel services, and not for all goods and services, which means that most prices still have to be collected in the traditional way.

Another aspect to consider is whether there are systematic differences between online and offline prices,[61] particularly concerning price trends. There is some evidence that online and offline price levels are often identical, although country, sector and retailer-specific differences do exist. Also, price changes are similar in average size and frequency, but the degree of synchronisation is low. Heterogeneity in online and offline pricing across retailers must be taken into account when only a few data sources are used.[62]

---

[57] See, for example, Eiglsperger (2019a) and Mehrhoff (2019). For practical recommendations on the use of scanner data, see Eurostat (2017a), and for web-scraped data, see Eurostat (2020b).

[58] See Eiglsperger (2019a).

[59] Product mapping can be particularly challenging with web-scraped data from fashion outlets. Products, for example a yellow t-shirt, usually stay in the (online) store for a certain period, during which the price decreases until the product is removed from the store. A potentially similar, yet different, yellow t-shirt with a different ID may then be introduced at a higher price. Treating all products as different products would result in a negative chain drift and constantly negative inflation.

[60] See, for example, the two-stage "C-approach" by Chessa (2016), which first bundles items that are sufficiently similar into products and secondly bundles products that are closely related to each other into a (consumption) segment.

[61] See, for example, Box 5 "Online vs. Offline Prices: Evidence from German CPI Micro Data" in the report of the Digitalisation Strategy Work Stream.

[62] See Cavallo (2017) or Gorodnichenko et al. (2018).

Further, in the recent past NSIs implemented scanner and web-scraped data according to their own methods and procedures. The intensity with which these data are incorporated, and the way this is done, depend to a great extent on the data availability, the data quality, the NSIs' capacities and resources as well as on conceptual and IT-related considerations. Scanner data may differ in terms of data reporting frequency, information about product characteristics, data provision within a reporting month (e.g. covering two or three full weeks). The aggregation may also differ, in particular with respect to whether or not monthly turnover weights are used (a popular approach does not use weights and applies geometric averaging combined with cut-off sampling; another option is to use constant weights and sample replenishment in case of product churn). Eurostat is addressing these aspects, with the aim of more harmonisation, by supporting efforts across countries and providing recommendations elaborated by ESS experts.

The COVID-19 crisis exposed NSIs to greater measurement uncertainty, making the use of new data sources such as scanner and web-scraped data even more relevant. Scanner data are particularly valuable in this respect, because they also contain information on changes in consumption patterns. This information may not only help to improve the representativity of the HICP but is also of economic significance in itself. To understand whether and how the inclusion of these new data sources affects aggregate price indices, the data and their behaviour need to be analysed thoroughly. The use of new data sources is likely to vary across countries and this could, if not addressed in a satisfactory manner, undermine harmonisation efforts in this domain.

Box 5
Inflation measurement challenges arising from COVID-19

The COVID-19 pandemic has posed several challenges for inflation measurement which are likely to persist for some time. The pandemic led to changes in households' consumption, and the possible implications of these changes for inflation have been documented in a number of publications.[63] Pandemic-driven changes in households' spending patterns in 2020 were not immediately reflected in the HICP's expenditure weights because the HICP keeps expenditure shares constant over a period of one calendar year. The shifts in consumption structures are therefore expected to affect the weights and numbers of some HICP categories in 2021.

The HICP weights for 2021 represent the consumption structures of 2020 (as required by the HICP's legal basis).[64] Due to the pandemic, however, the method used to derive HICP weights was considered to be less accurate than usual in 2020. Indeed, Eurostat has published new guidelines for the compilation of HICP expenditure weights for 2021.[65] Chart A shows the values of the HICP weights in 2020 and 2021 by special aggregates and their historical distribution over the 2012-19 period (the grey whiskers).[66] The largest weight decreases can be observed in services, in particular recreation, but also to some extent for transport, while the largest increases were seen in

---

[63] Contributions to this literature include Kouvavas et al. (2020), Cavallo (2020), Hood and Driessen (2020), Surico et al. (2020), Carvalho et al. (2020), National Institute of Statistics and Economic Studies (2020), Office for National Statistics (2020), Huynh et al. (2020), Reinsdorf (2020), and Zimmer and Jonckheere (2020).

[64] These are estimated with data primarily reflecting 2019-measured consumption expenditures and extrapolated using 2020 price developments.

[65] See Eurostat (2020a).

[66] See also Gonçalves et al. (2021).

the weight of food items and housing services. For some categories, the magnitude of the shifts was unprecedented; this can be better understood by examining the distribution of historical changes (Chart A, left panel).

The impact of strongly changing weights was partly reinforced by the HICP chain-linking procedure, which takes place in December. Generally, chain-linking over December could generate distortions in the measurement of year-on-year rates of inflation, especially when large consumption shifts occur between two consecutive years while price developments across items differ and show pronounced swings. The intuition is that even if the prices at the end of the year have returned to the same level as in the previous year, the year-on-year change of an aggregate index might show some increase/decrease due to changes in item weights across periods.[67]

Some of these weight changes led to significant shifts in annual inflation rates. This came as a surprise since full information about HICP weights usually only becomes available with the final January data, a factor which hampers inflation analysis and forecasting. A counterfactual exercise can be used to quantify the impact of the change in weights on HICP inflation. Applying the published HICP inflation figures for January 2021, a counterfactual HICP index has been constructed using the 2020 weights instead of the 2021 weights.[68] In January headline inflation is around 0.3 percentage points higher than the constructed counterfactual. HICP excluding energy and food (HICPX) inflation is around 0.4 percentage points higher as the impact on the weights of food and energy inflation is limited. At times when significant changes to HICP weights are occurring, it is particularly important to communicate possible HICP volatility early enough before the publication of the flash estimates. This was done, for example, with the publication of the estimated effects on inflation of the new methodology for the treatment of seasonal products.

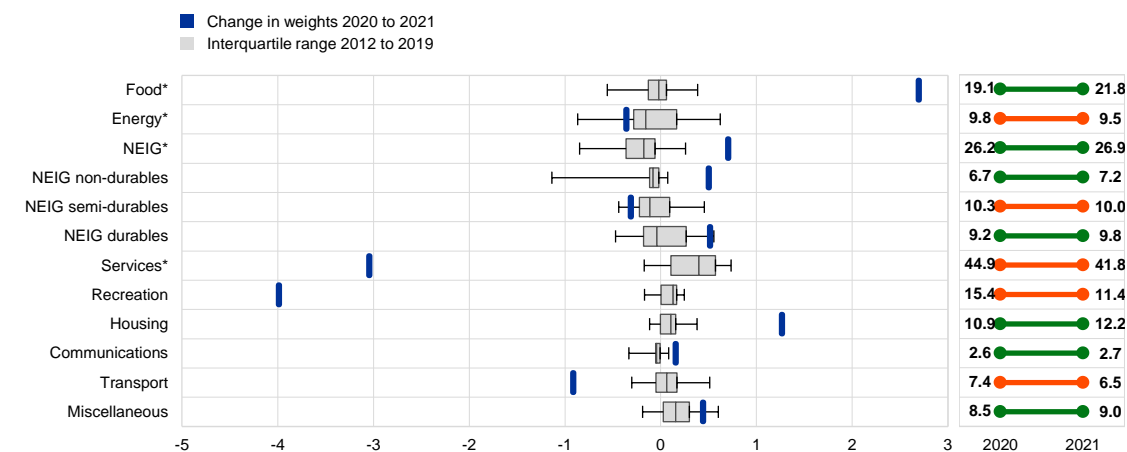[67] See, for example, the case of counterintuitive annual inflation rates following an exceptionally strong change in HICP weights for package holidays in Germany; Deutsche Bundesbank (2019b) and Deutsche Bundesbank (2021).

[68] To reproduce the exact HICP methodology, the counterfactual was constructed using data at the COICOP-5 level of aggregation. See Gonçalves et al. (2021).

Chart A

HICP weights: levels and historical changes in distributions

(percentage points)

■ Change in weights 2020 to 2021
▨ Interquartile range 2012 to 2019

	2020	2021
Food*	19.1	21.8
Energy*	9.8	9.5
NEIG*	26.2	26.9
NEIG non-durables	6.7	7.2
NEIG semi-durables	10.3	10.0
NEIG durables	9.2	9.8
Services*	44.9	41.8
Recreation	15.4	11.4
Housing	10.9	12.2
Communications	2.6	2.7
Transport	7.4	6.5
Miscellaneous	8.5	9.0

Source: ECB calculations based on Eurostat data.
Notes: The chart shows the HICP weights in 2020 and 2021 (right panel). A green/red line indicates whether the weight has increased/decreased. The chart also shows the changes in weights (blue markers in left panel) and their historical distributions (grey whiskers in left panel). The grey whiskers report the median, 25th percentile, 75th percentile, minimum and maximum of the historical distribution. NEIG stands for "non-energy industrial goods". Historical ranges are based on the sample from 2012 (the first year in which annual HICP weights became available) to 2019.

The lockdown periods also caused price collection issues, creating a need to impute the missing observations. The share of the headline HICP consumption basket that was imputed during the lockdowns was exceptionally high at times. At its peak in April, 32% of the overall HICP basket was imputed, a share that was even higher for some sub-indices (i.e. 40% for HICP services).[69] The imputation rate declined substantially by the summer but increased again to about 13% for overall HICP in January and February 2021.[70] Imputations hamper the interpretation of inflation rates to a certain extent but, since they were handled very transparently by the ESS, it has been possible to calculate rates without imputed components. However, the exact overall impact of price imputations on the change in HICP inflation is not fully known, since the detailed treatment of imputations at product level (below ECOICOP-5) was not revealed and is affected by the rules applied by NSIs in individual countries.

Overall, the challenges during the COVID-19 pandemic have highlighted the importance of transparency, particularly in times of shocks. The pandemic has shown the need to provide early guidance to producers of inflation statistics on how to deal with these abrupt and large economic changes when compiling the HICP, as was done by Eurostat with its guidance notes. The pandemic also revealed the importance of obtaining more timely representative information about consumption structures so that the impact of expenditure changes on HICP weights can also be gauged when considering the effect on inflation forecasting for the following year.

---

[69] The imputation shares in different countries and components since March 2020 can be found on the COVID-19-related Eurostat website.

[70] See also Kouvavas et al. (2020).

Box 6

Approximation of the effect of the weight change on the HICP rate

The HICP is a Laspeyres-type index, chain-linked over December of the respective previous year. Its annual rates of change reflect not only changes in prices, but are also influenced by weight changes, as HICP weights are updated every year. This box seeks to disentangle the impacts of price and weight changes[71] on the total index. The methodology used breaks down, in an approximate manner, the annual rate of change of the HICP for a certain month into changes in prices and in weights.

It can be shown that the annual rate of change AR^{mt} of the HICP of month m in year t (CP, with b as the index reference period) can be approximated by the fixed-weight Laspeyres change rate from m in year t-1 to December (LPC), a scaling factor (SF)[72], a price change during the reporting year (PC) and a combined price weight change (PWC).[73]

$$AR^{mt} \equiv \frac{CP^{b,mt} - CP^{b,m(t-1)}}{CP^{b,m(t-1)}} \approx LPC^{m(t-1),0(t-1)} + SF^t \cdot \left(PC^{0(t-1),mt} + PWC^{0(t-1),mt} \right)$$

The formula illustrates that the difference between the annual rates of change of the HICP and of a Laspeyres index based on fixed weights from t-1 is driven not only by the change in HICP weights, but also by the HICP's SF, which is determined by changes in HICP weights and the values of the elementary aggregates in the linking month.

We have calculated this approximate decomposition for all countries of the euro area[74] and for the euro area HICP.[75] Our results suggest that weight changes can have a substantial impact on annual HICP rates of change, especially at national level, while at euro area level the impact remained muted in the period observed. It was also found that the approximation works well at euro area level and for large countries with only small changes in the basket (differences of up to 0.2 percentage points between the official HICP and the approximated annual rate of change), while differences due to our approximation can be larger for small countries (we found differences up to 0.7 percentage points).

Euro area results

In Chart A, the left panel shows the approximative decomposition of the HICP annual rate of change. The chart suggests that the impact of the weight change is limited: at most, it leads to an increase in the annual rate of change of 0.13 percentage points (January 2019) or a decrease of 0.10 percentage points (July 2019), with an average effect close to zero, at 0.03 percentage points.[76] As the SF is close to 1, the weight change almost equals the difference (Chart A, right panel) between the annual rate of change of a pure Laspeyres index (LAR) and the HICP's annual

---

[71] For the effects of weight changes in times of high uncertainty and large shifts in the shares of consumption expenditures, as in 2020, see also Box 5 above.

[72] This is the core element of the chaining procedure. It drives the annual change rate, as it may scale up or down the index change from December (t-1) to mt.

[73] $PWC^{0(t-1),mt}$ reflects the second influence of the weight change; it adjusts the price change according to the Laspeyres formula for the changes in HICP weights which have occurred in $0(t$-1).

[74] With the notable exception of Latvia: confidential granular weights for a large part of the index make such an analysis impossible.
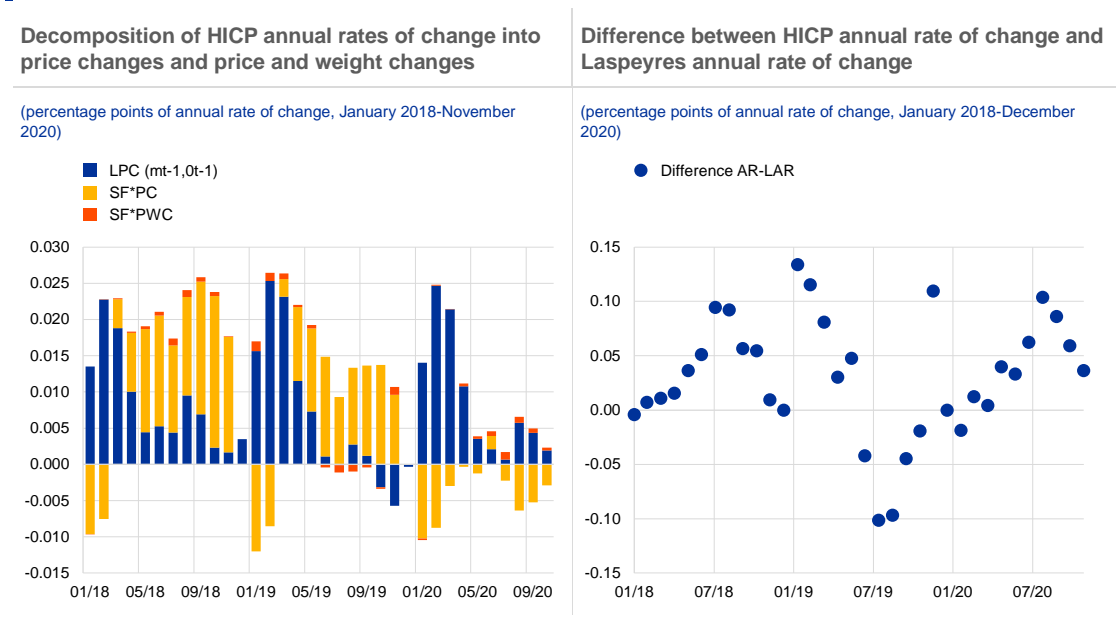
[75] The lowest-level indices and weights have been used. Lowest-level HICP sub-indices for all countries are only available from December 2016 onwards.

[76] PC and PWC need to be multiplied by SF to show meaningful results.

rate of change (*AR*).[77] The difference is positive in most cases, so annual HICP rates of change run higher than the respective rates of a fixed-weight Laspeyres index.

Chart A

Results of approximative HICP decomposition into price and weight change for the euro area

Decomposition of HICP annual rates of change into price changes and price and weight changes	Difference between HICP annual rate of change and Laspeyres annual rate of change
(percentage points of annual rate of change, January 2018-November 2020)	(percentage points of annual rate of change, January 2018-December 2020)

Source: ECB calculations based on Eurostat data.
Notes: Left panel: *LPC*(*mt*-1, 0*t*-1): Laspeyres price change from the same month of year *t*-1 to December of year *t*-1; *SF*PC*: Laspeyres price change component from December of year *t*-1 to the reporting month of year *t*, scaled with the *SF*; *SF*PWC*: component reflecting price and weight change from December of year *t*-1 to the reporting month, scaled with the *SF*. Right panel: *AR*: HICP annual rates of change, as calculated with the approximative decomposition; *LAR*: Laspeyres annual rates of change with weights from the previous year (same weights in each annual rate of change, no chain-linking).

Country heterogeneity

While this approximative analysis at the euro area level shows no substantial influence of the weight change, the situation differs at country level. While the maximum and minimum *SF* values influence the annual rate of change to a certain extent, the average value is very close to one. The weight changes exert a far greater impact. The maximum influence (seen in Malta, August 2018) increased the annual rate of change, relative to the rate of the fixed-weight Laspeyres index, of that country by 1.7 percentage points to 2.1% – almost all "inflation" can be attributed to the change in HICP weights. The maximum decrease (Germany, July 2019), of 0.5 percentage points, led to an annual change rate of 1.2%.[78] So this approximative decomposition already shows that the weight change can have a substantial impact on the HICP annual rate of change, while the impact is low on average over countries and months (0.05 percentage points).

---

[77]   AR as calculated with the approximative decomposition. It may deviate from the published annual change rate of the HICP, as it is an approximation.

[78]   The applied approximation method underestimates the HICP annual rate of change for Malta in August 2018 by 0.24 percentage points, while the HICP annual rate for Germany in July 2019 is overestimated by 0.10 percentage points.

3   HICP potential measurement bias and measurement uncertainty

The measurement of inflation by means of CPIs, and therefore by the HICP, is subject to various uncertainties and biases. The appearance of uncertainties and biases may have two main reasons: The first relates to the mode of data collection and to sampling, the second to the index calculation based on sampled data. In the mid-1990s, potential measurement biases in consumer price statistics were categorised by the Boskin Commission; the taxonomy underlying the Commission's bias estimation of the US CPI has become widely accepted since then.[79] This chapter addresses the upper and lower level aggregation bias, new products and quality adjustment bias as well as sampling and replacement biases – see also the overview in Figure 2.

The benchmark for an HICP bias in this chapter is its optimal COGI implementation. To estimate a bias for the current HICP implementation, it is necessary to determine a benchmark. The benchmark can be a narrow one – implementing the HICP according to its legal framework, assessing sampling and data measurement errors in this respect – or a wider one, based on an optimally designed COGI (see also Box 1 in Chapter 2). This COGI can be seen as the intended aim of the HICP. With regard to the upper-level aggregation bias, there is only one difference between these two benchmarks: an optimally designed COGI would also avoid the so-called representativity bias.[80] In this chapter, we compare the actual HICP both with the HICP according to its legal framework and with an optimally designed COGI.[81]

From the limited studies on measurement bias, we could conclude that there usually tends to be a small positive bias in consumer price indices (see also Box 7). A large majority of existing studies focuses not on the euro area HICP but on other CPIs and always uses the COLI concept as a benchmark.[82] The information content of these studies with respect to approximating the bias in the HICP is therefore limited, as the HICP addresses substitution, for example, better than a CPI with weights that are rarely adjusted. These studies can be classified in three groups: (i) studies using CPI granular data; (ii) meta studies, i.e. studies with estimates referencing other studies; and (iii) studies using alternative micro-data sources. Box 7 summarises these studies. Starting from the Boskin Commission

---

[79] Boskin et al. (1996).

[80] Representativity bias refers to the difference between the HICP and the situation where weights are intended to be fully representative of actual consumption. The representativity of weights is thus considered in a broader context and includes substitution bias. Diewert (2002) argues that substitution bias is part of representativity bias, but also includes any other changes in buyer behaviour.

[81] An optimally designed COGI is not bound by the practical constraints that the legal framework had to take into account, for example regarding the timeliness of weight availability.

[82] In practice, this is very similar to using an optimally implemented COGI as benchmark (for more details see Chapter 2, Box 1).

taxonomy and also acknowledging sampling errors, this section assesses the current measurement uncertainty surrounding the HICP.

Figure 2
Stylised overview: measurement bias
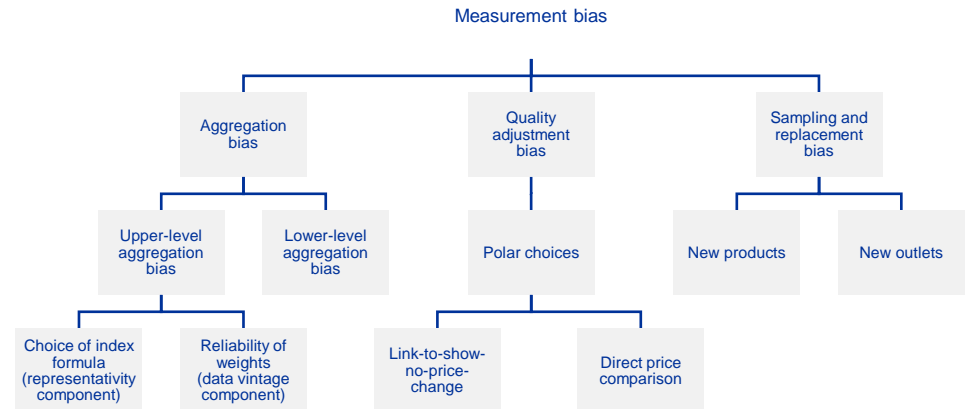
Source: ECB.

3.1 Upper and lower-level aggregation bias

Changes in economic conditions such as relative price or income can induce consumers to change their consumption patterns. Households can shift consumption across product categories (at the upper level) and within product categories (at the lower level). If such shifts are not reflected, the basket used to calculate the HICP is no longer representative of actual consumer expenditure, which in the COGI framework amounts to a representativity bias. The choice of boundary between the upper and lower-level aggregation also influences the bias.[83]

The mechanics of lower-level aggregation might lead to a larger bias than the choice of upper-level aggregation formula. ECB research[84] using household scanner data simulated differences between Laspeyres indices and superlative indices for a restricted set of products for three countries. Unweighted lower-level aggregation accounted for two-thirds of the total difference between both sets of indices, leaving only one-third to upper-level aggregation.

The introduction of a finer classification at the upper level of aggregation is likely to have decreased the overall aggregation bias of the HICP. Since December 2016, the HICP classification has been expanded to a more granular level.[85] As lower-level aggregation is the predominant source of bias, this

---

[83] Lower-level aggregation normally uses unweighted indices, while upper-level aggregation uses weights. If weights become available for lower levels also, a downwards shift of the border between upper and lower-level aggregation is possible. This "…would entail … multiple percentage points differences in measured price changes." Gábor-Tóth and Vermeulen (2018), p. 618.

[84] Box 15 in Chapter 5.

[85] The most granular harmonised level previously comprised 94 product categories, but the number has been expanded to 295.

methodological improvement should have reduced the overall aggregation bias of the HICP.

3.1.1 Upper-level aggregation bias

For price indices, weights that are not representative of actual consumption expenditure imply a bias. A price index built by comparing price levels for two periods, but using expenditure shares of only one of these two periods as weights, is biased compared with a price index that uses average expenditure shares of both periods as weights. This is the representativity bias mentioned above. A second source of bias can occur when weights based on vintage data are used instead of those from the latest revision: the data vintage bias.[86]

The introduction of annual reweighting in 2012 has reduced the representativity bias.[87] This is suggested by theory and is corroborated by recent work by Herzberg et al. (2021a) considering the German HICP. HICP weights are required to reflect expenditure shares of the previous year and are updated every year.[88] This procedure, which became mandatory in 2012, ensures a close proximity to actual consumption expenditure shares.[89] Annual weight updating by all Member States marks a significant advance towards a more comparable and representative overall HICP. However, as the annual update of weights relies on the first published estimates of the breakdown of private consumption expenditure as reported in national accounts, any inaccuracy in these can induce a bias.

Research indicates a small positive representativity and data vintage bias. An exercise conducted for this report to assess representativity bias looked at the difference between a Törnqvist index with weights from two consecutive years and the HICP.[90] The data vintage bias was also captured, by looking at later releases of national accounts data, which could not be taken into account for HICP production in real time. The results show both biases for the "big five" euro area countries and their aggregate for the years 2012 to 2019 (see Box 8). For the euro area aggregate including all countries, only the representativity bias was calculated. The results indicate an average combined bias for the HICP of less than 0.1 percentage points. The largest representativity bias was found for Germany, with 0.044 percentage points on average, and the smallest one for the euro area, with 0.022 percentage points, suggesting that contrasting developments in country HICPs tend to balance

---

[86]   Herzberg et al. (2021a, b).

[87]   Before 2012, minimum standards for HICP weights required that expenditure shares should be estimated from data not be more than seven years old. However, large changes in expenditure shares, which significantly affect total inflation, were required to be incorporated annually. In practice, in terms of intervals between updates of expenditure shares, NSIs followed different procedures, with a frequency ranging from one year to five. See European Central Bank (2012) for details.

[88]   For example, the weights for 2020 use the expenditure shares of private consumption from national accounts for 2018, which are updated to reflect consumption changes in 2019 and then price-updated to December 2019 with the respective sub-index of the HICP. For weighting in the HICP, sources and price-updating, see Eurostat (2018), Chapter 3, pp. 33-49.

[89]   The legal foundation for the procedure was Commission Regulation (EC) No 1114/2010.

[90]   Laspeyres indices use weights based on a single period from a past year. "Superlative indices" (e.g. Törnqvist or Fisher) use expenditure weights from two periods, a comparison period in the past and the most recent period (average weights in case of the Törnqvist index).

each other out on aggregate. The data vintage bias was in all cases smaller than 0.05 percentage points and even negative for the Netherlands (-0.068 percentage points).[91]

The representativity bias might vary over the business cycle, with the pandemic-related recession likely to be an outlier in this respect. While the representativity bias was measured as being 0.04 percentage points on average between 1997 and 2019, it peaks in the years of the global financial crisis and the subsequent recession (2008/2009), at almost 0.25 percentage points. This effect was even stronger for some euro area countries, such as Germany or France, where the representativity bias was as high as 0.3 percentage points during specific periods (Herzberg et al., 2021b). Similar results are found in a study for Sweden,[92] focusing on the difference between Sweden's CPI (a superlative Walsh index) and its HICP. Micro data studies, which focus on a subset of the HICP consumer basket and do not exactly replicate the split between upper-level and lower-level aggregation, find evidence for a business cycle dependence (see Box 15 in Chapter 5). Based on the results of these studies, the size of the bias increased in previous economic downturns. However, evidence for the euro area and Germany suggests that this was not the case in 2020. The non-availability of goods and services played a dominant role in the adjustment of consumption patterns during the COVID-19 crisis. This is in contrast to the effects of relative price and income changes that usually drive the representativity bias. Additionally, the high degree of price imputations in 2020 affected data quality, making it more difficult to draw firm conclusions about the underlying mechanisms.

3.1.2 Lower-level aggregation bias

Biases due to aggregation at individual product level may be related to the choice of aggregation formulae. Formulae based on arithmetic means may lead to biases as they imply an implicit weighting according to the price level. If there are large differences between the prices taken into account to calculate an arithmetic mean, the result will largely be influenced by expensive goods, leading to an upward bias.[93] Geometric means address this bias to a certain extent.[94]

Geometric average is the prevalent method of calculating elementary aggregates for the HICP. Overall, 11 countries, with a combined expenditure share of 63% in the euro area, use geometric means, while seven still use arithmetic means.[95] However, three of these countries[96] use lower-level weights to a large

---

[91] Based on the work of Herzberg et al. (2021b), a data vintage bias of 0.029 percentage points was calculated for an aggregate composed of the five largest euro area Member States ("Big-5 aggregate").

[92] We have analysed the difference between the Swedish CPI (a superlative Walsh index) excluding OOH and the Swedish HICP for the period January 1996 to December 2019. The results show that in the majority of cases (201 out of 276 months) the CPI excluding OOH exhibits smaller annual change rates than the HICP, as the theory of superlative indices implies.

[93] Leifer (1999).

[94] However, it may be noted that, in principle, the use of geometric means for the HICP is not consistent with its overall design as a Laspeyres-type index; for a discussion, see Diewert (2002), pp. 49-55.

[95] The information is taken from the HICP reference metadata (Eurostat, 2021).

[96] Germany, Estonia and Slovakia.

extent, thus limiting the bias.[97] Furthermore, the difference between unweighted elementary indices obtained as arithmetic and geometric means, respectively, seems to be rather small in practice.[98]

The use of scanner data provides opportunities to address lower-level aggregation bias, but also poses challenges.[99] With respect to the lower-level aggregation bias, the use of very granular expenditure data can help reduce this bias substantially. Scanner data contain both price and quantity information on all transactions by a given retailer, so they provide an almost complete picture of the basket of goods actually bought from a certain retailer. However, the main problem in processing such data is that the HICP is intended to compare like-with-like products; with many product varieties, of which several appear only during a limited time span, it is difficult to define sufficiently homogenous product items.[100] This can lead to a bias if the product definition is too wide (e.g. "yellow fruit") or too narrow (e.g. "baby banana air-shipped from India").[101] Furthermore, to avoid severe mechanical biases special index formulae must be used.[102] The application of "multilateral methods"[103] seems to be a promising avenue, while it should be mentioned that the use of different multilateral index concepts may lead to substantially different results.[104] The indices must therefore be assessed with respect to their robustness, along with other desirable properties. Research, in particular by NSIs in the EU, is very active in this field and should be closely monitored to avoid replacing old with new sources of biases.

3.2 Quality adjustment bias

In product replacement situations, a price index should only measure "pure" price changes; its change over time should not be affected by price differences arising from quality differences between the replaced and the replacement products. The principle of comparing like-with-like is a cornerstone of

---

[97] A special case is the Netherlands, where arithmetic means are used only for those products where neither scanner data nor web-scraping are applied

[98] See Goldhammer et al. (2019), Gábor-Tóth and Vermeulen (2018) and Leifer (1999).

[99] See also Box 4 in Chapter 2 on new data sources.

[100] See Eurostat (2017), p. 15.

[101] The resulting item "price" – a unit value obtained by dividing the turnover of a certain set of barcodes – assigned to the same item – by the respective number of units sold of that item – may suffer from "assignment bias". That means that the change over time of such a unit value may be affected ("biased") by changes in the composition of the product item (item description too heterogeneous). Alternatively, the unit value may suffer from an "assortment bias". This means that a price change of an identical item, occurring along with a change in the scanner data barcode of that item, would not be captured by the unit value (von Auer, 2016). Machine learning techniques may help solve that problem (Chessa, 2019a; Bertolotto, 2019).

[102] These biases typically increase over time and can occur for fixed weight indices chained and reweighted at a high frequency as well as for chain-linked superlative indices with monthly weights (de Haan and van der Grient, 2009; von Auer, 2019). For the reasons, see von Auer (2019).

[103] Price indices that include comparisons of all time periods within a certain time window (Chessa et al., 2017). They have been introduced for compiling price indices from scanner data with monthly weights, adapted from regional comparisons of purchasing powers.

[104] Białek and Bobel (2019), Chessa et al. (2017), Chessa (2019b) and de Haan et al. (2016).

constructing price indexes.[105] It requires to adjust changes in prices over time for quality differences that may exist between a replaced and a replacement product, for example a car model going out of production and its successor. Similar adjustments are necessary at a more granular level, for quality improvements in individual product items. The fundamental problem of quality adjustment is determining which part of the nominal price difference is related to the change in quality and which part represents the pure price change.

While it is well known that not carrying out any quality adjustment would lead to an upward bias,[106] **improper quality adjustment can also lead to a downward bias.**[107] Methods for quality adjustment are well established in the compilation of CPIs.[108] "Implicit methods" are based on general assumptions about certain relationships between overall product quality and price changes, while "explicit methods" provide a monetary estimate of the quality difference, by considering product specifications and quality aspects individually. In general, the use of different methods for quality adjustment may lead to different results. The two "extreme" methods either attribute the entire price difference between a replaced and a replacement product to the difference in product quality, or consider the entire price difference as pure price change, i.e. assume no quality differences. More sophisticated methods, in particular explicit ones, estimate quality differences that are often, but not always between the two extremes.[109] As product quality is a hidden factor, biases related to quality adjustment are hard to identify and to quantify, especially when explicit methods are applied. A downward bias is also possible, if quality adjustment is overdone.

For the HICP, there is no evidence that quality adjustment is systematically absent. Quality adjustment is a legal requirement for the HICP and is implemented for about one half of the HICP basket. The ESS offers a considerable amount of guidance in this direction,[110] but almost no binding rules. This leads to a variety of methods being used by NSIs, often related to national specificities in terms of data sources and access to related information. The heterogeneity in quality adjustment practices, being rooted in independent national decisions about which method to use, while reflecting national circumstances with respect to statistical data, may call for a stronger effort by the ESS towards more harmonisation. That in turn could increase the comparability of the country indices, prevent the generation of artificial

---

[105] "The matched model is the cornerstone of constructing price indexes. When products match over time, the characteristics of each product are held constant. Thus, any price change can only be attributed to inflation, and not to changes in characteristics." Groshen et al. (2017), p. 190.

[106] This refers to product replacement as discussed by Ahnert and Kenny (2004), p. 10, referring to the hedonic method of quality adjustment: "In the perceptions of many, hedonic methods are thought to yield lower results than other methods and reduce a perceived upward bias in price indices."

[107] If implicit methods are used in situations when their underlying assumptions are not met or if explicit methods are not correctly specified (Boskin et al., 1996, p. 27). Further examples can be found in Schultze (2003), Dalén and Tarassiouk (2013), Keating and Murtagh (2018), Goldhammer et al. (2019) and Bertolotto (2019).

[108] For an overview, see DESTATIS (2009), Chapter 5, pp. 29-43, and Eurostat (2018), Chapter 6, pp. 92-125.

[109] For a comparison of implicit and explicit methods, see Keating and Murtagh (2018) and Goldhammer et al. (2019).

[110] For example, the "Handbook on the application of quality adjustment methods in the Harmonised Index of Consumer Prices" (DESTATIS, 2009) and guidance in the HICP manual (Eurostat, 2018, Chapter 6, pp. 92-125).

cross-country inflation differentials in some HICP components where products across countries are very similar and avoid changes in measured inflation trends due to changes in quality adjustment methods.

A wide standard use of the same quality adjustment method for quite different product categories may influence the HICP inflation rate. Referring to the two "extreme" treatments, a country using link-to-show-no-price-change as a default method would see lower HICP inflation than a country using the direct price comparison method.[111] Research[112] has shown that systematic differences also exist between the bridged-overlap method and explicit methods.[113] The large differences in measured inflation for non-energy industrial goods across euro area Member States in the time period from December 2011 to December 2019 (from -22.3% in Ireland to +7.9% in Germany) may be partly explained by this factor.

While quality adjustment, being applied to about one half of the HICP basket, is not systematically absent, the question about the presence of downward biases remains. A simple method of calculating a bias does not exist. But it is possible to determine upper and lower bounds for quality adjustment in replacement situations (entailing quality changes but not constant prices) by assuming the two extreme cases, i.e. explaining all price changes by quality changes as the lower bound, and assuming no quality change as the upper bound.[114] Quality adjustment results beyond these bounds can be normally interpreted as bias (upwards or downwards, depending on the bound exceeded). Areas of strong technological progress, however, are often an exception in this respect. Quality-adjusted price indices for laptops, for example, may exhibit a downward trend, while the nominal price of the replaced and the replacement laptop has hardly changed. In such cases, the price effect attributed to improved quality is reasonably estimated to be larger than the difference in nominal prices.

Different quality adjustment strategies can lead to different inflation rates in the long term; however, there is no clear evidence about the size of the bias. Investigations with HICP micro data for Italy and Austria for several non-energy industrial goods (see Box 9 and Chart 4) support these conclusions. While Austria mainly uses explicit quality adjustment methods, Italy uses implicit quality adjustment, especially the bridged overlap method. An analysis of 12 product groups for Austria and ten product groups for Italy showed that adjustments remain within the bounds for Austria, with two, reasonable, exceptions. For Italy, the quality-adjusted index is outside the "corridor" between the lower and the upper bound for seven product groups. However, in all of these cases the distance from the corridor – which indicates a possible bias – is small (smaller than -0.1 percentage points for six product groups, and 0.1 percentage points for the seventh). The difference between the lower and the upper bounds can be substantial (up to 5.2 percentage points on

---

[111] Link-to-show-no-price-change: all nominal price change is assumed to be due to quality change. Direct price comparison assumes that all nominal price change is real price change.

[112] Dalén and Tarassiouk (2013) and Keating and Murtagh (2018).

[113] Bridged overlap: imputing the average price change of the other products in the sample. Explicit methods include the hedonics, option pricing and expert judgement methods.

[114] See Eurostat (2020c). The approach cannot be applied in the case of quality changes at constant prices, as the upper and lower bound would be identical in such a situation.

the yearly average for Austria and up to 0.5 percentage points for Italy). In such cases, the choice of quality adjustment method can have an impact in terms of different inflation rates in certain product categories of different countries, while the "corridor criterion" suggests that there is no clear evidence for a bias. The fact that the quality-adjusted indices for the Austrian example tend to the centre, while the Italian ones tend to the lower bound, may well explain some differences in cross-country inflation rates, and even trends, for these product categories.

Chart 4

Corridors for quality-adjusted indices for selected Austrian and Italian products

(index, January 2011=100)

- QA index
- LNP index
- DPC index

Italy: fridge, freezer

Italy: television

Italy: small electrical applicances (razors, etc.)

Austria: dishwasher

Austria: sofa set

Austria: notebook, tablet
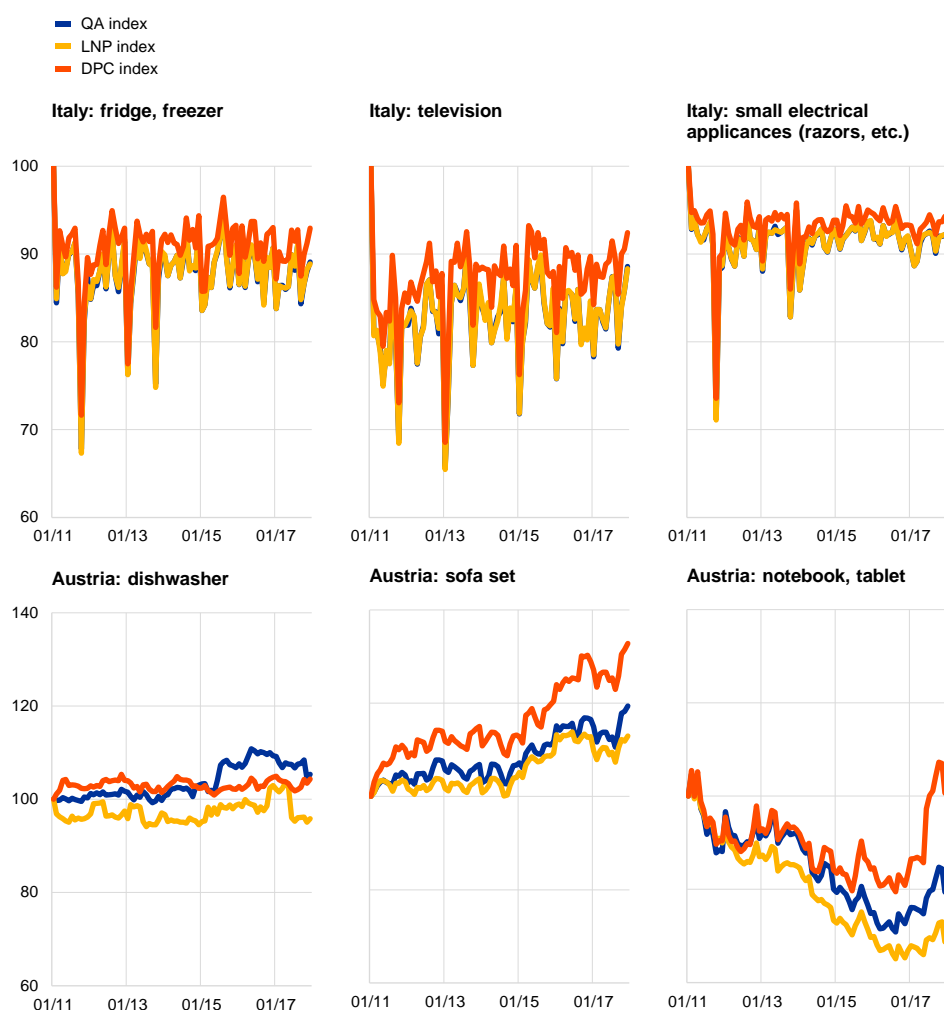
Sources: OeNB, BdI calculations based on Statistics Austria and ISTAT microdata.
Notes: QA index: quality-adjusted index, using quality-adjusted prices as for the HICP calculation. LNP index: index based on link-to-show-no-price-change, using a price difference of 0 as quality adjustment method in replacement situation. DPC index: index based on direct price comparison, using the full price difference in replacement situations (quality difference assumed to be 0). Data included from January 2011 to December 2017. In the charts for Italy, the QA index almost always matches the LNP index and is therefore "hidden" behind the LNP line in the charts.

At the euro area level, different methods of quality adjustment across countries may lead to upward and downward biases that could cancel out to a certain extent or could reinforce each other. There is not enough evidence to derive a sufficiently reliable estimate of the overall impact of potential quality adjustment biases on euro area HICP inflation. Both upward and downward biases

are possible, and they may generate diverging trends across countries for some sub-category indices. However, our investigation of Italian and Austrian microdata reveals that the use of different methods, even without necessarily exceeding logical bounds that would clearly indicate a bias, can drive HICP rates (and even the HICP trend) different ways. To that end, harmonisation of quality adjustment practices might lead to a reduction in divergent results.

3.3   Sampling and replacement

The collection of prices, product characteristics and expenditure data as well as the estimation of price indices on the basis of these data typically imply errors. When systematic, these errors imply biases. Two types of error can occur in data collection: sampling errors and non-sampling errors (observation errors, under-coverage, non-response). Errors related to random sampling are known to be of a non-systematic nature. However, it should be noted that pure random sampling is rarely used for CPI price collection.[115] As HICP inflation reflects changes over time, not only sampling at a given point in time plays a role, but also replacement strategies and the advent of new products and outlets, i.e. changes over time in the universe from which products are sampled.

As each HICP figure is only a point estimate, the reliability of each estimate with respect to sampling may be assessed by means of confidence intervals, based on variance estimates. However, we know of only a few examples from CPI practices, where at least some indicators for sampling errors have been computed.[116] Initiating such computations at ESS level would be a worthwhile effort to support the assessment of HICP reliability. However, the dominance of non-random sampling[117] across Member States, with approaches largely differing in several cases, and dynamic universes require innovative approaches and a significant effort in this respect.

Sampling errors are not constant over time as they depend on changes in the universe of outlets and products. For the HICP, an individual product is defined by "its characteristics, the timing and place of purchase and the terms of supply",[118] so the treatment of new products and new outlets poses statistical challenges. Biases related to the appearance of new products/outlets, for which expenditures by consumers have become significant, can occur because of a delayed introduction or

---

[115] The sampling design decides how representative the chosen sample is for the whole population. For probability or random sampling, it is possible to calculate the variance of the index (Dalén and Ohlsson, 1995). However, random sampling is hardly ever used in price statistics; target populations are typically unknown – in terms of size and composition – and are changing over time. In official price statistics, sampling of typical cases and judgement-based sampling are the predominant approaches. New data sources might make it possible to tackle this issue.

[116] Variances (Eurostat, 2018, p. 66) have been calculated for Sweden (Dalén and Ohlsson, 1995) and for the US CPI (Bureau of Labor Statistics, 2020).

[117] "Strictly speaking, calculating the sampling error in surveys by means of variance estimation requires probability sampling to be strictly valid. Since probability sampling plays a relatively minor role in the compilation of most CPIs and HICPs it is generally not possible to obtain reliable estimates of sampling error. Yet it is not entirely unreasonable to assume that samples are in many cases effectively random, i.e. the sample structure is such that it could have been generated by a random procedure." (Eurostat, 2018, p. 65).

[118] Definition according to Commission Implementing Regulation (EU) 2020/1148, Article 2(4) and (6).

replacement;[119] an aggregation bias may play a role in this context.[120] The absence of price comparison between different, but similar outlets and the appearance of innovative outlet types and products and of changes in price-setting strategies within and across outlets pose additional statistical challenges for price index compilations.

For delayed replacement and inclusion of innovative products, the HICP framework provides a generic treatment in the form of annual resampling and reweighting.[121] Remaining delayed replacement biases in the HICP should be present for less than 12 months and may be assumed to be small in normal times, when consumption baskets change slowly. They might become significant at times of big economic shocks that largely affect the allocation of consumer expenditure to different outlets, like the COVID-19 crisis.

No common methods have yet been developed for measuring price differences – adjusted for differences in outlet quality (e.g. in services provided by outlets) – between sampled and newly entering outlets as well as between different outlet types.[122] A comparison of prices between different types of outlet can be influenced by differences in the service quality provided by the respective outlet.[123] The same applies to innovative products, while several product features may have been provided by existing products. New outlets, as well as innovative products (e.g. smartphones, streaming services), enter the HICP usually in the course of its annual resampling. This implies that none of the prices affected are attached to the new entries.

New outlet types with different pricing policies may pose a challenge to price collection. Recent years have seen the rise of e-commerce as a new distribution channel, in which digitalisation enables fast price adaptation or personalised pricing. These pricing policies can lead to a bias in the price index, if not correctly approached.[124] Web-scraping of internet data[125] is an opportunity to obtain price and product-specific data quickly and at a high frequency, but aggregating these data into a price index poses specific challenges. These include the lack of weights, rapid product turnover and an unstable number of price quotations, which all can be potential sources of biases.[126] However, as the growing body of literature shows,[127]

---

[119] Outlets and product items in the sample may be replaced only when their turnover has become insignificant, while shifts in the actual turnover distribution would call for earlier replacement.

[120] If different price developments occur in different outlet types (such as discounters, supermarkets, department stores, e-commerce, etc.) and across individual product items; they should be weighted according to turnover shares. The non-availability of explicit weights may lead to lower-level aggregation bias.

[121] See Eurostat (2018), Chapters 3 and 4.

[122] Quality-adjusted price level differences between different outlets have been calculated for Iceland (Guðnason, 2003) for a limited number of cases, which where all evaluated in detail.

[123] Schultze and Mackie (2002), p. 168.

[124] Blaudow and Burg (2018) and Blaudow and Seeger (2019). For the monetary policy strategy review, a dedicated work stream on digitalisation is addressing these issues.

[125] See Box 4 in Chapter 2 on new data sources.

[126] National Institute of Statistics and Economic Studies (2019) demonstrates how web-scraped data on product characteristics of electronic goods can be used as input for quality adjustment. Blaudow and Seeger (2019) look at aggregation formulae addressing the lack of weights and Bertolotto (2019) discusses rapid product turnover.

[127] Blaudow and Seeger (2019), Blaudow and Burg (2018), Van Loon and Roels (2018), Griffioen and ten Bosch (2016) and Boettcher (2015).

NSIs are aware of these challenges and are developing solutions; web scraping has already become a normal data-collection method for the HICP. The application across Member States of sufficiently harmonised methods – in terms of applications, not necessarily with respect to product categories[128] – is supported by Eurostat in the form of recommendations and guidance.

3.3.1 New products and e-commerce: an ESS survey

The results of a survey on the treatment of new products and e-commerce in the HICP[129] indicate a large degree of heterogeneity among NSIs in terms of integrating new data. Technological changes and digitalisation pose a challenge for price collection. The survey, conducted with NSIs by Eurostat and the ECB,[130] shows significant differences in the procedures for including new products int the HICP consumer basket. While technological progress brings many innovations to the markets, only about two-thirds of NSIs in the euro area have established procedures to systematically identify new innovative products. In most cases, an innovative product becomes a candidate for being included in the consumer basket when it has reached the expenditure threshold[131]; often, the product is then included in the basket at the time of annual resampling.[132] A bias could arise in such cases as the price profile of an innovative product in the early stages of introduction could differ significantly from its profile at later stages. It is therefore likely that the timing of its introduction has an impact on the index.[133] Differences in the time of inclusion, while products may enter the market at similar points in time in euro area countries, may therefore introduce a bias in the euro area HICP.[134]

While the survey has shown that in most cases innovative products show a different price trend from the product class they have been added to, the size and direction of a potential bias arising from the inappropriate timing of product introduction remain uncertain. The direction of the bias can be both upwards (especially for goods) and downwards (for some services).[135] The survey results also show a positive correlation between the countries' income level (and

---

[128] This is methodologically correct to the extent that differences across countries with respect to product categories are justified by market conditions.

[129] Goldhammer et al. (2021).

[130] The survey was conducted in December 2020 and was intended to discover how NSIs include new products in the consumer basket and the significance of e-commerce in compiling the HICP. The answers were provided by 14 out of 19 NSIs from euro area countries.

[131] In accordance with Article 4(4) of the HICP implementing act, Commission Implementing Regulation (EU) 2020/1148 of 31 July 2020, the expenditure threshold is 0.1%.

[132] The introduction at the time of annual resampling (December of each year) means that the new product is included in an index-neutral way, i.e. the price is not compared with that of a previous product and potential price differences are not taken into account (Eurostat, 2018, p. 56).

[133] Sometimes new innovative products (especially goods) enter the market at a very high price, which drops significantly in the early stages of introduction. The opposite could apply for services that enter at a low introduction price and then steeply increase.

[134] Differences in the timing of inclusion can be caused by late adoption of the innovative product or delayed inclusion by the NSI. Delayed inclusion is the potential source of the bias.

[135] This separation depends on the respective pricing strategies applied to innovative products and services. Innovative services tend to be introduced at a lower price (e.g. streaming services) that increases over time; innovative goods (e.g. smartphones) are introduced at a high price that decreases over time.

degree of digitalisation) and the early inclusion of new products. This is also true for the share of e-commerce reported by NSIs. An explicit assessment of whether the timing of the introduction was appropriate would require additional detailed information for each country. However, the large disparity in the times products are incorporated into HICP baskets suggests that practices are not harmonised. The current state of inclusion of ten innovative products is shown in Chart 5.

The share of e-commerce in retail sales, and its inclusion in the HICP, varies substantially among euro area countries. While the survey provided only limited evidence for different price trends in different retail channels,[136] it points to a possible sampling bias for countries with a high e-commerce share but no inclusion of e-commerce in the HICP (see Chart 5, right panel). By contrast, NSIs that have already included e-commerce have also made substantial progress in this area, as they have increased the number of product groups for which prices at e-commerce outlets are available over time. Most countries use prices from the e-commerce channel for, on average, 39 out of 108 product groups of non-energy industrial goods. Hence, a possible bias for the euro area HICP due to the non-inclusion of e-commerce is expected to be limited.

Chart 5

Innovative products and e-commerce in the HICP

Diffusion rate of selected product innovations by HICP country weight for the euro area

(2020, weight percentages, 18 countries reporting)

Share of e-commerce and the inclusion of e-commerce prices in the HICP by country

(see notes for reporting period, percentages)
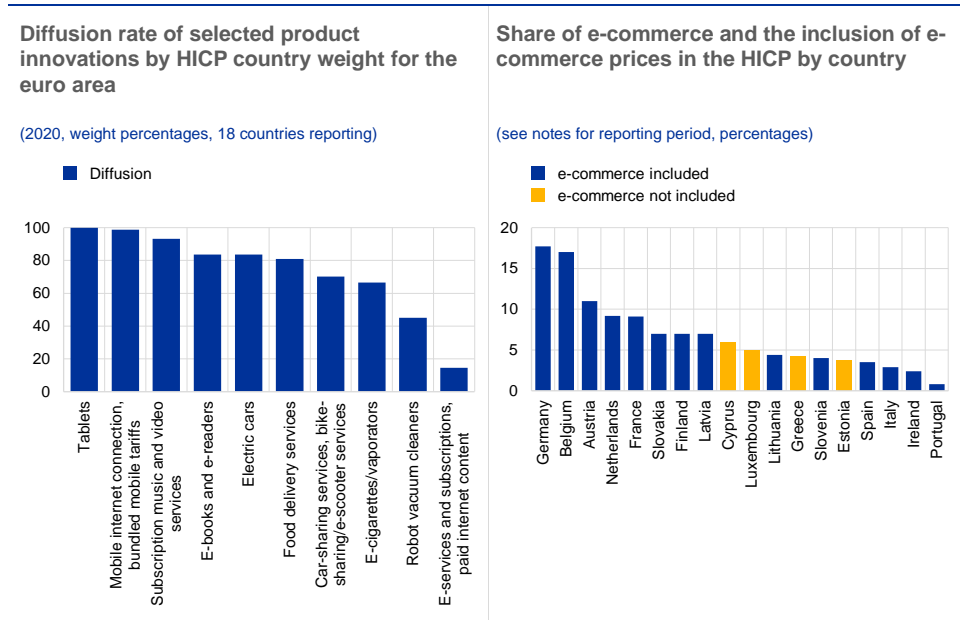
Sources: Eurostat and ECB survey, ECB calculations.
Notes: Left panel: the calculation is based on the country weights in the euro area HICP, re-scaled to the 18 euro area countries which reported in the survey. If a country had stated the inclusion of a certain innovation in its basket of goods, its country weight in the euro area was counted for this product and summed with the weights of other countries that had included this innovation.. Right panel: share of e-commerce in retail trade of non-energy industrial goods. Blue bar: e-commerce prices included in the HICP; yellow bar: prices not included. Reporting period for e-commerce share: 2015 (Germany, Luxembourg), 2018 (Belgium, Estonia, Ireland, Italy, Netherlands, Slovenia), 2020 (Finland, Greece, Latvia), 2019 (all other countries); inclusion: 2020.

---

[136] Only one country was able to compare the price trend for two product groups, and only over a period of two years, showing slightly lower price trends for e-commerce than for physical stores.

3.4   Assessment and recommendations

There is still considerable uncertainty around the overall measurement bias in the euro area HICP. It is not clear whether the overall bias has substantially declined since the last strategy review. Partial analyses suggest that counter-directional effects have been at play. Since 2003 considerable progress has been made in methodological terms and the HICP has been further harmonised. Methodological improvements in the HICP have focused on reweighting, resampling, quality adjustment and more granular breakdowns. All other things being equal, this should have reduced the biases in the HICP. However, these improvements may have been counterbalanced by issues related to the treatment in HICPs of rapid market developments, in particular the rise of e-commerce, the increasing numbers of product varieties and the faster pace of product and outlet replacement and innovation.

The upper-level aggregation bias is on average small but occasionally reaches a sizeable level under specific circumstances. A higher bias was experienced in the economic and financial crisis of 2008-09, for instance. The analysis in Box 8 shows an average upper-level aggregation bias of the HICP of less than 0.1 percentage points, including the bias related to the use of vintage data. However, evidence suggests that aggregation biases might be more pronounced at the lower level (see Box 15 in Chapter 5).

Wide heterogeneity in the application of the different quality adjustment methods might contribute to heterogeneous price index development for certain product categories and may, in some cases, even cause negative biases. Microdata analysis of selected products for Austria and Italy shows only marginal biases, if any, in a broad sense, referring to a meaningful corridor of annual change rates for quality-adjusted indices. In a narrower sense, inside the corridor, the analysis shows that the method used for quality adjustment can drive the size of inflation rates, which may result in heterogeneous HICP trends across countries.

The heterogeneity of inclusion practices for innovative products calls for a more harmonised and systematic approach. The size of a potential bias that could arise from an inappropriate timing of their inclusion and from e-commerce remains unknown. A survey conducted with the ESS showed no signs of systematically absent innovative products or inadequate treatment of the e-commerce distribution channel, but revealed heterogeneous practices which themselves could lead to biases.

The sampling error of collected prices and the data sources themselves also play an important role when assessing the accuracy of price indices. This has been widely neglected so far. Variances and confidence intervals of price indices and weights would enable more informed assessments of the measurement bias related to sampling and the related degree of uncertainty of the HICP.

Areas for further in-depth research and further action by the ESS have been identified. First, the aggregation bias in the HICP could be further reduced by investing in the development of more accurate HICP weights. Second, wide

heterogeneity in the application of quality adjustment points to a need for further harmonisation of methods and their implementation by the ESS. Further guidance on adopting improved and more harmonised quality adjustment and explicit methods, coordinated and monitored by Eurostat, could address the problems mentioned. The actions currently taken by Eurostat should be acknowledged in this respect. Third, the assessment of statistical uncertainty related to the sampling of products and their prices, expressed in the form of variances and confidence intervals, remains a largely unexplored area. Deriving a framework for calculating sampling errors for the HICP would therefore be beneficial to judge its statistical quality. However, as these kinds of uncertainties are not only related to sampling at certain points in time, but also to replacements and aggregation, related issues are highly complex. Hence, this endeavour would entail more statistical research. Overall, the analysis in this chapter suggests that there remains a knowledge gap concerning the size of the overall measurement bias of the HICP. This calls for further research by the ESCB and the research community at large, in close liaison with the ESS.

Box 7
Measurement bias in the literature

Studies using granular CPI data

Numerous studies using granular CPI data have been undertaken by researchers, also in central banks and in statistical institutes using microdata, published aggregates, methodological descriptions and/or example cases. They are useful to gauge the approximate size and range of potential measurement errors, while several findings are surrounded by elevated uncertainty.

One of the first studies of this kind – a study that initiated much of the discussion in the field – is the Boskin Commission report (Boskin et al., 1996), which suggested an average upward bias in the annual growth rate of the US CPI at that time of 1.1 percentage points using the conditional COLI framework as the benchmark. A follow-up of this study was undertaken by Lebow and Rudd (2003), who recognised some methodological improvements and estimated an upward bias for the United States of about 0.9 percentage points. Additionally, they introduced the concept of a weighting bias which refers to the use of vintage weights. Two other studies focused on specific biases in the US CPI. Greenless and McClelland (2008) did not find conclusive evidence for the presence of a new outlet bias, in contrast with the positive bias of 0.05-0.1 percentage points suggested by Lebow and Rudd (2003). Greenless and Williams (2009) investigated the upper-level substitution bias and found that more frequent reweighting can reduce this bias, but not eliminate it. There are several studies on biases for the Canadian CPI, in particular by Bank of Canada experts. Sabourin (2012) finds a positive aggregated bias of 0.43 percentage points and, notably, a negative quality adjustment bias. This negative bias was also detected by Kryvtsov (2006), who used CPI granular data for his investigation. For the Australian CPI, a study by Bishop and van Kints (2013) finds the upper-level substitution bias to be 0.24 percentage points for an index with weights updated only every six years.

For Europe, three studies used the taxonomy of the Boskin Commission. Hoffmann (1998) reports for the German CPI a bias of 0.75 percentage points; Lequiller (1997) investigates the French CPI,

which is very similar in concept to the HICP and reports a bias of 0.25 percentage points;[137] Neves and Sarmento (1997) find an upper-level substitution bias in the Portuguese CPI of between 0.04 and 0.57 percentage points.[138] Interestingly, in the 20-plus years since then, it seems that no studies based on national granular CPI data for European countries have been conducted. Altogether, the findings of existing studies suggest a positive HICP bias even when the COGI concept is applied as a benchmark.

Meta studies

Meta studies do not provide new bias estimates, but instead compare the properties of the index under consideration with that of the empirical studies mentioned above, and gauge the potential bias from such comparisons and analogies. This implies that the COLI concept is inherited as the benchmark. Brachinger et al. (2000) analysed the bias of the Swiss CPI in this way, arriving at a bias of 0.5-0.6 percentage points. Cunningham (1996) studied the UK Retail Prices Index (RPI) and concluded that the bias in the annual growth rate was 0.5-0.8 percentage points. Studies for the US CPI were updated by Moulton (2018), supposing a total bias of 0.85 percentage points

Studies using alternative granular data sources

For these studies, price indices were calculated on the basis of alternative granular data; the results were compared with the respective CPI. Common data sources are the internet (web scraping of price and product data) and supermarket or household scanner data.[139] The representativeness of the alternative data used in these studies is a general issue in these studies, since they normally cover only narrow segments of private consumption, while official CPIs are designed to be generally representative.[140]

Several recent studies calculate inflation rates using home-scan data with the caveats mentioned above. Kaplan and Schulhofer-Wohl (2017) calculate inflation rates at the household level and show that, on average, they vary with the US CPI. Braun and Lein (2021) use similar data to calculate biases potentially present in the Swiss CPI. They find a bias of 0.38 percentage points which, however, increases to 0.5 percentage points after 2015.[141] The study by Hayman (2006) focuses on price levels, not indices. It concludes that the CPI would not have missed cost decreases in households' budgets due to a shift of purchases to discount superstores, given that price levels in existing outlets were reduced when the discount superstores opened. Wynne (2005) delivers an estimate of the measurement bias in the HICP by comparing it with the qualitative inflation perception in the European Commission's business and consumer surveys. He finds a bias of 1-1.5 percentage points. Finally, several papers have been and are being written on biases arising from the COVID-19 pandemic (for example, Cavallo, 2020; Dunn et al., 2020; Surico et al., 2020; Carvalho et al., 2020). However, as many of these papers use experimental data and methods, and their focus is less on systematic biases and more on those caused by a major, unique

---

[137] Excluding a quality adjustment/new products bias, for which no estimate was provided.

[138] Estimated sizes of biases are typically related to the overall inflation level and may depend on the available breakdown.

[139] Special surveys and inflation perception data were used as well.

[140] The findings of some studies might have been adversely affected by chain drift. In such cases the biases are mainly negative, which may be related to not recording price changes when products are replaced.

[141] This refers to the classic bias estimates which directly compare a Fisher price index with the results of a newly calculated CPI (which does not use the Fisher formula); the latter however tracks the official CPI well.

crisis, we have decided not to consider their contributions as part of our core analysis. Some of the issues are discussed in Box 5 in Chapter 2 of this report.

Box 8
Empirical findings on upper-level aggregation issues in the HICP[142]

Most consumer price indices are based on a fixed basket of goods and services to measure pure price changes between two points in time. When prices change – in absolute or in relative terms – households tend to adjust their consumption, implying that the money they spend for individual goods and services usually varies over time. Hence, the assumption of a fixed basket (or constant expenditure shares) raises a measurement issue at the upper level of aggregation since, at this level, product categories are aggregated using expenditure weights. In the measurement literature, the focus is usually on the bias, which results from averaging the deviations of a price index from a reference over time. As price stability in the euro area is defined in terms of HICP inflation, the relevance of measuring existing upper-level aggregation bias is indisputable, although other sources of bias are likely to be quantitatively more important.

According to Herzberg et al. (2021a), the upper-level aggregation bias can be decomposed into two components: the choice of the index formula (from which a representativity bias may result) and the reliability of the weights (from which a data vintage bias may result). The HICP as a chain-linked Laspeyres-type index is constructed using disaggregated relative prices without explicitly accounting for quantity changes from the previous to the current year, whereas superlative index types, such as Fisher, Törnqvist or Walsh indices, incorporate the quantities of both years. An annual updating using data from national accounts ensures that HICP weights reflect current consumption structures to the extent possible. However, the preliminary status of national accounts data at the time they are used for deriving HICP weights implies measurement uncertainty and the potential of a bias, since final (or at least revised) national accounts data would have led to more accurate and reliable weight estimates.[143] While theory suggests that the representativity bias is positive, the sign of the data vintage effect is generally unknown at the outset. In addition to the bias, it is worth looking at the root mean square deviation as a measure for inaccuracy.

With a comprehensive set of national accounts vintage data at hand, it is possible to estimate the upper-level aggregation bias retrospectively. The bias was estimated for Germany, Spain, France, Italy and the Netherlands ("the Big-5") in the period from 2012 to 2019 (see Table A). The database also makes it possible to quantify this bias for an aggregate representing 82% of the euro area HICP (Big-5 aggregate). Measured in terms of annual HICP rates, the total upper-level aggregation bias of the Big-5 aggregate clearly falls short of one-tenth of a percentage point. This is a rather small number, in particular compared with evidence existing elsewhere. The representativity and the data vintage components contribute to the overall bias to fairly similar degrees. As expected, the representativity component is positive for all countries under consideration. Interestingly, the representativity effect for the euro area HICP is also markedly smaller than in the Big-5 aggregate.

---

[142] This box is a short version of Herzberg et al. (2021b).

[143] In Herzberg et al. (2021a), the reference is a superlative index based on full information weight, suggesting a closer proximity to the unknown "truth" than the final weights. With full information weights, however, the analysis can be carried out only for the German HICP.

Data vintage components are positive in all countries but the Netherlands. Owing to the negative data vintage component, the overall upper-level aggregation bias is negative for the Dutch HICP.

Table A
Metrics for bias and inaccuracy in HICP measurement.

Mean deviation (percentage of Törnqvist index with final weights) and root mean square deviation (percentage points per annum), 2012-19

Country	Mean deviation			Root mean square deviation		
	Representativity	Data vintage	Total effect	Representativity	Data vintage	Total effect
Germany	0.044	0.046	0.089	0.062	0.062	0.111
France	0.027	0.029	0.056	0.035	0.039	0.070
Italy	0.031	0.037	0.068	0.043	0.052	0.091
Spain	0.042	0.025	0.067	0.052	0.045	0.088
Netherlands	0.040	-0.068	-0.028	0.052	0.081	0.060
Big-5 aggregate	0.037	0.029	0.066	0.043	0.035	0.075
Euro area	0.022	-	-	0.028	-	-

Source: Deutsche Bundesbank calculations based on Eurostat data.

Considering the root mean square deviation, the results for the Big-5 aggregate, as well as for the individual countries, confirm that the representativity and data vintage components contribute with fairly similar sizes to HICP inaccuracy at the upper level of aggregation. While the root mean square deviations differ somewhat from one country to another, the inaccuracy of the Big-5 aggregate is relatively small.

Box 9
Is there a measurement bias from quality adjustment in Austria and Italy?

Quality adjustment in inflation measurement is a widespread practice by statistical institutes to account for quality changes when comparing prices of comparable, but slightly different, products over time. There are implicit and explicit methods of quality adjustment.[144] In this box, we assess quality adjustment practices in Austria and Italy based on CPI micro data for a number of selected products from the non-energy industrial goods component.[145] As Austria uses mainly explicit quality adjustment methods and Italy relies exclusively on implicit ones, both countries can be seen as exponents of their chosen methods, which helps gauge quality adjustment biases for other euro area countries as well.

To assess whether the quality adjustment by Statistics Austria and the Italian National Institute of Statistics (ISTAT) might lead to a bias in the measurement of inflation, we calculate inflation using

---

[144] For more details, see HICP manual (Eurostat, 2018, Chapter 6, pp. 92-125). See also Consumer Prices Indices Technical Manual, 2019 (UK Office for National Statistics), Consumer Price Index Manual (IMF, 2020).

[145] For Austria, the entire database includes monthly observations from January 1996 to December 2017 for over 1,000 products and services; the exercise is carried out on 12 non-energy industrial products. For Italy, it includes monthly observations from January 2011 to December 2018 in 267 eight-digit COICOP categories; the exercise is carried out on ten non-energy industrial products.

three different assumptions about the value of quality differences.[146] One extreme case is to assume that the difference in quality between a replaced and a replacement product is identical to the price difference between these two products. This approach is equivalent to assuming that the value of the new quality-adjusted price is equal to the old price at the point in time the replacement takes place; this method is called link-to-show-no-price-change (LNP). The other extreme case is to assume that the quality difference is either negligible or zero, such that no quality adjustment of prices is necessary. This approach is equivalent to directly comparing the old and the new price and is called direct price comparison (DPC). The third version we calculate uses the actual quality adjustment conducted by Statistics Austria and ISTAT.

Under reasonable assumptions,[147] the new quality-adjusted price should lie somewhere between the old and the new prices. Thus, inflation actually calculated by Statistics Austria and ISTAT should move between inflation calculated under the two extreme assumptions, which represent an upper and lower bound for reasonable quality adjustment. Put differently, when actual inflation systematically and permanently exceeds the range spanned by the upper and lower bounds, we would conclude that a measurement bias arising from quality adjustment cannot be ruled out.

In Table A and Table B, we report average annual inflation for the period 2011-17 for Austria and 2011-18 for Italy, for the three versions of constructed indices. The results for Austria indicate that for most of the products, the quality adjustment based on information from Statistics Austria produces inflation rates that are lower than if no quality adjustment had been implemented and higher than if quality adjustment were so strong as to completely offset the price changes. It is difficult to assess a priori whether a small quality adjustment is preferable to or more correct than an extensive quality adjustment, so for Austria a large corridor of reasonable quality-adjusted inflation remains. For Italy, where ISTAT mainly uses implicit methods of quality adjustment, results are reasonable and we can conclude that there are no major biases in price indices and inflation rates. Inflation rates derived from the official index are always close to the lower bound, which means that price changes in replacement situations are mainly attributable to changes in quality.

---

[146] We follow the proposal by Eurostat (2020c, Annex 4) for deriving upper and lower bounds of meaningful quality adjustment. However, Eurostat (2020c) notes with respect to the bridged overlap method: "Values outside these boundaries do not necessarily imply that the bridged overlap method is incorrect."

[147] The assumption is that the difference in the price of the replaced and the replacement product is larger than the difference in quality.

Table A

Austria: average inflation for different quality adjustment assumptions

2011-17, percentage per annum

Product	DPC	QA	LNP
Bedroom furniture	4.3	3.8	1.7
Sofa sets	2.9	1.8	1.2
Dishwashers	0.1	1.2	0.4
Electric razors	-1.0	0.1	-2.0
Toothbrushes	0.6	0.3	-0.3
Washing machine	0.2	0.1	-0.6
Lawn mowers	0.7	0.4	-0.2
Sinks	3.2	1.0	0.8
Laundry detergent	-1.2	-5.3	-6.0
Notebooks/tablets	-0.1	-3.1	-5.3
PCs	1.3	0.6	-1.7
Men's jeans	-0.6	-0.5	-1.4

Source: OeNB calculations based on micro price data 2011-17.
Note: DPC stands for direct price comparison, QA for the quality adjustment according to information from Statistics Austria and LNP for link-to-show-no-price-change.

Table B

Italy: average inflation for different quality adjustment assumptions

2011-18, percentage per annum

Product	DPC	QA	LNP
Men's trousers	-0.016	-0.017	0.07
Women's pullovers	-0.22	-0.22	-0.09
Washing machines/dryers and dishwashers	-0.003	-0.013	-0.010
Bedroom furniture	0.164	0.225	0.199
Laundry detergent	0.375	-0.067	-0.066
Fridges/freezers	-0.932	-1.337	-1.322
Appliances for heating and air conditioning	0.608	0.705	0.689
TVs	-0.383	-0.631	-0.638
Small electronic devices	-0.063	-0.081	-0.093
Jewellery and clocks	0.312	0.264	0.265

Source: Banca d'Italia calculations based on micro price data 2011-18.
Note: DPC stands for direct price comparison, QA for the quality adjustment according to information from ISTAT and LNP for link-to-show-no-price-change.

For Austria, taking the average for all products, quality-adjusted inflation is exactly half-way between the two extreme cases represented by DPC and LNP. For Italy, it seems that for seven out of ten products the quality-adjusted price index is outside the corridor, with five cases sitting below the lower bound. However, in all cases the distance from the corridor – and the possible downward bias – is small; with the exception of bedroom furniture, the bias would be within the rounding margin. The differences found between the explicit and implicit methods used by Statistics Austria and ISTAT, respectively, could mean that there is a need for more harmonisation of quality adjustment methods among euro area countries in the future.

4 The role of OOH

The inclusion of OOH in the HICP is desirable but faces implementation challenges. Internationally, many methods are used to pursue an adequate treatment of OOH in inflation indices. The NA and rental equivalence (RE) approaches are two possible avenues that could be followed in the euro area. Both methods have conceptual and practical pros and cons, which are discussed in detail in this chapter.

In terms of communication and consumer perceptions, NA is the preferred approach, as it tracks conditions in the housing market more closely (even though it refers only to part of the market; see Section 4.2). This chapter includes an outline for a roadmap to integrate OOH into the ESS's HICP using the NA approach. The complete development of a Eurostat official price index including OOH is likely to be a lengthy process. In the interim, steps can be taken to consider the corresponding experimental price indices (produced by the ESCB or the ESS) as part of the monetary policy strategy.

The implementation period for the new price index incorporating OOH needs to be carefully managed, especially from the communication point of view, as the existing HICP would continue to serve as the main reference for monetary policy purposes.

4.1 Conceptual considerations

OOH costs are defined as those costs associated with owning, maintaining and living in one's own home. These costs represent a significant share of households' consumption in the euro area (about 13% according to the national accounts) and are significant in predicting their perceptions of consumer inflation (see Box 10).

Currently, the HICP only partially includes changes in the prices related to living in an owned dwelling; it covers only the cost of maintenance and minor repairs and other running costs, both for tenants and for owners.[148] Quite apart from practical measurement issues, OOH costs have not been incorporated because conceptually they are not fully aligned with the HICP: the HICP captures only those changes in the prices of goods and services whose purchase generates monetary transactions for consumption purposes. OOH does generally not generate these types of monetary transaction and thus does not fit into the HICP concept in this regard. Different types of OOH indices have been proposed but none of them meet these conditions. Depending on the compilation approach, they either are not based

---

[148] As a reference, the weight of rents for primary residences actually paid by non-owners to landlords up to 2020 was about 6% in the euro area and increased to 7% in 2021. It is above or close to 10% in Germany and the Netherlands.

on actual monetary transactions or pose serious challenges in disentangling the consumption from the investment component in housing.

Internationally, and in the national CPIs of some euro area countries, there is considerable variation in the way OOH is treated in CPIs (Table 1). Where OOH is included, in most cases the related costs are imputed using the RE approach, and sometimes using the acquisition approach. Another approach is based on the user cost of capital, which quantifies the costs of OOH from the perspective of investment theory (i.e. buying a dwelling at the beginning of a period and reselling it at the end of that period). Finally, in some cases indices are compiled on the basis of the payments actually incurred each period by private households for their own housing (payments approach).

Table 1
Possible methods to include OOH or consider OOH costs in inflation indices

	Monetary policy purposes (1)	Other purposes (1)
User cost of capital	Canada, Iceland, Sweden (CPI with a fixed mortgage interest rate)	Canada, Iceland, Sweden
RE approach	Czech Republic, Japan, Norway, Switzerland, US (Personal Consumption Expenditure Price Index)	Cyprus, Czech Republic, Denmark, Germany, Japan, Netherlands, Norway, Switzerland, UK, US (Personal Consumption Expenditure Price Index or CPI)
NA approach	Australia (2), New Zealand (2)	Euro area Member States: HICP – separate OOH price indices, Australia (2), Finland (3), New Zealand (2)
Payments approach		Austria (for flats), Ireland
Not included	Euro area, UK	Most national CPIs in EU Member States not mentioned above; many others

Source: Eiglsperger et al. (2020).
Notes: (1) Where EU Member States are mentioned explicitly, the reference is the national CPI and not the HICP. (2) Consumer price indices provided with a quarterly frequency. (3) Updated monthly with a proxy index.

International experiences with the measurement of OOH costs offer some lessons for the euro area. For example, following the approach of the UK's Office for National Statistics (ONS),[149] several methods to calculate OOH costs could be considered simultaneously even if some or all are only provided on an experimental basis. While having multiple inflation measures produced by the national statistics offices could potentially be confusing, household living cost indices are published in several countries in addition to the main consumer price inflation measure (e.g. in Australia, New Zealand and the United Kingdom). The United States has two main consumer price indices (CPI and Personal Consumption Expenditures Price Index, which is the index used by the Federal Reserve System (FRS) to set its target) that both include an OOH cost component based on the RE approach.

The NA and RE approaches are two possible avenues for including an OOH component in the euro area HICP and will be discussed at length below. Conversely, the inclusion of OOH in the user cost and payments approaches can produce counter-intuitive results. In the former, for example, the inclusion of a capital

---

[149] The ONS publishes an official consumer price index, the "CPIH", which is the HICP for the United Kingdom including OOH, according to the RE approach. Alternative approaches to OOH, i.e. the NA approach and the payments approach – both with a monthly reporting frequency – are also considered with data published once per quarter.

loss/gains term – according to the comprehensive definition of OOH user costs from an investment perspective – would mean that when house prices are accelerating, the cost of capital component should be falling. In the latter, given mortgage interest payments, an increase in the interest rate could have a positive, direct effect on inflation. Even if, in practice, a number of fixes are used to neutralise these impacts, both methods are more difficult to implement than NA and RE for the euro area.
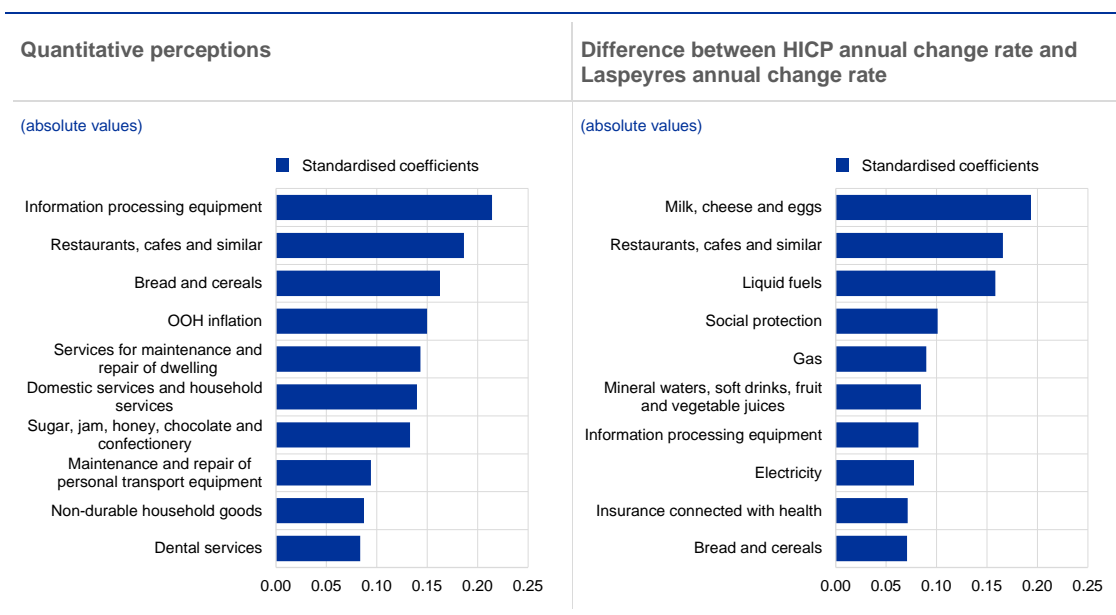
Box 10

Consumers' inflation perceptions and OOH costs in the euro area

ECB Listens events suggest that the cost of housing has significant effects on people's perceived levels of inflation. In some instances, the public voiced the need for a more realistic representation of housing costs, including those of homeowners, in the ECB's measure of inflation. However, the literature on the role of OOH costs in explaining consumers' inflation perceptions, which are typically higher than observed inflation, is scant and provides mixed evidence. House price inflation, as a proxy for OOH costs, is one of the factors considered to explain perceived inflation. Research shows that house price inflation is a significant determinant of inflation perceptions (Döhring and Mordonu, 2007; Abildgren and Kuchler, 2019; Zekaite, 2020b). Nevertheless, the size of the role that house prices play varies across studies. For example, Del Giovane et al. (2009) find that recent experience in dwelling transactions is irrelevant for inflation perceptions. Other authors analyse the perception gap and argue that OOH costs do not help explain it (Aucremanne et al., 2007).

The analysis in this box follows the methodological approach set out in Zekaite (2020b) to examine the role of OOH costs, measured using OOHPIs, in euro area consumer-perceived inflation. The results show that OOH inflation is one of the important predictors of quantitative inflation perceptions but does not play a significant role for qualitative perceptions. The model regresses perceived inflation on lagged inflation rates of a range of pre-selected HICP components (see Zekaite, 2020b, for details) and on euro area OOHPIs. The most relevant predictors are then identified using elastic net regularisation, which is a combination of the LASSO and ridge methods. OOH inflation is found to be among the top ten predictors, while actual rents for housing are less important, albeit still significant. Over the sample period the average OOH inflation contribution to the median perceived inflation rate is 0.7 percentage points (the full-sample average of median perceived inflation is 6.3%). In the "ordinary least squares" (OLS) model, where the top ten predictors are the only determinants, the contribution is 0.5 percentage points The difference between the fitted values from the models with and without OOH is zero on average but varies from -0.2 to 0.2 percentage points (-0.5 to 0.5 percentage points with OLS). The results are similar if no pre-selection is carried out and 94 HICP sub-indices, mostly at four-digit level, are used. Regarding qualitative perceptions, the food and energy HICP components are the most significant predictors. Chart A shows the top ten predictors in the models for quantitative and qualitative inflation perceptions with no pre-selection.

Chart A

Top ten predictors of consumer inflation perceptions

Quantitative perceptions	Difference between HICP annual change rate and Laspeyres annual change rate

(absolute values)

■ Standardised coefficients

Information processing equipment	
Restaurants, cafes and similar	
Bread and cereals	
OOH inflation	
Services for maintenance and repair of dwelling	
Domestic services and household services	
Sugar, jam, honey, chocolate and confectionery	
Maintenance and repair of personal transport equipment	
Non-durable household goods	
Dental services	

0.00   0.05   0.10   0.15   0.20   0.25

(absolute values)

■ Standardised coefficients

Milk, cheese and eggs	
Restaurants, cafes and similar	
Liquid fuels	
Social protection	
Gas	
Mineral waters, soft drinks, fruit and vegetable juices	
Information processing equipment	
Electricity	
Insurance connected with health	
Bread and cereals	

0.00   0.05   0.10   0.15   0.20   0.25

Sources: Eurostat, European Commission, Lünnemann and Wintr (2021) and own calculations.
Notes: Left panel: standardised coefficients in absolute value of ten predictors of quantitative inflation perceptions with the largest coefficients. Right panel: standardised coefficients in absolute value of ten predictors of qualitative inflation perceptions with the largest coefficients. For comparison reasons, the coefficients in both charts are based on the model where inflation perceptions are regressed on 94 sub-indices of the HICP. If pre-selected sub-indices are used in the case of quantitative perceptions, the results are similar to those in the left panel. Estimation method: elastic net regularisation. Sample period: March 2004- March 2020.

4.2   The NA and RE approaches

In the NA approach, the purchase of a dwelling is recorded as consumption at the time the transaction takes place, as is done with other durable goods (e.g. cars). In this sense, it is aligned with the overall HICP concept. As the motivation for the purchase is usually not recorded, there is no distinction between the household sector buying a new house as investor or as consumer. Once a dwelling has been acquired by the household sector, all further, secondary, transactions between households involving this dwelling will not be considered as consumer costs. According to Eurostat, in the euro area the share of home sales that are new to the household sector is smaller than the share of the market for existing dwellings (Eurostat, 2017).

The introduction in the consumer basket of an item whose purchase could be partly motivated by investment considerations could render NA-based OOH price indices prone to asset-like behaviour. In principle this could have consequences for the conduct of monetary policy (see Box 11). In practice quantifying the relevance of the asset component would be challenging (see Section 4.3.2). One suggested way of dealing with this problem is to exclude changes in the value of land from the house price index. However, for all practical purposes distinguishing the price of the structures from the price of the underlying land implies complex measurement issues and may reduce the informational content of house prices (see Box 12). In the short to medium run, the building and the land

are effectively bundled goods, the price of which cannot easily be separated other than in specific cases, such as self-built dwellings. One alternative is to estimate the OOHPI starting from construction costs, adjusted for builders' margins, thus leaving out the location element. But location is a very important housing characteristic that affects prices and often determines the difference between the price dynamics of existing and new dwellings (new dwellings usually being outside urban centres, for example). Further, construction output price indices are typically compiled from survey data, with surveys currently conducted once per quarter. Currently, these indices are compiled only in some euro area Member States. Comprehensive and monthly reporting would require extending and adapting the surveys accordingly.

Under the NA approach, the weight of the components of the OOH price index could vary significantly with the construction cycle. This is because purchases of newly built dwellings, while being a small part of the entire housing stock, are an important element of the OOH price index. The related index weight can be expected to vary considerably over the construction cycle, reflecting both quantity and price effects.[150] As such, the weighting shares may also vary significantly over time across countries and locations (e.g. big cities vs. others).

Underlying the RE approach is the premise that households, as house owners, have a dual role: they are both producers and consumers of housing services and there is only an implicit (shadow) price of the housing services produced/consumed. Therefore, such prices are generally not included in transaction-based consumer price indices, as they involve no purchases during the reference period. When they are included, the most common way of pricing OOH services is by imputing what the market rent would be for an equivalent dwelling in the same area (also called "owners' equivalent rent" or "OER"). This is the approach taken in the national accounts of most euro area Member States.

There are significant measurement difficulties in applying the concept of OER. NSIs will need either to survey owner-occupiers or to match the characteristics of an owned dwelling with a comparable rented dwelling in the same or equivalent location. This can be particularly difficult when the rental market is thin or segregated from the market of owned dwellings. In some cases, an OER approach might in practice not be implementable.[151] Indeed in national accounts, when the share of actual rents in the economy is small and below some conventionally accepted threshold, the RE approach is replaced by the user cost approach. Close matching of owner-occupied and rented units becomes all the more difficult when housing market conditions vary substantially across locations.

Some euro area countries include RE-based OOH indices in national CPIs, providing a basis on which to assess the approach. In Germany, the CPI rent index (including both actual and imputed rents) closely resembles the HICP rent index (covering only actual rents). This is also the case in the Netherlands. Notably, the presence of long-term rent contracts and of regulations that protect sitting

---

[150] The housing price cycle in principle differs from the construction cycle, although they tend to be significantly correlated.

[151] See Eurostat (2017b).

tenants from fluctuations in housing market conditions implies that an OER-based index could differ substantially from what a rent index for the new tenants only may look like.[152] Rent stickiness is common in many countries. This may be a result of regulations, or of contractual clauses that link rent increases to inflation (Hoeberichts and Zekaite, 2020) or to market practices.[153] The overall rent inflation index for the euro area is rather flat, close to or below 2% (see Chart 6), and its evolution differs significantly from the dynamics of house prices.

Chart 6

HICP actual rent index and house price index (euro area)

(year-on-year percentage changes)
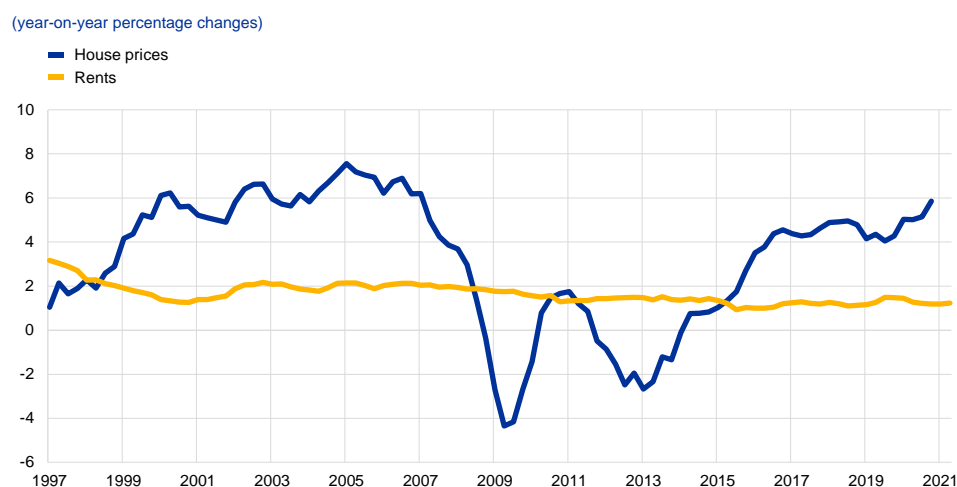
Sources: ECB calculations and Eurostat.
Note: The latest observations are the fourth quarter of 2020 for house prices and the second quarter of 2021 for rents.

Box 11

Implications of an asset-like component in the monetary policy target inflation index

An important consideration for the conduct of monetary policy is that an OOH component computed according to the NA approach would introduce an asset component to the HICP basket. A durable consumption good like a house is different from other shorter-lived durable goods, such as cars, whose value predictably decreases as time goes by. Housing, however, like other long-lived assets, can be subject to large price fluctuations, both upwards and downwards. Housing prices reflect changes in fundamentals such as demographics or local conditions, which are also reflected in actual and imputed rents. But they are also likely to respond to a variety of monetary and financial factors: changes in risk appetite, interest rates, or the beliefs of market participants, which might be

---

[152] Arguably, an index referring only to new or renegotiated rents would better reflect current housing market conditions. In Germany, rents for new lettings track the housing price cycle more closely than is the case for existing rent contracts. However, the use of only new or renegotiated rent contracts to impute the OER would present additional measurement difficulties, in particular in terms of small sample sizes and sample selection biases. There are also conceptual issues on the suitability of this measure for an OER index, as both tenants and owners tend to live in the same dwelling for a certain amount of time, and therefore retaining some rent stickiness might be a proper way to measure such prices. In the United Kingdom, the ONS uses both new and existing rent data to impute owner-occupiers' rents in their preferred CPIH inflation measure.

[153] Rent control intensity has generally declined since the 1980s, but rent controls have recently been strengthened in several countries, for example Germany and France. While the literature on the impact of rent controls on actual rents is quite limited, results for Germany and Ireland show significant effects of rent controls on actual rents and rent increases.

exuberant in the presence of bubbles. These financial factors would then be directly reflected in the HICP.

From a monetary policy point of view, the inclusion of an asset price in a central bank's target has long been debated. On the one hand, there have been arguments for using interest rate policy "to "lean against the wind" in the case of asset price booms, and some hold the view that asset prices should be considered as part of the price level (see Goodhart (2001) for an overview of the theoretical and empirical arguments underpinning this view). However, the conventional view is that financial developments are important for policymaking and warrant policy responses, but macroprudential policy, rather than monetary policy, is the right tool to directly address them. Including a volatile asset component in the target inflation measure might blur the lines between macroprudential and monetary policy and imply some trade-off between having a more representative inflation measure and its informational content for the conduct of monetary policy.

This trade-off might be problematic for policymakers should the asset component steer inflation up or down significantly at the euro area level (for example in the case of a synchronised housing-price boom or bust cycle) while the remaining components of inflation send out opposing signals. In general, since the credit and housing cycles that usually drive the OOH component are long, a well-developed analytical framework is required to be able to interpret such developments and calibrate policy appropriately. This would pose the challenge of engineering appropriate policy actions to respond to fluctuations in housing prices, which affect medium-term inflation, without causing detriment to the anchoring of inflation expectations.

Further challenges may arise during the implementation period. For instance, market participants could interpret the switch to a new target, consisting of the HICP augmented with an OOH component based on the NA approach, as an implicit weakening of the central banks' commitment to price stability in favour of financial stability considerations. There is a debate as to whether such a rebalancing is desirable, depending on circumstances. However, it would represent an important change from the current policy framework, and should be acknowledged as such. Overall, a decision of this nature would benefit from an explicit and targeted communication effort in order to avoid sending mixed signals to the market.

In a broader sense, both issues bring to the fore the need to provide clear communication on the reasons for the decisions taken and to address the implications for interactions between different policy areas and for external communication and the debate on the monetary toolkit.

Box 12
The treatment of land in OOHPIs

A household's decision to buy or sell a dwelling is determined both by its need for housing services and by its expectations of future returns. It is therefore partly a consumption and partly an investment decision. There are two main ways to approach the issue of including house prices in a price index. The first is to accept that housing costs also include investment expenditures. The second option is to measure house price inflation excluding any speculative elements not arising from changes in fundamentals.

The idea of separating the cost of building structures from that of land and assigning the properties of a durable consumption good to the former and the properties of an investment good to the latter has a long history (ILO et al., 2004). The main argument whereby the exclusion of land may be justified is to exclude the most volatile component of price changes for dwellings.[154] [155] However, if the objective is to "smooth" volatility in changes of house prices, then there are other statistical methods for doing so, which are easier to implement, understand and communicate. A second argument is that, while excluding land would probably eliminate a large part of asset price inflation, it would also disregard components related to fundamentals, since land prices also react to demand and supply shocks and include a location component.[156] Third, the exclusion of land costs involves severe measurement issues, as structures and land are to all intents and purposes bundled goods, at least in the short to medium run. This is related to the issue of land-location confounding factors and the representativity of the land data used as input to separate them, as markets for land are very thin in some locations.

In practice, one way of decomposing the cost of building structures from that of land could be to use supplementary information. Instituto Nacional de Estatística (Statistics Portugal, 2009) investigated several regression approaches, some of them used as input price data for residential land in Finland's indices. However, convincing separate price indices for structures and for land were not found. Diewert's "Builder's Model" (2013) was estimated for specific segments of the housing market in Tokyo (2015) and in Ottawa (2020). These studies suggest that observable data, such as construction cost indices, are needed for addressing statistical issues. However, these data may not match the actual, unobservable component in a house price index.[157] Representing purchases of dwellings new to the household sector, as required by the NA approach, requires that all relevant market segments be covered. Applying the Builder's Model to such data is likely to introduce more biases than clean information.

A somewhat different but related approach is to use, from the outset, price indices which do not include land cost, as was done by the Australian Bureau of Statistics.[158] These indices exclude elements related to location so in terms of their construction they are not suited to representing different regional characteristics. Construction cost/price dynamics may differ substantially from actual price movements on housing markets, both over time and across locations. Hence, in the context of the euro area, they would not be easy to explain to the public and might bear very little resemblance to house prices dynamics as observed by households.

At present there is an unfulfilled need for representative data on land prices that could foster a better conceptual and practical understanding of house prices. House price statistics could benefit from more long-term joint research projects (for example with the ESS, the IMF and the OECD) to shed more light on optimal measurement methods. At the current juncture further substantial

---

[154] This is because construction land is typically very inelastic in the short to medium run, and often also in the long run.

[155] Another argument often mentioned is that land represents the investment element because it does not depreciate. However, this view is controversial. It could be argued that building structures also contain an investment element, even though they deteriorate over time. This interpretation is applied, for example, in national accounts where residential buildings form a component of gross fixed capital formation.

[156] Extracting land prices from house prices would also remove volatility due to demand, location, demographics, etc.

[157] A purchase price index for dwellings broadly covers different locations and various sizes and qualities of dwellings. By contrast, construction price indices reflect changes in prices of dwellings whose characteristics are set to be constant over time.

[158] See Australian Bureau of Statistics (2017).

analytical research is clearly required before concluding that separation should be a requirement for OOHPI indices. At this stage, pursuing such a separation could raise more conceptual, measurement and communication issues than it would potentially solve.

4.3 Assessing OOH-augmented HICP indices

4.3.1 Eurostat OOHPIs

Since 2014, Eurostat has compiled and published an NA-based OOHPI for all euro area countries except Greece. The main, but not sole, components of this index are the expenditures on the purchase of new dwellings, including transaction-related costs (e.g. taxes and estate agent fees). OOHPIs are quarterly and are disseminated one quarter and one week after the end of the reference quarter (about 95 days after the end of the reporting quarter). In addition to the acquisition component (which, in the euro area as a whole, accounts for a weight of about 80% on the index), the OOHPI also includes an ownership component (about 20%), which records expenditure related to owning and maintaining the dwelling (such as repairs and maintenance, and insurance on the dwelling). The OOH-acquisition index can be broken down into four main components: acquired new dwellings, acquired existing dwellings that can be included in the owner-occupancy framework, self-built new dwellings and major renovations, and services related to the acquisition of a dwelling.

The composition of OOH-acquisition indices is heterogeneous across euro area countries. In several member states self-built houses have a large weight (almost 80% in Germany, 66% in France and 60% in Italy), while in other cases the greatest weight is associated with the "acquiring new dwelling" component (45% in Spain). While prices indices for the acquisition of new dwelling include the price of land, the indices for self-built houses do not. As a result, the OOHPI acquisition sub-index mainly reflects changes in housing prices (including land prices) in countries such as Spain and the Netherlands but is more akin to the implicit residential investment deflator (excluding land) in France, Germany and Austria. While reflecting the different underlying properties of the new dwellings, this implies a limited degree of harmonisation in the construction of the index across the ESS, related to the different treatment of land prices. Current work on the legal basis for compiling OOHPIs and HPIs in the EU (OOHPI/HPI Implementing Act) is expected to strengthen methodological harmonisation on the price side of the index.

4.3.2 HICP augmented with an NA-based OOH measure

Including the OOHPI in the HICP would yield an augmented index (HICP-H). To publish a euro area-wide HICP-H, the weights for all jurisdictions would need to be

computed. The statistical challenges surrounding this issue suggest that this step still requires some efforts.

At present, the period covered by OOHPI time series is short, which poses challenges regarding its validity. Specifically, the OOHPI series provided by Eurostat only started in the first quarter of 2010. This means that the HICP-H measures are currently of limited research use as, in order to evaluate the properties of an augmented index of this type, a period covering at least a complete house price cycle should ideally be considered. Similarly, to assess whether the inclusion of an HICP-H measure would lead to different policy signals than the HICP, as suggested by monetary policy rules, the sample should include periods of both decreasing and increasing policy rates and extend to periods in which unconventional monetary policy tools were not present.

The results of a back-casting exercise on OOHPIs suggest that including a measure of OOH in the HICP would have had rather a small impact on average inflation in the euro area over the period 1999-2019. Lünnemann and Wintr (2021) back-cast the OOHPIs to 1999, exploiting the close correlation between OOHPIs and national house price indices, which are available for longer periods (Chart 7). With these longer estimated OOHPI series, they derive an aggregate HICP-H starting in 1999. While including an OOHPI measure in the HICP would have had a small impact on average euro area inflation over the last two decades as a whole, it would have led to persistently higher inflation in recent years (0.2-0.3 percentage points on average from 2018 to 2020). In addition, including OOH in the HICP would have increased the dispersion of inflation across countries, reflecting the sizeable impact of OOH measurements on measured inflation around the peaks and troughs of housing cycles in countries that experienced pronounced swings in house prices.[159]

---

[159] For example, housing prices in Spain were declining by about 15% year-on-year in 2012. This would have substantially decreased HICP-H inflation with respect to HICP inflation in that country, by 1.5 to 2 percentage points, depending on how weights are measured (Kalantzis et al., 2020a). Even stronger differences between HICP and HICP-H (about 3 percentage points) would have been recorded in Lithuania and Estonia in 2006 (not members of the euro area at the time), when the annual growth rate of residential property prices in the two countries exceeded 40% in annual terms, and in 2009, when house prices collapsed by more than 30% (Lünnemann and Wintr, 2021). For Germany, however, an experimental back-calculation of the OOHPI results in differences of around +/- 0.2 percentage points between HICP and HICP-H during the period from 1998 to 2019, with an average difference of 0.0 percentage points (Herzberg et al., 2020).

Chart 7

OOHPI and HICP (euro area)

(year-on-year percentage changes)
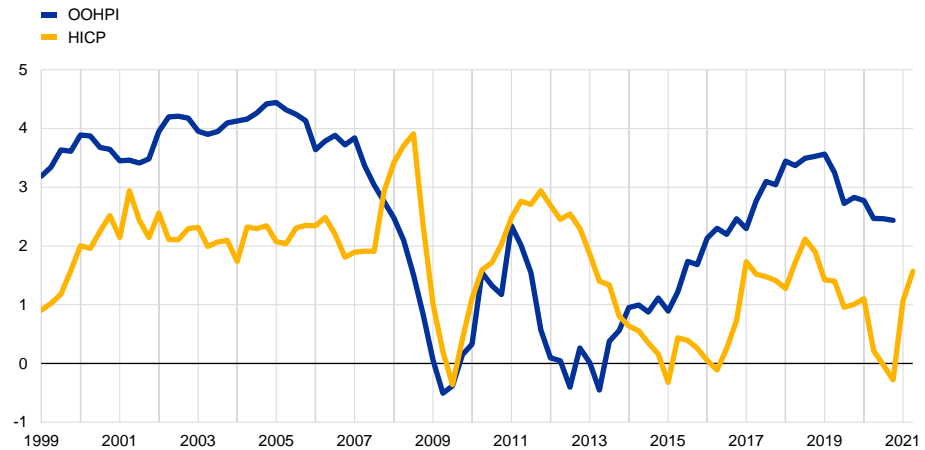
- OOHPI
- HICP

Sources: Lünnemann and Wintr (2021) and Eurostat.
Notes: OOHPIs are back-casted for 1999-2010. The latest observations are the fourth quarter of 2020 for OOHPI and the second quarter of 2021 for HICP. For HICP, the latest observation reflects the average for the available months (i.e. April 2021 and Eurostat's flash estimate for May 2021).

In the hypothetical case of a synchronised housing price cycle in the euro area, including OOHPIs in the HICP could lead to a more sizeable and asymmetrical impact on measured inflation. The limited average impact of OOH on the HICP owes much to the overall small degree of synchronisation of housing prices across euro area countries to date. Therefore, Lünnemann and Wintr (2021) simulate a scenario in which a majority of euro area countries (accounting for about 80% of euro area GDP) record housing price growth at the 80th percentile of their historical distribution. In such a scenario, the annual rate of change of the HICP-H would be 0.4-0.5 percentage points higher than that of the HICP. Conversely, in a synchronised housing price downturn at the 20th percentile of each country's residential property price change distribution, the year-on-year change of the HICP-H would be around 0.1 percentage points lower than that of HICP inflation.[160]

Identifying how large the asset component is in OOHPIs and to what extent it drives the difference between the HICP and the HICP-H is challenging. First, it is difficult to capture the asset price component of a house price index conceptually in an appropriate, meaningful and sufficiently accurate manner, as previously discussed in this Occasional Paper. This is especially true when land price data at the country and euro area level are lacking. Second, what constitutes an asset in this framework is to some extent arbitrary (as structures could be included, according to some definitions). With these important caveats in mind, an extremely rough approximation could be provided by selecting those subcomponents of the OOHPI

---

[160] The estimated impact of synchronised housing cycles on HICP-H is subject to uncertainty and caveats along several dimensions. First, it is based on an estimation of the historical relation between residential property prices and OOHPI, which might change, for instance, if statistical offices change their procedures or the data-generating process changes. Second, HICP is assumed to be 2% and the weight of the OOH component in HICP 12%. Third, the historical distributions may not be representative of a housing cycle.

that in principle should display asset-like dynamics,[161] and computing their contribution to the headline OOHPI. Such an exercise, admittedly coarse, suggests that in 2014-19 these components, whose weight is about one-third of the total OOHPI, are responsible for about half of its average year-on-year growth and slightly more than half of the average difference between the year-on-year growth rates of the HICP-H and HICP (see Section 4.3.4). As a comparison, an upper bound for the asset component would be represented by the whole OOHPI acquisition item, which has a weight of almost 80% of the total OOHPI and contributes about 90% to the difference between the HICP-H and the HICP.[162]

To inform monetary policy in a more timely manner, a quarterly time-lagged HICP-H would need to provide a reliable medium-term forecast and would need to have the potential for gauging inflationary/deflationary pressures, especially with regard to turning points in the housing price cycle. The HICP-H could be used in the monetary and economic analysis in a similar way as other variables, with at least one additional quarter to be forecast for HICP-H because of the publication time lag. Using simple autoregressive models, van Overbeek and Hoeberichts (2020) suggest that for the period from the first quarter of 2014 to the third quarter of 2019, the OOHPI can be predicted one quarter in advance. The prediction error in the HICP-H is relatively small compared with the loss in representativeness caused by not including the OOHPI, measured by the difference between ex post HICP-H and HICP. This applies to both headline and core inflation. Considering the six-week frequency of Governing Council monetary policy meetings, further efforts might be needed to improve HICP-H forecasting models. The aim here is to enhance their precision and provide monthly predictions at different time junctures, especially when the publication lag for the OOHPI with respect to the monthly HICP is significant.[163]

4.3.3 HICP augmented with an RE-based OOH measure

The alternative method for incorporating the costs of OOH in the reference inflation measure would be to use an RE-based method and compute an HICP augmented with an OER component (HICP-R). No official pan-European method for an owner-occupied RE-based statistical method currently exists in the context of CPIs. The implementation and harmonisation of such a method could only be accomplished by a change in the HICP framework, achieving which would probably be a lengthy process.[164]

---

[161] Namely, the items "purchases of new dwellings", "existing dwellings new to households" and "services directly related to the acquisition" (i.e. the whole acquisition component without the item "self-build dwellings and major renovations").

[162] An alternative rough and imprecise, albeit useful, approximation of the size of the asset component could be made by comparing the OOH index based on the NA approach to that based on the RE approach (no asset component). The euro area does not have a properly compiled OOH index following the RE approach, but the ONS produces an index of this type for the United Kingdom, together with the OOH index based on the NA approach. On average, the asset component is around 40% in the OOH-NA for the United Kingdom and accounts for almost all of its variance.

[163] Assuming that, once available, the quarterly OOHPI could be decomposed in monthly indices, the lag with respect to the HICP could vary by three to five months.

[164] A fairly harmonised approach exists in national accounts. However, this approach is not granular enough to fully capture the changes in housing costs within the different locations in each country.

The main advantage of using imputed rents is that timeliness and frequency issues are unlikely. Monthly indices should be relatively easy to construct with the same timeliness as the HICP. Moreover, an HICP-R would not introduce asset prices to the inflation measure. Constructing an HICP-R would, however, pose non-trivial challenges. As explained in Section 4.2, the HICP-R would depart conceptually from the HICP framework, given that prices are not principally based on market transactions, and, in practical terms, it would be difficult to compute RE indices in countries with a thin or heavily regulated rental market.

The need for OERs of a rent market that is representative both in depth and in composition in order to reliably impute equivalent prices in OOH indices could give rise to policy-related issues. When the base is narrow and possibly subject to rent controls, there is a risk of creating a potentially unwanted link between government-controlled rents and the central bank's inflation target. Indeed, the imputation method used for the OER could have a large impact on inflation in the presence of administrative regulations that have minor or no overall economic effects.

Euro area citizens might view the limited impact on HICP of including the OER as unsatisfactory, with regard to including housing costs in the target inflation index. This would apply especially if house prices were rising (or falling) significantly, while rents were not changing much.

4.3.4 Comparing the HICP-H and HICP-R

Chart 8 shows a comparison of the HICP-H and HICP-R with the HICP. To proxy HICP-H and HICP-R, assumptions on the weights had to be made. The weight of the OOHPI in the HICP-H has been set to 13%, reflecting the estimated share of imputed rents in private consumption expenditure; the actual weight of the OOHPI would in fact probably be lower when properly computed, at 9 or less%.[165] The HICP-R was obtained by expanding the weight of the actual rent component in the HICP from around 6% to 17%.[166] Indeed, in the absence of a true OER comparable to the OOHPI, this measure should be regarded as a provisional HICP-R proxy. As such, its similarity to the true OER would crucially depend on the method agreed to

---

[165] In general, with the HICP, the OOHPI and an imputed rent index at hand, it will be possible to compile the HICP including OOH according to the NA and the RE approach, if the following three expenditure categories are annually available at current prices. First, household final monetary consumption expenditure incurred on the domestic territory (as the scope of the HICP). Second, household expenditure on self-builds and the net acquisition of dwellings (excluding land) by the household sector, with all dwellings built and purchased for owner-occupancy, as well as major repairs, maintenance, insurance, fees and other services related to the ownership of dwellings (as the scope of the OOHPI). And third, the dwelling services of owner-occupiers estimated by linking the actual rents paid by those renting similar properties in the rented sectors to those of owner-occupiers.

[166] See Eiglsperger et al. (2020) for details on the weight computations. Kalantzis et al. (2020a) studied how sensitive HICP-H might be to different OOHPI weights. They define three alternative weighting schemes for the euro area, based on the share of imputed rents and housing investment, and find that the resulting variants of HIPC-H are relatively similar in 2011-19. This could suggest that measurement issues related to OOH-weights might be benign in practice for the euro area. However, it should also be noted that in the period under consideration there was no synchronised construction and housing price cycle, which would have made the weights more variable for the whole of the area.

compute the OER. The approximation in the chart is good if OER is based on the average of actual rent data, less so if based only on new rents.

Based on these preliminary computations the HICP-H would have a higher amplitude but fairly similar average (taking data from 2011) compared with the HICP-R. Importantly, in the period considered the RE component leads to a dampening of deviations of the HICP-R from the target (the rent component having always been close to 2%) and the NA component to an amplification of cycles in the HICP-H. This implies a more important role of economic slack in Phillips curves for the HICP-H. Moreover, if coupled with non-synchronised housing price cycles across euro area countries, it also entails a larger risk in the future of cross-country inflation dispersion. Volatility could also be passed on to other sub-components of the HICP through indexing. As mentioned in Section 4.3.2, there are a number of non-trivial issues regarding the consequences that new indices could have on ESCB procedures and analyses – especially in relation to forecasting and monitoring. For example, the stronger de-coupling of the HICP-H and the private consumption deflator should be taken into account for forecasting models.

Chart 8
HICP, HICP-H and HICP-R

(year-on-year percentage changes)

- HICP
- HICP + OOH
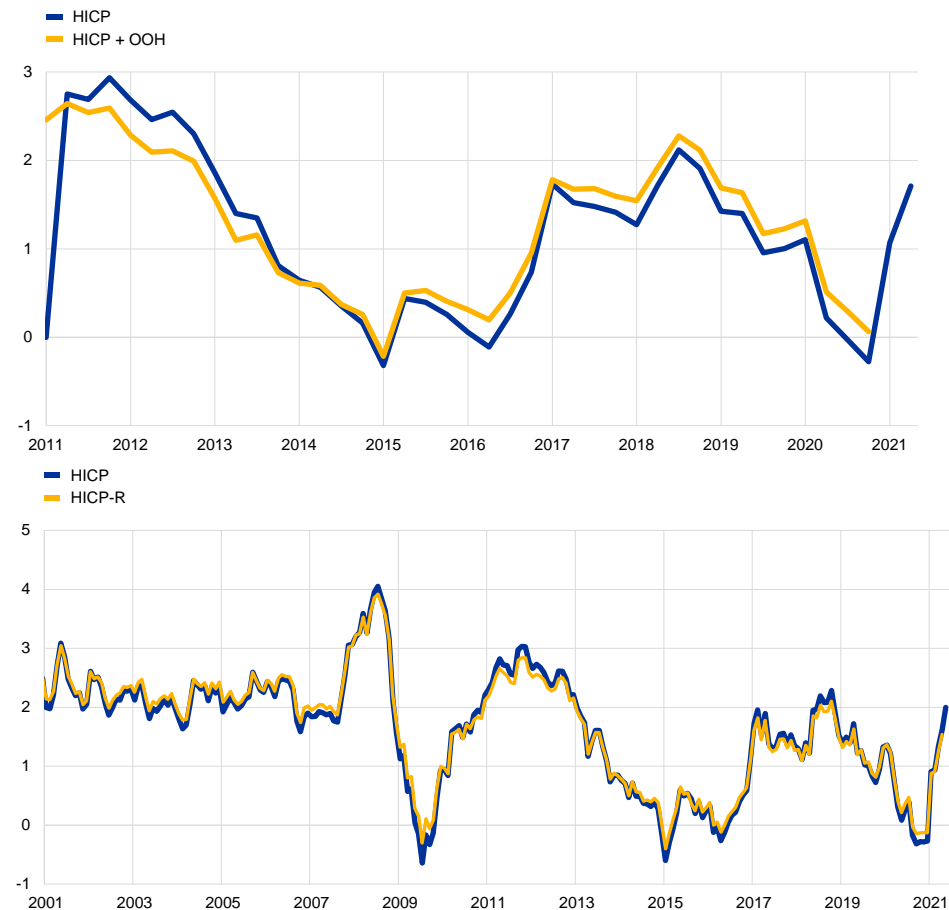
- HICP
- HICP-R

Sources: ECB calculations and Eurostat.
Notes: HICP-H refers to HICP augmented with Eurostat OOHPIs; HICP-R refers to HICP augmented by increasing the weight of actual rents from 6% to 17%. The upper panel reflects quarterly data while the lower panel reflects monthly data. For quarterly series, the latest observations are for the fourth quarter of 2020 for HICP + OOH and the second quarter of 2021 for HICP. For HICP, the latest observation reflects the average for the available months (i.e. April 2021 and Eurostat's flash estimate for May 2021). For the monthly series, the latest observations are April 2021 for HICP-R and Eurostat's May 2021 flash estimate for HICP.

Public communication on Governing Council decisions regarding new indices will pose challenges. The nature of these challenges will differ depending on whether the indicators are taken as the target measure of inflation, or if they are considered as auxiliary indicators for monetary policy. This is the case even leaving aside the issue of frequency (HICP-H being published when the monthly HICP is already available for part of the reference quarter) and timeliness (HICP-H referring to a quarter already fully concluded).

The use of the HICP-H as an auxiliary analytical indicator has some clear advantages and some drawbacks. The advantage for communication purposes is that the OOHPI refers directly to observed transaction prices for new dwellings that are clearly perceived as a cost by the household sector as a whole.[167] It is therefore easy to explain and probably reflects housing market conditions more closely than

---

[167] Households tend to purchase dwellings from other households, but as long as the OOHPI is relatively close to the HPI this would not create significant perception issues.

an OER would do in many euro area countries. On the downside, the lower weight of OOH using this approach than in the RE approach, and its dependence on construction cycle activities, could also be a communication hurdle, more acutely felt if the HICP-H is used as the reference for the ECB target inflation measure. Alternative and more stable weights, such as those connected to housing stock-related data from national accounts or those implied by some multi-year smoothing, may therefore need to be considered. Stable weights would be particularly valuable at turning points in the construction cycle.

4.4 Evaluation and recommendations

The inclusion of OOH in the HICP is in principle desirable but poses implementation challenges. The two most suitable measures for including OOH costs in the HICP, based on NA or RE, both involve significant trade-offs.

Weighing the specific pros and cons of the NA and RE methods, the case for the NA approach seems stronger. Given its reliance on transaction prices rather than imputed rents, the NA approach is likely to enhance the information content of the HICP. It may also better address the concerns of euro area citizens. The summary results of the ECB Listens events confirm that the cost of housing has significant effects on people's perceived levels of inflation. The results also suggest that the general public expects a stronger and more realistic representation of housing costs, including those of homeowners, in the ECB's measure of inflation.[168] However, the public's perceptions are to some degree based on a limited understanding of the HICP as a measure of inflation. A careful communication campaign would be needed to avoid raising unrealistic expectations.

The roadmap in Figure 3 illustrates the necessary steps and an expected timeline for the implementation of a quarterly price index that includes OOH expenditures based on the NA approach in the HICP basket. To implement this process, which has implications for monetary policy and communication, statistical and legal aspects would also need to be addressed.

An experimental quarterly price index including OOH expenditure may be published by Eurostat in 2023. This schedule takes into account the time needed to complete the current work of the ESS, particularly on weights for OOH. In the meantime, the ESCB should endeavour to compile proxy indicators for analytical purposes.

---

[168] See "ECB Listens – Midterm review summary report".

Figure 3

Possible roadmap for the inclusion of OOH based on the NA approach
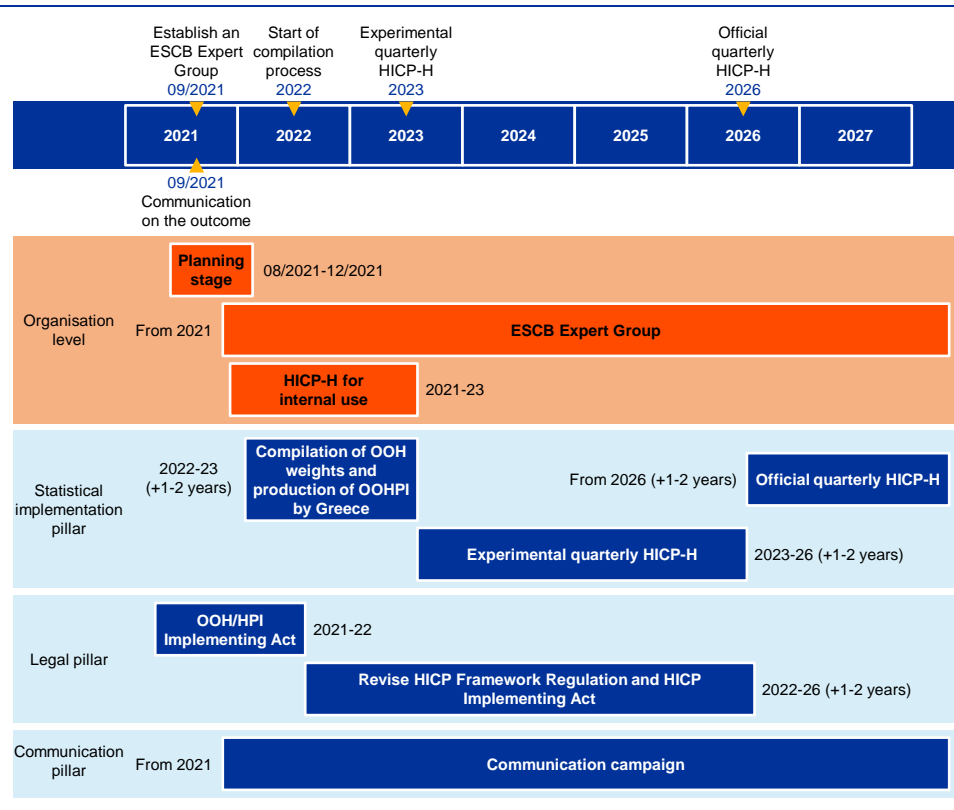
Source: ECB.
Note: Bars reflect the median timeline expectation.

For the ESS to develop and publish a new official quarterly index, the regulations governing the HICP framework would need to be changed, a process that could take an additional three years or so. To reduce implementation delays, work on the legal aspects could start immediately after the ongoing work on the OOHPI/HPI Implementing Act has been finalised (expected for 2022 – see Figure 1).

The time needed to develop and implement an official index that includes OOH could be fruitfully used by the ESCB to investigate the potential monetary policy implications of such an index and design a communication campaign in collaboration with the ESS. After the strategy review, a communication campaign could be launched to inform the public of the forthcoming changes. In addition to existing communication channels such as speeches, the ECB Blog and so on, a dedicated area on the ECB's website could be set up to explain the most significant aspects of the decision. The CES could also be used to gauge the level of consumer awareness and understanding.

Another desirable, yet major and extremely challenging, advance would be the inclusion of OOH costs in the HICP at a monthly frequency and in a timely manner, ideally close to the current release schedule of all-items HICPs. This might require the collection of additional and timely information, for example from new ESS surveys, as well as the more timely collection of data on transactions

involving dwellings, for example using digital data. The application of estimation techniques could also be explored, also in forms not yet established in HICP compilation. This suggests that the ESS might have to further develop the HICP measurement framework and enter new territory in HICP compilation, similar to the approaches applied in national accounts and for flash estimates (GDP, HICP). This would require an adaptation of the HICP's legal basis. ESCB experts would be available to support the ESS in creating estimation methods of this type. The ESS and the ESCB, in their respective fields of competence, would have to assess whether feasible monthly data estimations are sufficiently accurate and reliable to have the status of official statistics and be used for monetary policy purposes.

The compilation of an official quarterly HICP-H for the euro area should not be seen as a final outcome with regard to correctly measuring price pressures deriving from house prices. The need remains for representative land price data that could lead to a better conceptual and practical understanding of those prices. Long-term joint research projects with the ESS, IMF and OECD would shed more light on optimal measurement methods and should be encouraged, with the aim of better isolating the asset component in NA-based OOH indices. A reduction in the current long time-lag in data availability (95 days) should be also be pursued, and the feasibility of a monthly HICP-H explored in depth.

5 Measuring the cost of living

Previous chapters reviewed the reasons why the HICP was chosen as the most suitable measure of inflation for quantifying the ECB's price stability objective, its improvements over time and important areas for further enhancements. This chapter discusses the measurement of the cost of living in general and different types of COLI, the effort that would be required to construct such an index for the euro area and how a COLI could enhance the economic analysis underlying the ECB's monetary policy.

COLIs are theoretically appealing because they derive from welfare theory, but they have practical limitations. They are considerably more comprehensive in scope than HICPs, but three measurement issues related to imputation and revisions speak against using COLIs as a policy objective.

First, producing a timely COLI requires past and current weights, which is a hard requirement to meet. Current expenditure weights can be estimated but this requires enhanced and more frequent/timely data collections than the HICP as currently specified. A COLI of broader scope would also require more data. These can be estimated and extrapolated, but this may also entail extra work that would delay the publication of the COLI relative to the HICP. In addition, the use of estimates implies revisions once the final expenditure data become available.

Second, a COLI is not very precise, especially at shorter horizons and if defined very broadly. Prices and quantities of non-market goods must to a large extent be estimated. Some publicly provided goods might be valued at cost (as in national accounting), or non-market prices and expenditures could be extrapolated over shorter time horizons. Thus, especially at short horizons, the signal-to-noise ratio of a broad COLI is lower than that of more focused indices such as the HICP.

Third, any feasible COLI is not bias-free. The COLI could in principle be defined as the gauge against which to measure substitution bias in the HICP but feasible COLI implementations are not themselves free from bias.[169] In particular, the more frequent updating of expenditure weights requires attention to avoid chain drift, as discussed in Chapter 3. Moreover, compared with an ideal COLI, any feasible, actual, COLI might not be able to cover the full scope that would be desirable in theory.

A single COLI representing average consumption patterns does not address all welfare concerns. The theoretical appeal of a COLI is to represent welfare, but preferences are heterogeneous across households, and so are their living standards

---

[169] A feasible COLI is in practice similar to an optimally defined COGI, and thus shares many of its advantages and disadvantages. See Box 1 in Chapter 2 and in Chapter 3. Because COLIs tend to be updated retroactively as more information becomes available, COLI estimates based on incomplete early data differ from the "best" estimate at later points in time. In other words, they are typically revised, reducing the vintage bias only over time.

and consumption baskets. A single COLI can only track the welfare of the average euro area consumer.

Taking all this into account, the next two sections assess what building blocks would be needed to compute a single COLI for the euro area and how it could be useful for the ECB's economic analysis in the integrated assessment underpinning monetary policy decisions.

5.1 The effort required to construct a COLI for the euro area

5.1.1 The main building blocks of a COLI: frequently updated weights and non-market prices

Which data are necessary to construct a COLI depends on whether the aim is only to estimate more frequent weights or also to include non-market prices in the index. The two dimensions along which a COLI differs from a COGI, i.e. the frequency of consumption weight updates and the scope of the index, are separate issues and can be treated independently. In other words, it is possible to construct a COLI with time-varying weights that has the same scope as the HICP (for example to help quantify the substitution bias), while treating the inclusion of non-market prices separately.

The construction of current weights requires information on households' spending at the same frequency as prices.[170] To capture substitution effects adequately, past and current weights should ideally cover all categories of household spending at the most disaggregated level possible. If spending data on some categories are unavailable at the desired frequency a statistical model needs to be developed to estimate them.[171] A proof of concept of a time-varying weights COLI for the euro area is presented in Box 13. Using publicly available data to estimate monthly weights during the COVID-19 period in 2020, a feasible COLI approximation was derived, which differed on average by 0.2 percentage points from the HICP over the period. There are two important caveats to this analysis. One is that the only data available were aggregated at a higher level than those at which much substitution takes place,[172] so that this gap could be a lower bound. The other caveat runs somewhat counter to this: as highlighted in Chapter 3 the pandemic period is highly exceptional and substitution across categories has been more extensive than in normal times even at higher levels of aggregation.

---

[170] Expenditure weights can be at a monthly or quarterly frequency.

[171] For example, the Bureau of Economic Analysis (BEA) in the United States collects monthly data using several sources to estimate monthly PCE data. For instance, retail sales data are taken from the United States Census Bureau Monthly Retail Trade report. Monthly estimates of margins on used cars and used light trucks are extrapolated from the annual estimates using data on retail sales of used car dealers from the monthly retail trade survey. For housing services, the BEA employs a model based on demographic trends.

[172] See Box 13 on the cyclicality of upper and lower-level substitution bias.

The inclusion of non-market goods means that suitable proxy data must be collected. Measuring the non-market price of, for example, free public libraries, public health and safety, and potentially even environmental goods, requires either imputation or specific model assumptions. For some non-market services, the standard practice is to value output in terms of the input costs incurred in production.[173] An example of a COLI implementation that includes non-market prices is the US personal consumption expenditures (PCE) price index. Box 14 elaborates on the US experience.

5.1.2 A COLI for the euro area: what is needed in addition to the data collected for the HICP?

Constructing an official euro-area COLI would require additional resources from NSIs and from Eurostat. Based on currently available data and information, three main steps are required to construct a euro-area COLI. The first two steps are methodological: establishing the scope of and the formula for the index and identifying the set (and sources) of non-market prices required. The third step requires additional data collection efforts, i.e. collecting and/or estimating infra-annual, ideally monthly, consumption weights.

The data currently available on household consumption in the euro area are incomplete and not sufficiently timely for constructing monthly weights. At present, the measurement of households' consumption for the purposes of computing HICP relies on national accounts "as well as any available and relevant information from household budget surveys and other data sources which are sufficiently reliable for the purposes of the HICP.[174]" These can include market intelligence information, retail trade data, various administrative sources and, more recently, scanner data, as long as they meet reliability standards. Most of these sources have limitations for estimating current expenditures. For instance, household budget surveys differ across countries in terms of frequency and timeliness, while the national accounts data include quarterly data but only at a three-digit (rather than the five-digit) COICOP level.[175] Therefore deriving COLI weights at the same granularity as those of the HICP would require a considerable investment in data collection.[176] Should the HICP weights become more timely, the COLI would naturally be built on this improvement, drastically reducing the necessary investment.

---

[173] These costs include labour, materials, supplies, and the use of fixed capital.

[174] Article 3(2) of Commission Regulation (EU) No 1114/2010 of 1 December 2010 laying down detailed rules for the implementation of Council Regulation (EC) No 2494/95 as regards minimum standards for the quality of HICP weightings.

[175] Obtaining detailed data at COICOP-5 level for all countries would require annual or infra-annual household budget surveys, and/or expanding national account data to that level of disaggregation. The decision would necessarily take the associated costs into account.

[176] One alternative would be to estimate household consumption patterns using disaggregated data on retail trade and services turnover published by Eurostat (as described in Box 12). The advantage is that these data are published on a monthly basis and become available with a one-month delay for retail trade data and a two-month delay for other services. The limitation is that they are compiled according to the Statistical classification of economic activities in the European Community (NACE Rev. 2, Eurostat). A matching NACE-COICOP would therefore be necessary, with all the problems entailed in matching a production classification to a consumption classification.

Additional data are needed to estimate non-market prices, which means that quality and reliability issues will need to be tackled. The list of non-market prices included in a COLI, and thus the need for additional data to estimate them, depends on the desired scope of the index itself. The natural starting point would be the current national accounts-based approach, which includes non-market prices in the private consumption deflator.[177] One possible direction for broadening the scope beyond the current ESA concepts could be to use producer price indices.[178] In any case, the measurement of non-market prices involves non-negligible measurement issues because some non-market services, such as the quantity of education consumed and its price, are measured on the basis of cost in national accounts. It is even more complicated to measure the quantity consumed and price of, for example, illegal activities.

5.2   The usefulness of a COLI for economic analysis

Many authors have argued in favour of a COLI, highlighting its methodological and conceptual value-added in a wide range of applications.[179] **However, a COLI reveals its full value-added as a complement to narrower indices.** Its benefits, like its implementation costs, would derive from its use of current weights and its broader scope. Using both past and current weights makes it possible to capture the impact of substitution on inflation, which can be particularly large in severe downturns. Having a wider scope makes it possible to take into account the impact of price trends in non-market goods and services that are relevant for consumption decisions. A complementary index of this type could serve as a useful communication tool, embedding the ECB's policy decisions in a broader context and signalling robustness (as long as clarity is maintained regarding the different roles of the HICP as a target and any other price and cost measures, especially when their inflation rates deviate from each other).

This section reviews how the provision of a COLI as an experimental index can help the ECB increase the robustness of its economic analysis of the risks to price stability. It starts with an illustrative example and goes on to review the five most important benefits. The first four aspects relate directly to the use of a COLI as a supplementary indicator for the economic analysis underlying monetary policy, while the fifth considers communication and accountability aspects.

---

[177] In the national accounts, household final consumption expenditure covers the goods and services consumed by domestic households. In addition to actual purchases, it also includes certain imputed purchases, such as consumption by entrepreneurs and the value of OOH. Final consumption expenditure of non-profit institutions serving households (NPISHs) is another part of private consumption and includes the value of the goods produced by them, less capital formation and sales, and, where relevant, expenditure on goods provided to households for their consumption as social transfers in kind.

[178] Insights into practical approaches for including non-market goods and services in a conditional COLI compiled at a monthly frequency are provided by the BEA PCE price index. For non-market products produced by NPISHs, for instance, social assistance services, including homes for the elderly, output price dynamics are often approximated by referring to input costs incurred in their production. In the EU, non-market production is primarily quantified in the context of national accounts. As in the United States, input approaches are typically applied. However, in several cases input is estimated by referring to physical quantities.

[179] See, for example, Schultze and Mackie (2002).

By combining information on market and non-market goods, COLIs can account for their interaction, in particular for substitution and bundles of market and non-market goods. Consider, as an example, the non-market good, public health, and the market good, candy. In normal times, a given amount of candy and a given health risk (which includes hospital capacity and waiting times) provides the household with a certain level of utility.

A deterioration in the quality of non-market goods (or an increase in the effort needed to obtain them) in this framework is like an increase in their price and increases the demand for market goods, thus adding inflationary pressure to their prices as well. In the case of the COVID-19 pandemic, the household could compensate for the utility lost (higher health risk, capacity issues at hospitals, etc.) with additional consumption of, say, baking ingredients. This substitution of a non-market good (public health) by market goods (baking ingredients) might trigger subsequent price increases.

If the non-market prices remain unobserved, the change in the cost of living cannot be measured by the change in the market prices alone: the change in expenditure would be needed. This example vividly illustrates the merits of monitoring a COLI. Frequent updating of weights ensures that the weight on baking ingredients does not become too stale. This alone, however, would not capture the increase in the cost of living and thus the household's additional expenses to barely maintain their living standard. The broader scope of a COLI, in this case the inclusion of public health, provides a theoretical framework which properly assigns the increased baking ingredients consumption to an increase in the cost of living.

Understanding the price, and in particular the quality changes, of non-market goods can help assess and predict price trends in market goods. The origin of consumption shifts in market goods can be long-term trends (e.g. climate change) and short-term fluctuations (e.g. pandemic) in non-market goods. Monitoring non-market goods can therefore help predict changing consumption and hence price pressures in the HICP. This applies not only to the short run, as in the case of an ongoing pandemic. An improvement in the quality of public health after the pandemic might have the opposite, predictable effect. More subtle are the effects of climate change on non-market goods, which in turn affect long-term trends in the demand for market goods, and thus HICP inflation.

How can monetary policy steer non-market goods prices? Transmission of policy effects via costs would be similar to transmission for market goods, while the expectations channel would be more indirect. When non-market prices are measured in feasible COLI implementations from the cost side, transmission via cost is taken into account.

5.2.1   Possible uses at the ECB

By adding to the information set underlying ECB monetary policy, a COLI could help monitor its long-term adequacy and convey a sense of robustness when communicating that policy to the public. A COLI would directly benefit ECB

monetary policy by improving the data basis of the its policy models and supporting robust policy decisions as well as communication. As described in the example in the previous section, developments in non-market goods through substitution affect the demand for and thus the prices of market goods. A COLI is therefore an additional indicator of future price developments.

First, economic models that are routinely used as part of the information set for economic analysis build on the COLI concept.[180] Calibrating or estimating some of these models with a COGI as proxy for the COLI introduces an avoidable source of error which renders interpretation difficult, and might potentially even give misleading results.

Second, an experimental COLI used as a supplementary inflation measure improves the robustness of monetary policy decisions. No single inflation metric captures all aspects of inflation or has all of the desirable features of an inflation index for policy (for example accounting for heterogeneity, distortions, etc., as also described in Chapter 6). If, for example, HICP and COLI inflation pointed in different directions, this could prompt a deeper analysis of the underlying reasons and possibly suggest more cautious policy action than if both indices sent out similar signals.[181]

A COLI would enrich the information set of cost and price indicators used in the economic analysis pillar informing monetary policy.[182] The HICP has many advantages (as outlined, for example, in Chapter 2), and thus remains the policy target variable. But these advantages in terms of transparency and reliability come at the cost of only partially accounting for important aspects of inflation, in particular the non-monetary nature of many of the goods and services that households consume and the substitution bias. It would therefore be useful to have an index that takes these effects into account. Even if a COLI was available on a less timely basis than the HICP and was subject to revisions and to the imprecision generated by imputing non-market prices, it would still allow an ex post comparison of the more timely HICP with the wider-scoped and revised COLI.[183]

For example, the FRS has both COGI-like and COLI-like inflation measures at its disposal and clearly communicates that it targets the COLI-like PCE deflator. The CPI and the PCE price index are published on a monthly basis by two separate statistical agencies in the United States. The Bureau of Economic Analysis (BEA) also publishes the decomposition of the CPI-PCE gap on a regular basis, highlighting the contributions of the different effects (Chart 1). The FRS often discusses comparisons of alternative inflation measures in terms of their

---

[180] These are micro-founded DSGE models where consumption aggregates are based on CES preferences, such as the NAWM (see Christoffel et al., 2008, and Coenen et al., 2020). Many trade models that are routinely used to assess welfare gains from trade are also based on CES preferences.

[181] The robustness gained from cross-checking the economic analysis using multiple indices might make it unnecessary to introduce inflation ranges, which have been suggested by the FRS (Chung et al., 2020).

[182] Further research on a COLI is already envisaged in the HICP framework regulation (Regulation (EU) 2016/792 of 11 May 2016), which suggests in recital 12: "In addition to the HICP, research on a harmonised cost of living index should be initiated."

[183] The revisions themselves might be predictable and state-dependent, which can in itself inform policy decisions (Croushore, 2019).

complementarity in supporting the economic analysis.[184] In addition, the BEA regularly publishes a decomposition of the gap between CPI and PCE inflation in terms of scope, weight, formula and other effects. The blue bars in Chart 1 show the formula effect; this gives an indication of the substitution bias, which has been positive as expected, apart from the two quarters of extreme contraction of GDP in 2008 and 2020. The red bars refer to the scope effect, which in the United States has been mostly negative.

Chart 9

CPI-PCE gap decomposition in the United States
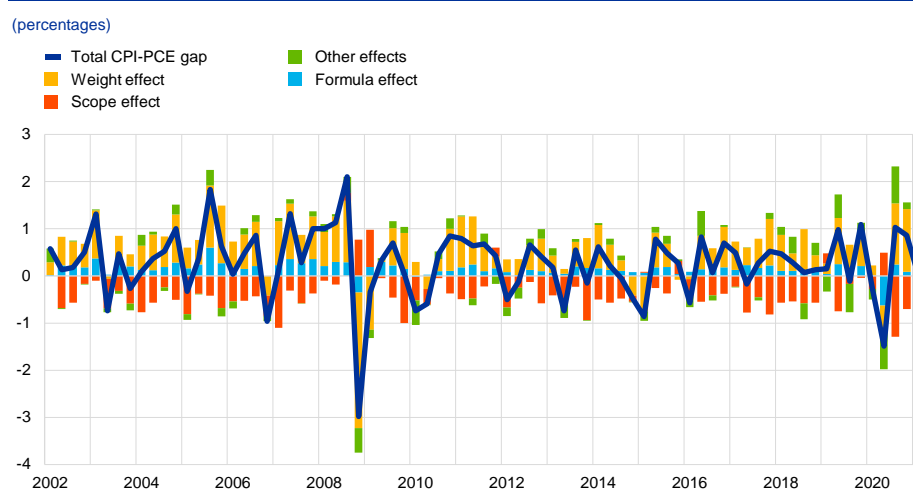
(percentages)

Sources: BEA and NIPA Table 9.1U ("Reconciliation of Percent Change in the CPI with Percent Change in the PCE Price Index").
Notes: The chart shows the difference (the blue line) between the quarter-on-quarter percentage change at annual rate of the (overall) US CPI and (overall) US PCE price index. The bars show the contribution to the CPI-PCE gap of the "formula effect" (blue bars), the "weight effect" (yellow bars), the "scope effect" (red bars) and other effects (green bars). The formula effect is calculated as the difference between the US PCE price index (which is a monthly chain-linked Fisher index) and the US CPI calculated using a Laspeyres index formula. The weight effect reflects the fact that the same item has different weights in the CPI and in the PCE price index. Finally, the scope effect accounts for items that are either out-of-scope in CPI or in PCE space (such as non-market prices or medical prices). The latest observation is for the first quarter of 2021.

If a COLI was available a decomposition of this type would also be available for the euro area, bringing two closely related advantages. First, it would provide continuous, close to real-time guidance on the magnitude of substitution bias in the HICP series, as the formula effect of Chart 1 does for the United States.[185] Second, if the HICP index and the COLI drifted apart this would provide a signal to look more deeply into the components of the wedge: a large change in the formula effect would indicate more intense substitution than usual, possibly indicating a turning point in the cycle. The analysis in Box 15 suggests that the substitution bias may be cyclical; this property could be exploited in the business cycle analysis underlying the assessment of the risks to price stability.[186] Second, a drift in the scope effect would indicate a persistent trend in the relative prices of non-market goods; this information

---

[184] One example is Luciani and Trezzi (2019). Box 12 of this document discusses the experience of the FRS in detail.

[185] To our knowledge, all the statistical literature on substitution bias takes a COLI of some form as the benchmark against which this bias is measured. This is related to the choice of formula, using both current and past weights.

[186] The analysis also has some results pointing to the need for further research on the impact of the utility function assumptions on practical COLI implementation.

could be useful, for example, in the context of infrequent strategy reviews to assess the optimal inflation rate derived from COLI-based models.

Third, a COLI helps to cross-check the impact of ECB monetary policy with longer-term, broader inflation developments. Over the long run, biases in the HICP can accumulate and the prices of non-market components not included in the HICP can change substantially, for example due to demographic, societal and environmental trends.[187] Because the ECB's inflation objective is defined over the medium term, such longer-term developments (if any) must be taken into account, especially in the context of regular strategy reviews, as they might have a bearing on the calibration of the inflation target.[188]

Fourth, a COLI can support communication around monetary policy. The COLI is a considerably broader inflation concept than the HICP, being directly derived from welfare theory. The ability to refer to the COLI on occasion would embed policy decisions linked to the HICP target in a broader context. In crisis events monitoring a COLI would show transient but important substitution effects. This might not trigger a policy action, but adequate communication could reassure the public that policymakers were aware of developments. This was shown in 2020 by the attention paid by central bankers, statistical institutes, the financial sector and even newspapers to alternative indices proposed by academics. In the absence of a reference experimental index the communication on that discussion was complicated. For this reason, the COLI would be a valuable communication tool, strengthening the ECB's accountability and transparency vis-à-vis the public. In this respect the availability of representative timely weights would be of paramount importance.

Another way in which reference to a COLI can contribute to the ECB's public accountability is by complementing the economic analysis with an examination of broader market and non-market price developments. These have – of course – always been important but are now attracting increasing attention from central bankers and the public. The COLI concept easily accommodates items that stretch the HICP framework, such as "paying with data", public health and environmental goods.[189] These have a close link to welfare but are not recorded in monetary transactions. If the ECB monitored the impact of these developments on a broad, welfare-related measure of inflation, this would result in a deeper understanding of the transmission of monetary policy to the consumption decisions of households and support more informed and possibly more robust monetary policy decisions. Knowing that such important price developments are not ignored in the

---

[187] See Costa and Kahn (2003); the ability of a COLI to account for these changes would be related to its feasible scope.

[188] Omitting, for example, environmental goods could lead to understating the welfare-relevant inflation rate if the growth of these prices was higher than the inflation target. In this sense a COLI could potentially be used to assess the ECB's monetary policy stance with respect to the possible risks to price stability posed by non-market prices related to climate change, thus addressing the issue within the monetary policy framework.

[189] Specifically, the concept facilitates discussion of the effect of climate change on the cost of maintaining a given standard of living. The effect of disasters and regulation on consumer prices is already part of the policy discourse (see, for example, the recent blog post by Frank Elderson), but this might be perceived by the public as a deliberately narrow perspective.

information used to reach policy decisions would further enhance the credibility of the ECB.

Finally, as an inflation index external to the HICP framework and derived directly from consumer welfare maximisation, the COLI can increase transparency. Because the proposed COLI is not the policy target, any debate over its calculation can focus on its practical implementation. This might facilitate agreement on a common, fully transparent, bottom-up calculation approach, featuring complete documentation and public access to all non-confidential components.

5.3 Conclusions and recommendations

Overall, an experimental euro area COLI could be a valuable economic analysis and communication tool. It would transparently seek to integrate all prices of consumer goods and services that create utility for consumers (ideally including non-market ones). It would also support the recommendation expressed in Recital 12 of Regulation (EU) 2016/792[190], the HICP Framework Regulation, to the effect that "research on a harmonised cost of living index should be initiated".

The effort involved in computing an experimental COLI would depend on its scope. If limited to a scope similar to that of the HICP, there would be synergies with the investment needed to obtain more timely weights for the HICP based on more frequent collections of consumption pattern data. An "HICP-scope" COLI would be the natural gauge for quantifying substitution bias on a continuous basis.

The calculation of a broader-scope COLI requires substantial investment in gathering data, deciding which non-market goods to include and developing methods for valuing non-market goods and services. The effort would be reduced if the existing quarterly COICOP-3-level information used by NSIs to construct the personal consumption deflator was made available; this information is already used in production at a quarterly frequency but only published annually.

HICP, national account, and COLI calculations all benefit from ongoing projects tapping into alternative data sources and new technologies. The ESS is making great strides in exploring new data sources and technologies such as web scraping and the use of scanner data. These may greatly reduce the effort required to measure consumption patterns going forward than would have been the case a few years ago, although the coverage of the consumption basket is uneven, especially with respect to services.

---

[190] Regulation (EU) 2016/792 of the European Parliament and of the Council of 11 May 2016 on harmonised indices of consumer prices and the house price index, and repealing Council Regulation (EC) No 2494/95 (OJ L 135, 24.5.2016, p. 11).

Box 13

An experimental euro area (conditional COLI) monthly weighted CPI[191]

In this box, we have constructed a proof of concept of a euro area COLI. This experimental index assumes the same scope as the HICP but allows consumption weights to change from month to month.[192] The index covers the period during which households were forced to change their consumption baskets as a result of the COVID-19 pandemic. Overall, this exercise is a proof of concept and shows that an experimental COLI for the euro area can be created using publicly available data. Nevertheless, the estimation of an official euro area COLI would require additional resources and information.

A growing body of literature has estimated experimental price indices across European countries. Using high-frequency data, several studies have identified large changes in spending across product categories.[193] These time-varying expenditure shares[194] have been used to quantify the difference between published consumer price indices and the inflation rate for items actually purchased by consumers.[195]

---

[191] This box is based on Kouvavas et al. (2020).

[192] In other words, the items included in the experimental index are the same as for the HICP. Therefore, the experimental euro area COLI does not include any non-market items.

[193] Contributions to this literature include Cavallo (2020) and Dunn et al. (2020) for the United States, Hacioglu et al. (2020) for the United Kingdom and Carvalho et al. (2020) for Spain. Consumption of food items has increased and remains relatively high because households are spending more time at home (effectively switching from food served in bars, restaurants and cafés). See Rubene (2020).

[194] These weights are approximations and do not necessarily match the accuracy of the weights derived from national accounts.

[195] See Jaravel and O'Connell (2020) for the United Kingdom and Huynh at al. (2020) for Canada. Following this approach, two statistical agencies have also published experimental price indices with monthly time-varying weights showing how the pandemic has affected consumer spending and that a gap has opened up between CPI-type inflation figures and the inflation rate of the items actually purchased by final consumers. See National Institute of Statistics and Economic Studies (2020) and Office for National Statistics (2020).

(percentages)
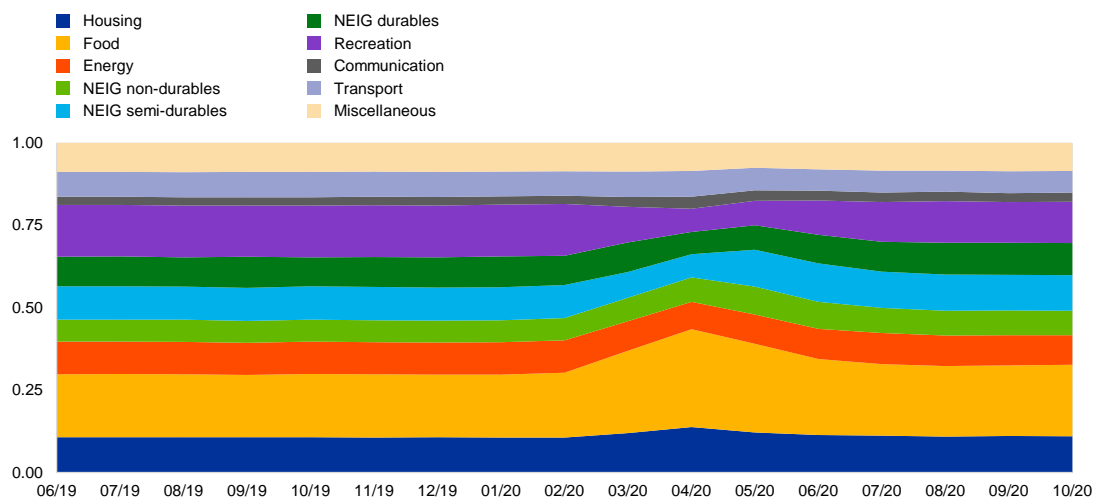
Source: Authors' calculations based on Eurostat data.
Notes: The chart shows the evolution of estimated relative spending. Spending patterns are calculated using the latest HICP weights as a starting point and applying growth rates based on turnover data for the retail trade and for other services. "NEIG" stands for "non-energy industrial goods". "Food" refers to food items and does not include eating out.

Several steps are necessary to construct a monthly-weighted CPI. For our experiment, we first estimated monthly household consumption weights. To do so, we identified proxies for each household spending category and matched them to the corresponding HICP (COICOP-5 level) category, using monthly turnover data. While retail trade turnover data primarily reflect transactions driven by household purchases, turnover of other services includes business-to-business transactions, which need to be stripped out. Then, taking the latest HICP weights as a base, we estimated the evolution of the spending categories using the corresponding nominal turnover growth rates.[196] Chart A shows the resulting weights data.[197] Finally, we used the estimated relative weights to construct the euro area COLI.[198] By design, our index captures part of the changes in consumption and therefore comes closer than a fixed-basket index to the rate of change in the prices of items actually purchased by consumers during this period.

Chart B shows the gap between the annual rates of change (year on year) of the experimental index and the HICP. Since the beginning of the pandemic, inflation as measured by our experimental index has been running higher than HICP inflation, although the difference has remained broadly stable in recent months. This gap started to open up in March and increased to

---

[196] The latest HICP weights reflect expenditure shares calculated using data mainly from 2018.

[197] As Chart A reports relative weights, some categories show an increase in March/April because the nominal spending in that category contracted less than overall consumption. Our estimate of the contraction in overall spending is in line with published Eurostat statistics for private final consumption.

[198] The match to HICP categories is therefore imperfect, particularly where monthly turnover data are used, as these data also include transactions between firms. The sources used to estimate monthly weights are less reliable than official HICP expenditure sources. Our (unchained) index assumes the same scope and coverage as the HICP using the same price data. In other words, the main difference between our index and the HICP is that we allow the spending weights to change from one month to the next, while the HICP keeps them constant. A Fisher index formula was used.

about 0.2 percentage points in April.[199] Intuitively, this reflects consumers switching from lower-than-average inflation categories to higher-than-average inflation categories.

Chart B

Difference between changes in a monthly-reweighted index and the HICP

(percentage points, year-on-year changes)
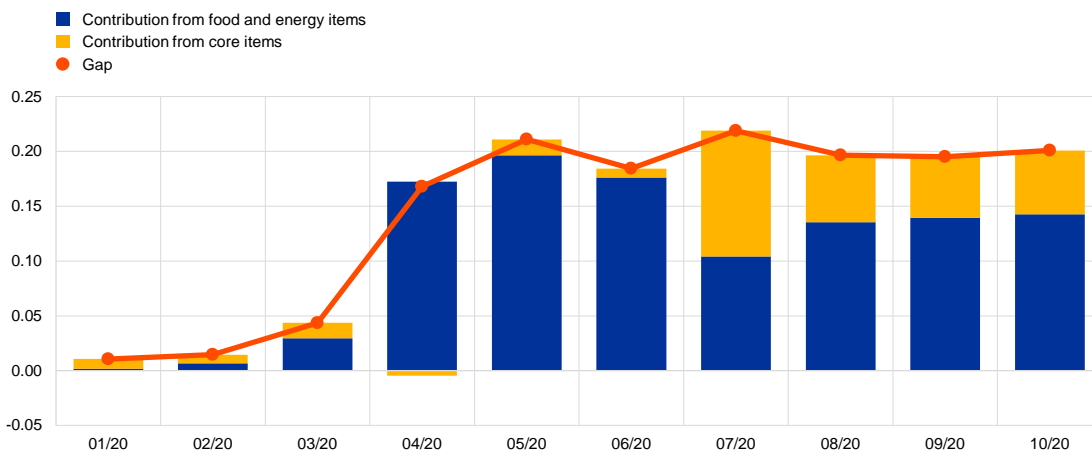
Source: Authors' calculations based on Eurostat data.
Notes: The orange line depicts the difference between year-on-year HICP inflation and the year-on-year change in our alternative index. The blue bars show the contributions of food and energy items and the yellow bars show the contributions of core inflation items.

Box 14

Approach and experience of the FRS with a COL-type index

In the United States, there are two separate monthly consumer price indices. One is closer to the definition of a COGI, while the other is closer to the definition of a COLI. Two statistical agencies (the Bureau of Labor Statistics (BLS) and the BEA) publish separate price indices on a monthly basis. The two main indices are the CPI and the PCE price index.[200] The latter is the estimated deflator of personal consumption in the gross domestic product (GDP) statistics. The BEA publishes monthly spending data that are used as weights in the PCE price index, whereas the weights in the US CPI are updated every other year. The CPI and the PCE price index differ in several respects, including their scope, the formula used to aggregate elementary items, and the weight of each item.[201]

---

[199] Because our analysis is at COICOP-2 level, the gap between our index and HICP inflation reflects only the upper level of changes in consumption patterns during lockdown (in particular, the gap reflects switching between food and energy items and core items at COICOP-2 level). The gap between our index and the HICP has remained fairly constant in recent months because the change in weights has resulted in a lasting shift in the level of the series.

[200] On top of the Consumer Price Index for All Urban Consumers (CPI-U), the BLS publishes four alternative consumer price indices: (i) the Consumer Price Index Retroactive Series (CPI-U-RS); (ii) the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W); (iii) the chained CPI (C-CPI-U), which aims to account for substitutions that consumers make across item categories in response to changes in relative prices; and (iv) the experimental Consumer Price Index for Older Americans (CPI-E).

[201] See Bureau of Labor Statistics (2011). For CPI, see Chapter 17 of the BLS CPI handbook of methods and for the PCE price index, see Chapter 5 of the BEA NIPA handbook of methods.

The PCE price index is a COL-type index and is the FRS's target.[202] In its Monetary Policy Report submitted to the US Congress on 17 February 2000, the FRS stated its preference for the PCE price index over the CPI because the former addresses substitution biases, eliminates self-reporting biases in the survey-based weights, and ensures that the information set available to the central bank at any given point in time is larger than if using the CPI.[203]

The CPI and the PCE price index tend to exhibit a gap that persists even over long horizons. Chart A shows the 12-month changes of the CPI and the PCE price index from January 2003 to the present (left panel shows headline inflation and right panel the evolution of the indices excluding food and energy items). While the overall developments are broadly similar, the CPI tends to run higher than the corresponding PCE price index, with the gap averaging around three-tenths over long time periods.[204] This applies to both headline inflation (the FRS's policy target) and the indices, excluding food and energy items. The difference in inflation between a COG-type and a COL-type index[205] is relatively small but can matter for policymakers, depending on the economic environment: when inflation is low and the Phillips curve is relatively flat the difference can entail large adjustments on other margins. Further, the CPI-PCE price index gap is one-third of a percentage point but occasionally pushes CPI inflation above the FRS target while keeping PCE price inflation below it, therefore possibly complicating the central bank's communication and its monetary policy response.[206] Nevertheless, even though a COL-type of index raises additional measurement challenges, the FRS choice has not been reported as having generated communication or credibility issues.

---

[202] The PCE price index is an empirical approximation of a theoretical cost of living indicator. While being the most advanced such indicator published by a statistical agency to date, the PCE price index does not include certain items (such as environmental goods).

[203] See The Federal Reserve Boards (2019).

[204] Most of the CPI-PCE gap is explained by the "formula effect", i.e. the fact that PCE accounts for households' substitution across items.

[205] While the COGI-COLI gap is well known and studied in the United States, not much can be inferred for the euro area because of the different structures of the two economies (especially in terms of share of non-market items that can potentially be included in a euro area COLI).

[206] In the last 20 years or so, average CPI inflation has been around 2%, while the mean of PCE inflation has been significantly below 2%.

Chart A

US CPI versus PCE price index

Headline inflation	Excluding food and energy items
(year-on-year growth rates)	(year-on-year growth rates)
■ CPI ■ PCE price index	■ CPI excluding food and energy ■ PCE excluding food and energy
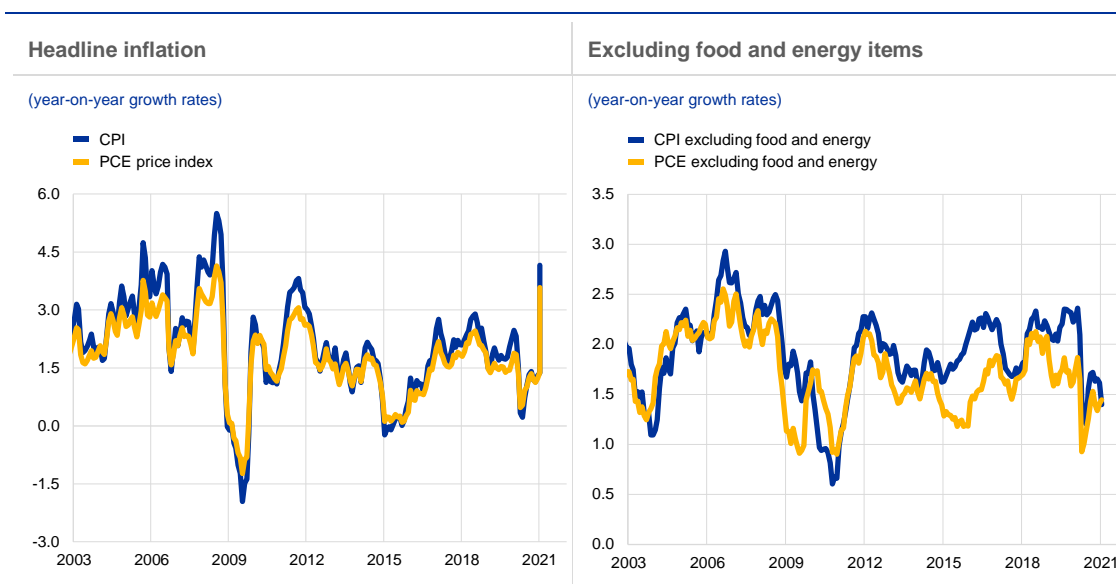
Source: ECB calculations based on Bureau of Labor Statistics and BEA data.
Notes: The charts show the 12-month percentage changes in the US CPI (blue line) and the US PCE index (yellow line). Left panel refers to headline inflation, right panel refers to the indices excluding food and energy. The latest observation is for April 2021 for all series.

The experience of the US Federal Reserve System shows that having one index of reference complemented by auxiliary indices can provide a rich signal to the central bank about the state of the economy. In the euro area, the economic analysis underlying monetary policy would greatly benefit from the development of an experimental index designed to approximate a theoretical COLI. While significant measurement challenges would need to be overcome, an experimental COLI would be a useful complement, especially at times of large cyclical movements.

Box 15

The cyclicality of upper and lower-level substitution bias

Households switch goods and varieties of goods in reaction to exogenous shocks.[207] The substitution bias (and whether it is cyclical) can be estimated by comparing the behaviour of a COL-type index with its COG-type counterpart constructed using microdata. Amann et al. (2020) have assessed the time variation of the substitution bias using household scanner data.[208] A superlative Törnqvist index with analogous inflation is compared with a Laspeyres index, created from the same dataset and mirroring the construction of the HICP in terms of its infra-annual properties.[209] Attention is restricted to the food, alcohol and tobacco categories, which together accounted for

---

[207] Recent research has shown that this substitution tends to be correlated with the business cycle. Coibion et al. (2015) and Kaplan and Menzio (2015) have shown that households tend to substitute cheaper products and outlets over time and that substitution is more intensive in downturns.

[208] The baseline dataset is the household-scanner dataset from GfK/Kantar for France and Germany, which has been acquired by the ECB through the price-setting microdata analysis (PRISMA) network. The dataset documents purchases of around 18,000 and 35,000 households, respectively. The household-level data cover all types of supermarket (including hard discounters). The US evidence is based on a store-level scanner dataset from IRi and covers sales of over 3,000 stores.

[209] The HICP is constructed as an annually chain-linked Laspeyres-type index. Within a calendar year, weights are kept fixed and are applied to the prices of the previous December. An HICP time series is obtained by annually chain-linking the Laspeyres-type indices of each calendar year over December.

17% and 21% of the German and French HICP baskets respectively in 2018. The overall bias is decomposed into upper-level (across category) and lower-level (within category) biases.[210]

Under quite general conditions, a superlative index approximates economically relevant COLIs. This relies on the assumption of homothetic household preferences. However, it fails if households switch to higher-quality, more expensive goods as their income rises.[211] If the assumptions are upheld, Törnqvist inflation rates are never higher than Laspeyres ones, contrary to what Amann et al. (2020) occasionally find, especially in economic upturns.[212] This points to an imperfect approximation implied by the use of superlative indices, probably due to a failure of the homotheticity assumption. In other words, contrary to assumptions household preferences may well vary with income. However, this is not necessarily a critical problem: at any given time, analysts have many indications of what is happening in the economy and are able to use their judgement to interpret these effects. For this reason, having auxiliary superlative indices can foster more robust economic analysis.

---

[210] The upper-level bias is measured as the difference between annual Laspeyres inflation and annual inflation according to a superlative Törnqvist index, across categories but not products (similar to the C-CPI-U). The lower-level bias is measured as the difference between annual inflation based on time-invariant weights across products within categories, but without chain-linking (as applied in both the HICP and the C-CPI-U).

[211] See, for example, Kouvavas (2019).

[212] Kaplan and Schulhofer-Wohl (2017) confirm that substitution towards cheaper goods is not uniform across households either; more than 40% of households fail to substitute in the expected direction and have a higher Törnqvist index than Laspeyres index.

Time variation of superlative Törnqvist and fixed-weight Laspeyres inflation

(percentage points, year-on-year changes)

■ Laspeyres inflation
■ Superlative Törnqvist inflation

Notes: The figure shows the evolution of the superlative Törnqvist (yellow line) and the Laspeyres (blue line) food annual inflation rates for France, Germany and the United States. The figures illustrate the time variation of the substitution bias (the difference between the two inflation rates). The shaded areas depict periods when food inflation rates were low, and in particular when they were below the first quartile of the historical food inflation distribution. The figures show that the bias tends to be positive when the inflation rate is low. Symmetrically, the bias tends to be negative when the inflation rate is high.

Table A

Estimated nominal household spending

Standard deviation of the bias	Overall bias	Lower-level bias	Upper-level bias
France	0.61 p.p.	0.53 p.p.	0.14 p.p.
Germany	0.94 p.p.	0.68 p.p.	0.38 p.p.
United States	0.90 p.p.	0.76 p.p.	0.26 p.p.

Notes: The table shows the standard deviation of the substitution bias (difference between Laspeyres and superlative Törnqvist annual inflation rates) in food prices over time in France (2009-18), Germany (2006-18) and the United States (2002-12). It also shows the standard deviation of the lower-level bias (difference between the Törnqvist index and an index that uses uniform weights across products within categories) and the upper-level bias (the difference between the Laspeyres index and an index that uses superlative Törnqvist weights across categories but not products).

Time-varying substitution bias is a robust feature observed in the micro data. The magnitude of the time variation can be sizeable and appears to be correlated with the business cycle. Chart A shows the evolution of the Törnqvist and Laspeyres year-on-year inflation rates in France (2009-18), Germany (2006-18) and the United States (2002-12). The average bias inferred from the behaviour of supermarket prices is zero in France, tiny in the United States (5 basis points), and sizeable in Germany (-67 basis points), but its time variation is apparent in all countries (see Table A). Consistently with economic theory, substitution is time varying. Moreover, it can become substantially larger during downturns when inflation is in any case low.[213] For instance, when inflation was at its lowest, the bias was 1.21 percentage points in France (in 2015), 1.85 percentage points in Germany (in 2009) and 1.28 percentage points in the United States (in 2002). Its decomposition indicates that the lower-level bias accounts for more than two-thirds of the time variation of the bias. The findings of Amann et al. (2020) support the theoretical intuition that substitution is more prevalent between goods within the same product category, which are close substitutes, than across categories.

The substitution bias can be relevant for policymakers, especially during downturns. The gap can also have implications for monetary policy in specific circumstances, for example when the Phillips curve is relatively flat and price inflation is below target. This dimension is especially relevant in policymaking: in practice the information provided by the gap between the HICP and a COLI serves not only as a useful, albeit imprecise, estimate of the substitution bias, but also as a form of early signal for policymakers to identify turning points in real time.

---

[213] When Laspeyres inflation was below its first quartile the bias was 21 basis points in Germany (compared with -67 basis points on average), 44 basis points in France (0 on average) and 48 basis points in the United States (3 basis points on average). Research (Handbury et al., 2013) based on Japanese data suggests that the measurement bias is time-varying and high at low levels of inflation for Japan as well.

6 The role of underlying inflation

The current inflation objective is formulated in terms of headline HICP over the medium term. Headline inflation is affected by short-term volatility induced by temporary idiosyncratic shocks, so it can be noisy, blurring the medium-term inflationary pressure signal that is relevant for monetary policy. To help distinguish signal from noise in the data, central banks monitor measures of underlying inflation. Several concepts of underlying inflation have been developed to abstract from short-term volatility in headline inflation.[214]

In this context, it could be argued that an inflation objective should be set in terms of underlying inflation. However, an underlying inflation objective may pose important communication and transparency challenges. First, there are many ways of measuring underlying inflation. Second, some underlying inflation measures require the estimation of econometric models, which the general public would find hard to understand. Third, many underlying inflation measures are produced by central banks rather than by an independent statistical agency. Finally, while food and energy prices are very important from the consumers' perspective, many underlying inflation measures either exclude these items or assign a lower weight to them. This could affect the credibility of the price index. Consequently, it could be problematic to define the Eurosystem's inflation objective in terms of underlying inflation.

This chapter concludes that the best course of action would be to monitor a broad range of measures of underlying inflation as there is no single measure that emerges as optimal across different criteria. These underlying inflation measures are important indicators for monetary policy. They can give useful signals as to how HICP inflation is likely to evolve over medium-term horizons. These measures should be discussed on a systematic and consistent basis in the Eurosystem's public communication.

The chapter is structured as follows. The first section presents four groups of underlying inflation measures. The second section assesses the properties of these measures. The third and final section concludes and provides recommendations.
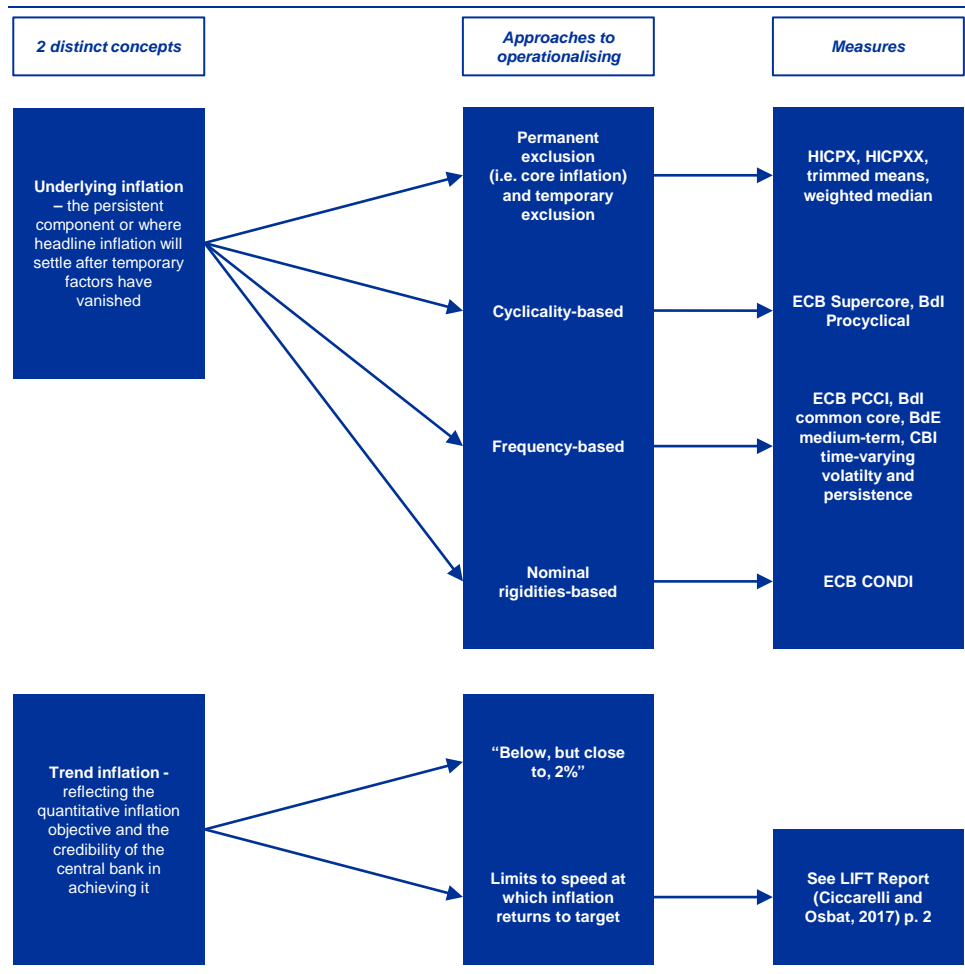
6.1 Measures of underlying inflation

As there are many ways of measuring underlying inflation, it is important to understand the construction and properties of these different measures. For ease of presentation, underlying inflation measures can be classified as simple measures, which do not require the estimation of an econometric model, and model-based ones (Figure 4). The former group can be further distinguished as measures that exclude – on either a permanent or temporary basis – some items and those

---

[214]   See, for example, Ehrmann et al. (2018).

that re-weigh items based on the frequency of price adjustment or the statistical properties of price changes. Model-based measures can be distinguished depending on whether they consider the response to the business cycle or try to approximate the slow-moving component of inflation. In addition to these measures the GDP deflator is also sometimes proposed as an alternative measure of underlying inflation. Box 16 provides a comparison of the GDP deflator with the standard underlying inflation measures.

Figure 4

Underlying inflation – one concept, several ways to operationalise

| *2 distinct concepts*	*Approaches to operationalising*	*Measures*
Underlying inflation – the persistent component or where headline inflation will settle after temporary factors have vanished	**Permanent exclusion (i.e. core inflation) and temporary exclusion**	**HICPX, HICPXX, trimmed means, weighted median**
	Cyclicality-based	**ECB Supercore, BdI Procyclical**
	Frequency-based	**ECB PCCI, BdI common core, BdE medium-term, CBI time-varying volatilty and persistence**
	Nominal rigidities-based	**ECB CONDI**
Trend inflation - reflecting the quantitative inflation objective and the credibility of the central bank in achieving it	**"Below, but close to, 2%"**	
	Limits to speed at which inflation returns to target	**See LIFT Report (Ciccarelli and Osbat, 2017) p. 2**

6.1.1 Exclusion-based measures

The simplest measures of underlying inflation are exclusion-based, that is, they exclude selected volatile items from the HICP basket. The underlying assumption of these approaches is that developments in certain items mostly reflect noise. These measures can be further classified into permanent exclusion measures, where excluded items are always the same, and temporary exclusion measures, where excluded items may change from month to month (Chart 10). A commonly

used permanent exclusion measure is HICPX, which covers around 70% of the HICP.[215] By excluding the volatile components of energy and food, this measure abstracts from changes in relative prices unrelated to underlying inflation. In recent years, large idiosyncratic movements in HICPX inflation have led to the construction of even narrower measures. For instance, the ECB publishes an HICP excluding energy, food, travel-related items and clothing (HICPXX), which covers about 60% of the HICP. An additional measure, called fine core, has been proposed by the Banque de France.[216] This permanent exclusion-based measure omits items with the highest volatility at the most disaggregated level available. It represents 70% of the HICP – the same weight as the HICPX – and behaves more or less similarly to trimmed-mean inflation for the group of temporary exclusion-based measures. The advantage of these measures is that they are intuitive, transparent, and can be broken down by subcomponents or countries. A disadvantage is that by excluding items (including those that tend to be frequently purchased), they may not be perceived as fully representative.

Chart 10

Exclusion-based measures

(annual percentage changes)
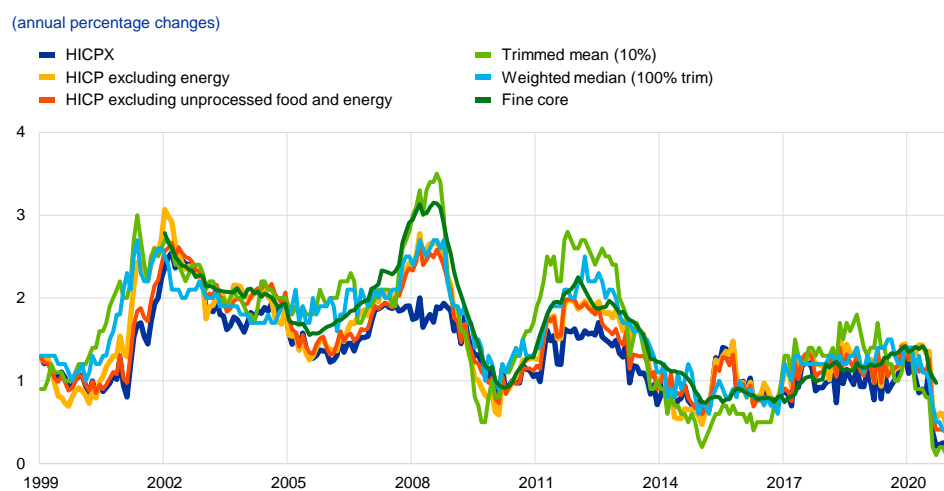
Sources: ECB and Banque de France.
Note: The latest observations are for December 2020 for fine core and April 2021 for the rest.

By contrast, temporary exclusion-based measures have a changing composition, as they exclude different items each month. For instance, trimmed-mean measures exclude items with the highest and lowest month-on-month change, representing a given weight of the HICP at the tails of the weighted distribution of month-on-month changes. The weighted median keeps a single item each month located in the middle of the weighted distribution. These measures are appealing since they do not assume a priori which items are volatile and remove large one-off price changes in typically non-volatile items. However, they do not allow for a decomposition by subcomponents or countries.

---

[215] Other similar aggregates published by Eurostat are the HICP excluding energy and the HICP excluding energy and unprocessed food.

[216] See Lalliard and Kalantzis (2020).

6.1.2 Costs of nominal distortions measure

The costs of nominal distortions index (CONDI) has a different theoretical underpinning from the measures discussed previously, since it is based on welfare considerations. This index (Chart 11) assigns a larger weight to "stickier" HICP components (i.e. those with a lower frequency of price change) relative to their expenditure weights. A CONDI has recently been constructed for the euro area.[217] The CONDI approach is based on the premise that a stable price index should allow households and firms to make better consumption and investment decisions, allocating their spending efficiently across goods and sectors. It emphasises the benefits of price stability precisely in terms of minimising inefficient price dispersion, a distortion in the pricing system that acts much like a tax on the entire market allocation process. This would be a theoretical argument in favour of using a CONDI for defining the target inflation rate. However, some economists argue that CONDIs should, in principle, also perform better in forecasting future inflation.[218] The rationale is that firms with stickier prices have an incentive to anticipate accurately and, indeed, build in inflation expected over the lifetime of the prices they set.

Chart 11
Costs of nominal distortions measure

(annual percentage changes)

Source: ECB.
Note: The latest observation is for November 2020.

6.1.3 Cyclical measures

Some underlying inflation measures go beyond a purely statistical basis by making an explicit link to macroeconomic conditions.[219] Specifically, the ECB's Supercore measure and the Banca d'Italia's procyclical measure exploit the Phillips curve relationship to assess the responsiveness of inflation at the disaggregate level to developments in the real economy, as measured by the output gap or

---

[217] Data refer to the COICOP-3 level. See Nakov (2020).

[218] See, for example, Bryan and Meyer (2010).

[219] See Ehrmann et al. (2018) and Conflitti (2020).

unemployment gap (Chart 4). These measures distinguish those items of the price index that respond to the economic cycle (procyclical) from those that do not (acyclical). The latter are driven by specific developments that are independent of the state of the economy as a whole and are not therefore considered in the cyclical measures of underlying inflation.

6.1.4   Frequency-based measures

Alternatively, some underlying inflation measures filter out transitory components and only retain persistent[220] **components of inflation rates to abstract from idiosyncratic shocks affecting consumer prices (Chart 12).** The ECB's persistent and common component of inflation (PCCI) belongs to this group and is intended to capture medium to long-run inflation dynamics. This measure reflects the common component that accounts for changes in prices due to shared shocks while the idiosyncratic component, i.e. price movements due to non-pervasive shocks, is removed.[221] The Banca d'Italia's common core measure uses a similar methodology to the one used for the PCCI, but the model is applied on aggregate euro area HICP items.

Other proposals have also been made as to how to capture longer-term inflation developments.[222] For instance, the Banco de España's medium-term inflation measure isolates inflation cycles lasting more than eight years using a low-pass Butterworth filter. This measure is smooth and tracks the medium-term mean of HICP inflation well.[223] A somewhat different measure is the time-varying volatility and persistence-based measure (TVVP) proposed by the Central Bank of Ireland. All HICP items are re-weighted, simultaneously taking into account each item's persistence[224] and volatility. Overall, this measure is similar to trimmed mean measures, but appears somewhat smoother. While it is relatively easy to construct, it may be difficult to explain to the general public.

---

[220] Note that HICP fluctuations can be broken down in terms of all different frequencies. Trend inflation measures only refer to the very slow-moving part.

[221] Common and persistent components of inflation rates across over 1,000 HICP items from 12 countries are estimated using a generalised dynamic factor model and capture cycles longer than three years. The final measure is a three-month moving average of the weighted average of these low-frequency components. Medium-term shocks affecting food and energy items are captured by this measure, while short-term fluctuations from items such as services are removed. The persistent and common component of inflation excluding food and energy components is constructed to exclude food and energy components. See Bańbura and Bobeica (2020) and Conflitti (2020).
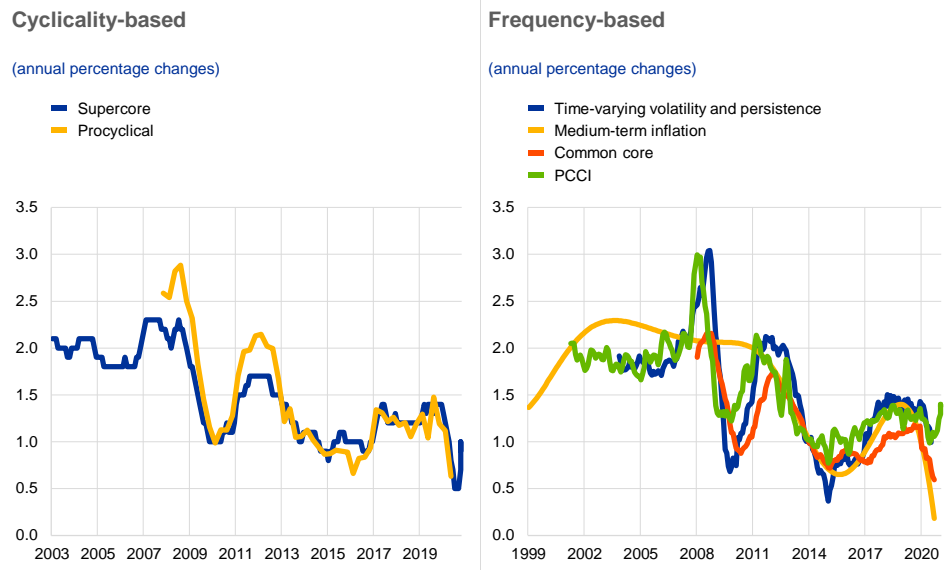
[222] See Zekaite (2020a). For a summary of the Banco de España measure, see the background slides by Alvarez and Cuenca (2020). Note that this measure is revised as new data are included.

[223] The number of years considered in this measure depends on the researcher's preferences. Note that the measure may be subject to the end-of-sample problem. This impact can be minimised by extending the inflation series with accurate forecasts. Indeed, if perfectly accurately forecasts were available the end-of-sample problem would disappear.

[224] Persistence is measured as the sum of autoregressive coefficients in an AR model. Volatility is measured as the standard deviation of the difference between the given item's inflation rate and HICP inflation. Both characteristics are calculated over rolling windows of 24 months.

Chart 12

Cyclicality- and frequency-based measures

Cyclicality-based

(annual percentage changes)

- Supercore
- Procyclical

Frequency-based

(annual percentage changes)

- Time-varying volatility and persistence
- Medium-term inflation
- Common core
- PCCI

Sources: ECB, Banca d'Italia, Banco de España and Central Bank of Ireland.
Notes: The latest observations are for April 2021 for both Supercore and PCCI, the third quarter of 2020 for procyclical, July 2020 for time-varying volatility and persistence and September 2020 for both medium-term inflation and common core.

6.2   An evaluation of underlying inflation measures

An empirical assessment can help discriminate between measures of underlying inflation. The aim of these measures should be to track the evolution of medium-term headline inflation. An assessment can be conducted based on empirical criteria, including unbiasedness and overall precision. For this purpose, an estimate of the persistent component of inflation is used as a benchmark.[225] As this component is unobservable and surrounded by high uncertainty, we use the 24-month centred moving average of monthly headline HICP inflation as a proxy.[226]

The accuracy of the measures of underlying inflation is episodic. The relative abilities of the measures to track the future development of the benchmarks vary considerably over different subsamples (see Tables 2a and 2b). Over the full sample, HICPXX has the lowest RMSE, while TVVP and medium-term inflation have the lowest bias.[227] Over the pre-global financial crisis sample, the PCCI and the GDP deflator, which is also considered in this comparison, tend to perform best in

---

[225]   Other criteria are conceptually possible, such as forecasting performance. This is left for future research.

[226]   The 24-month centred moving average should be sufficiently long to smooth out high-frequency fluctuations, yet short enough to reflect the horizon at which monetary policy operates over the business cycle. Given that constructing this measure requires the use of future values of inflation, it has limited conjunctural use, but it can serve as a benchmark to assess other indicators. The latest available series of the medium-term inflation measure developed at the Banco de España is also used as a benchmark – the results are broadly similar and are not shown here.

[227]   The bias and the RMSE are defined in terms of approaching the benchmark series.

tracking the benchmark two-year moving average of inflation.[228] While some measures, such as the weighted median, have only a small bias, their overall precision is relatively poor. Over the post-global financial crisis sample, the HICPXX has a somewhat lower RMSE than the other measures of underlying inflation, with the performance of the measures generally deteriorating compared with the earlier sample.

Table 2a

Accuracy of measures of underlying inflation

(units)

	HICP	HICP excluding energy	HICP excluding energy and unprocessed food	HICPX	HICP services	HICPXX	10% trimmed mean	30% trimmed mean	Weighted median
Full sample									
RMSE	1.14	0.90	0.86	0.79	0.85	0.73	1.01	0.93	0.86
Bias	0.08	-0.05	-0.07	-0.21	0.33	-0.18	0.11	0.06	0.04
Pre-GFC									
RMSE	0.98	0.81	0.78	0.70	0.84	0.65	0.89	0.77	0.72
Bias	0.28	-0.03	-0.05	-0.24	0.50	-0.19	0.18	0.06	0.00
Post-GFC									
RMSE	1.25	0.95	0.92	0.85	0.86	0.78	1.08	1.02	0.95
Bias	-0.07	-0.06	-0.09	-0.18	0.21	-0.18	0.06	0.05	0.07

Sources: Eurostat, Central Bank of Ireland, Banque Centrale de France, Banca d'Italia, Banco de España and ECB calculations.
Notes: Bias defined as the average difference between the underlying inflation indicator and the benchmark series (24-month centred moving average of monthly headline HICP inflation). GFC stands for: global financial crisis.

---

[228] While the GDP deflator differs in some important respects from measures of underlying inflation, it provides useful signals of emerging price pressures in the economy and is therefore an important complementary indicator of medium-term consumer price pressures.

Table 2b

Accuracy of measures of underlying inflation

(units)

	GDP deflator	PCCI	Super-core	CONDI	TVVP (CBI)	Medium-term inflation (BdE)	Common core (BdI)	Fine core (BdF)	Pro-cyclical (BdI)
Full sample									
RMSE	0.85	0.81	n.a.	0.85	0.89	1.36	0.91	0.95	0.89
Bias	-0.08	0.23	n.a.	-0.30	0.01	0.01	-0.31	0.09	0.07
Pre-GFC									
RMSE	0.63	0.65	n.a.	0.75	0.77	0.82	0.94	0.89	0.87
Bias	0.09	0.23	n.a.	-0.34	0.03	0.12	-0.56	0.14	0.16
Post-GFC									
RMSE	0.98	0.90	1.02	0.92	0.97	1.63	0.88	1.00	0.91
Bias	-0.20	0.23	0.06	-0.27	-0.01	-0.07	-0.14	0.05	0.01

Sources: Eurostat, Central Bank of Ireland, Banque Centrale de France, Banca d'Italia, Banco de España and ECB calculations.
Notes: The RMSE is computed by evaluating the error incurred by each of the different measures at each specific time t to capture the inflation trend at time t. To the extent possible, the measures of underlying inflation are computed in "real time", i.e. by considering only the information that would be available to the forecaster at time t. GFC stands for global financial crisis. Pre-GFC: January 2002 to August 2008; post-GFC: September 2008 to December 2019; full sample: January 2002 to December 2019. Supercore results not available for pre-GFC period.

Measures of underlying inflation should also demonstrate a link to business cycle conditions. In a reduced-form Phillips curve regression based on the output gap, the short-run slope is highly significant in all regressions except in the case of the HICP, PCCI and medium-term inflation measures. The estimated long-run slope lies in the range 0.2 to 0.5, with the slope at the upper-end of the range for the fine core, trimmed means and cyclicality-based measures (see Chart 13).[229] Overall, the measures of underlying inflation generally show a strong link to the business cycle.[230]

Measures of underlying inflation should satisfy some practical criteria and be sufficiently transparent. First, they should be available on a timely basis, since HICP data are quickly available. Many of the measures can only be constructed when the detailed monthly data have been released. Most of the measures are not revised but there are notable exceptions, including the PCCI and cyclical measures. It is also important for measures of underlying inflation to be sufficiently transparent so that they can be easily communicated to the public. In this regard, developments in permanent exclusion measures are advantageous, as any divergence from headline inflation can be attributed to certain sub-components. Developments in temporary exclusion measures can be more challenging. Frequency-based measures, particularly those derived from a model, can pose greater communication challenges. In addition, the results can sometimes be challenging to interpret.

---

[229] The long-run slope is calculated as the slope of slack divided by one minus the autoregressive coefficient in a Phillips curve regression.

[230] Recently, Fröhling et al. (2021) proposed an alternative measure of underlying inflation which uses import shares of consumption components as an exclusion criterion. The authors show that their measure of domestic inflation performs similarly to core rates in terms of accuracy and displays a stronger link to the business cycle at least in the time after the great financial crisis.

Chart 13

Long-run slope in Phillips curve regression

(slope)
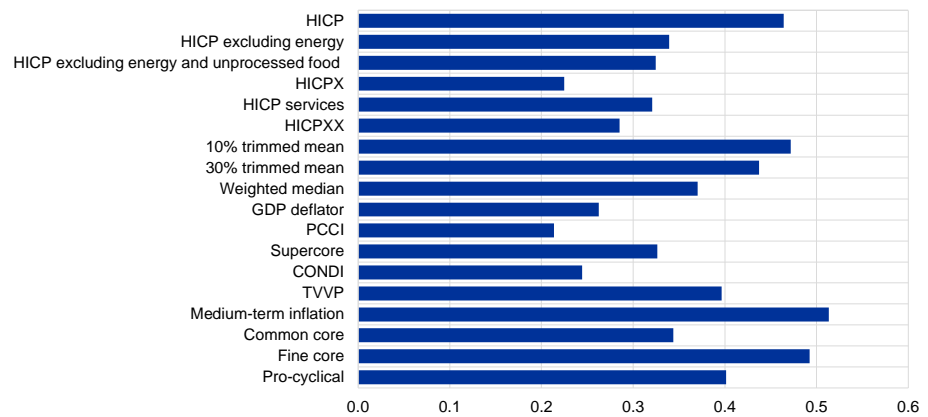
Sources: Eurostat and ECB calculations.
Notes: Based on output gap (results for unemployment gap also available). All corresponding short-run slopes are significant at the 1% level except for HICP and PCCI. The sample period is the second quarter of 2003 to the fourth quarter of 2019, using latest available vintages.

6.3   Conclusions and recommendations

Several underlying inflation measures have been put forward and assessed.
The aim of underlying inflation measures is to help distinguish signal from noise in inflation developments, thus abstracting from short-term volatility. These measures are useful in order to understand and explain HICP developments and should be given proper attention in the communication of monetary policy analysis and decisions. However, they are not well suited for use as an inflation objective, since they may pose significant communication and transparency challenges. The existence of many alternative measures of underlying inflation make it problematic for its use as the Eurosystem's inflation objective. Our assessment leads us to conclude that a broad range of measures of underlying inflation should be monitored, as individually the measures do not consistently give very precise or reliable signals. These underlying inflation measures can provide useful signals as to how inflation is likely to evolve over the medium-term horizon, which make them important indicators for monetary policy.

The Eurosystem's public communication should regularly and consistently refer to measures of underlying inflation.[231] As a tool, underlying inflation measures should feature in policy discussions when they provide additional relevant information. However, references to these measures in public communication would need to be regular and consistent because in recent years underlying inflation has been used in the context of assessing the sustained adjustment path of inflation and state-dependent forward guidance. This implies an important role for underlying

---

[231] To date, the ECB has referred fairly regularly to measures of underlying inflation in official publications such as the Economic Bulletin and the Accounts. References in the Introductory Statements to its Press Conferences have been less systematic.

inflation as an operational device in achieving the price stability objective. Consistent use is also warranted because the Eurosystem/ECB now regularly publishes projections for some exclusion-based measures of underlying inflation to highlight the relevance of conditioning the projections for headline inflation on food and energy prices.

Box 16
The GDP deflator and measures of underlying inflation

In the 2003 strategy review, the GDP deflator was discussed as an alternative concept to consumer price inflation in formulating the price stability objective. This may be appropriate from a perspective where monetary policy has the task of financing monetary transactions, as in nominal GDP targeting, which is a much broader perspective than a focus on consumption. In this box, by contrast, it is taken as given that the relevant concept is consumer price inflation and the information content of the GDP deflator is examined in that context.

The GDP deflator is an important indicator in the analysis of price developments in the euro area. Developments in the GDP deflator are closely monitored and analysed on a regular basis together with those of a wide range of other indicators in order to obtain a comprehensive assessment of the drivers and developments of inflation. Beyond its role in the overall price analysis, the GDP deflator is sometimes also proposed as an alternative measure of underlying inflation. It is, however, currently not part of the standard range of measures used to gauge underlying inflation developments in the euro area. As described in this section, the standard measures of underlying inflation comprise model and non-model-based measures. This box compares the GDP deflator with the standard measures of underlying inflation and considers its merits as a complementary indicator of medium-term consumer price pressures.

The GDP deflator differs conceptually in some important respects from measures of underlying inflation. The GDP deflator and, for example, the HICPX differ in concept in the following ways. First, the GDP deflator captures value added and indirect taxes but not intermediate consumption, whereas HICPX measures the entire selling price.[232] Second, the composition of products is quite different. The GDP deflator is a broad measure covering all domestically produced goods and services (i.e. including export prices, while import prices are netted out, and also including prices of goods and services other than for consumption purposes as well as non-market items). The HICPX, on the other hand, refers to core consumer goods and services that originate in the domestic economy and also from abroad.

While there is a notable degree of co-movement between the GDP deflator and the HICPX over the longer term, larger deviations are not uncommon over the short to medium term (Chart A). Over the longer term, there is a clear alignment between the two indicators and it does not make much difference whether price developments of value added for the whole economy are considered, or underlying consumer price developments. However, persistent and large deviations can occur, as was notably the case from 2005 to 2007 and from mid-2017 to mid-2020.[233] Such episodes can

---

[232] In this respect, the HICPX is representative of the other standard measures of underlying inflation as all of them are constructed using HICP items only.

[233] Over recent quarters, the gap between the GDP deflator and HICPX inflation has widened substantially; annual growth in the GDP deflator was 2.4% in the second quarter of 2020, up from 1.9% in the previous quarter, whereas HICPX inflation was 0.9% in the second quarter of 2020 and 1.1% in the previous quarter.

partly be explained by developments in the terms of trade, which in turn mainly reflect the impact of oil price and euro exchange rate movements. Indeed, the monitoring of terms of trade effects is part of the regular assessment of consumer inflationary pressures.[234]

Chart A

GDP deflator and HICPX
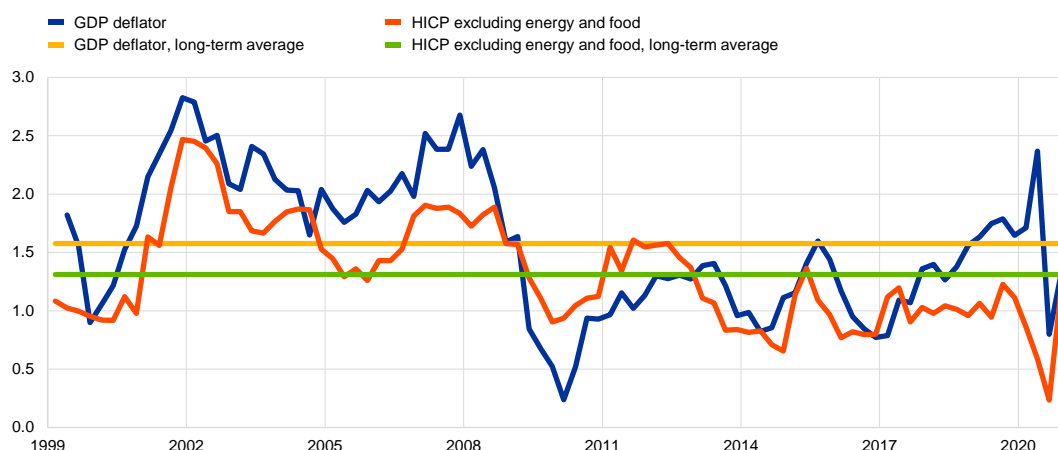
(annual percentage changes)

Sources: Eurostat and ECB calculations.
Note: The latest observation is for the first quarter of 2021.

Divergences between the GDP deflator and the HICPX may also reflect prices movements in the GDP deflator that only have a tenuous link to consumer prices. Aside from consumer prices, the GDP deflator also includes prices for government consumption and investment and net exports.[235] The government consumption deflator can at times play an important role in driving the divergence – this was evident in recent quarters, when productivity in the public sector declined sharply (Chart B).[236] Divergences can also sometimes be attributed to abnormally large movements in the investment deflator that tend to be driven by booms and busts in the construction sector.[237] Government consumption and investment prices can provide an additional view of where price pressures are emerging in the economy. However, it is unclear whether and to what extent these pressures would feed through to consumer price pressures over the medium-term through the various cost and demand channels.

Empirical tests suggest that the signalling power of the GDP deflator with respect to medium-term inflationary pressures in consumer prices is comparable to that of underlying inflation indicators. Ultimately, it is an empirical question whether the GDP deflator can provide useful leading information on medium-term consumer inflationary pressures beyond that provided by standard

---

[234] For more on the role of the terms of trade in explaining divergences, see, for example, European Central Bank (2016b).

[235] Diev et al. (2019) show that core inflation can be broken down algebraically into four main factors: (i) unit labour costs; (ii) margins; (iii) terms of trade excluding energy and food; and (iv) price differentials between personal consumption and government consumption, investment and exports.
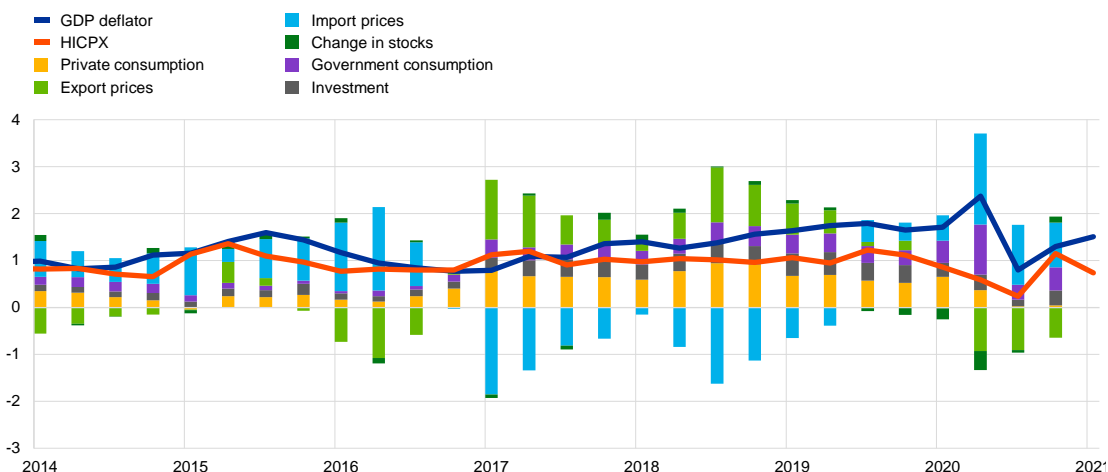
[236] The government consumption deflator increased strongly over recent quarters (at 4.4% year-on-year in the second quarter of 2020, up from 2.8% in the previous quarter and 1.7% in the fourth quarter of 2019). This mainly reflects a decrease in labour productivity in the public sector. It is worth recalling that there are recognised difficulties in measuring prices of government services such as health and education (see also European Central Bank, 2003, op. cit.).

[237] The gap between the GDP deflator and HICPX over recent years is also partly explained by the investment deflator, which is making a higher contribution than in 2016/2017.

HICP-based measures of underlying inflation. One approach is to assess in pseudo real-time the performance of these indicators in tracking the evolution of a measure of trend HICP inflation.[238] According to bias and RMSE metrics, the GDP deflator does not consistently outperform the standard measures of underlying inflation but shows a comparable performance to them.[239]

Chart B
GDP deflator and HICPX

(annual percentage changes)

Legend:
- GDP deflator
- HICPX
- Private consumption
- Export prices
- Import prices
- Change in stocks
- Government consumption
- Investment

Sources: Eurostat and ECB calculations.
Note: The latest observations are for the first quarter of 2021 for the HICPX and the GDP deflator and the fourth quarter of 2020 for the rest.

The GDP deflator has some disadvantages from a practical perspective compared with the standard measures of underlying inflation. The GDP deflator is available only at a quarterly frequency and has a release delay of more than two months, whereas all of the standard measures of underlying inflation are available monthly and with a release delay of no more than two to three weeks.[240] Furthermore, the GDP deflator can be subject to revisions, while this is not typically the case for most of the standard measures of underlying inflation.[241] Another potential downside of the GDP deflator is that it could be more challenging to explain to the public than permanent exclusion measures of underlying inflation such as the HICPX.

In conclusion, the GDP deflator differs in some important respects from measures of underlying inflation. It provides valuable additional information for price analysis and useful signals of emerging price pressures in the economy. It is therefore an important complementary indicator of medium-term consumer price pressures.

---

[238] The trend inflation target in month t is defined as the annualised HICP growth rate over the subsequent two years, i.e. it is equal to 1,200*(pt+H – pt)/H, where H is 24 months.

[239] For the GDP deflator (taking the quarterly value for each month of the quarter), the RMSE is 0.85 and bias -0.05 for the full sample (January 2000-April 2018), while the RMSE is 1.04 and bias 0.04 for the shorter sample of July 2007 to April 2018. This compares with an RMSE of 0.84 and bias of -0.25 for HICPX over the full sample and RMSE 0.87 and bias 0.01 for HICPX for the post-GFC sample.

[240] The HICPX is already available around the end of the reference month on a flash estimate basis.

[241] Some measures of underlying inflation can be subject to revision – the HICPX flash is sometimes revised with the full release. Model-based measures, including the PCCI and Supercore, are also subject to revision (as a result for example of revisions in seasonal factors, or in the case of PCCI re-estimation of the factors, and in the case of Supercore revisions in the output gap).

7   Summary and conclusion

By considering experiences with the HICP and its improvements since 2003, this report finds that the HICP continues to fulfil the prerequisites for the index underlying the ECB's definition of price stability. It confirms that the HICP still performs well in terms of credibility, reliability, comparability, and timeliness, coming out favourably across all these dimensions. While this high-level finding is essentially the one communicated in 2003, in this review it is based on another 17 years of practical experience with the HICP and its use in policy analysis and communication. Over these years, the HICP framework has seen substantial further improvements but it is natural that both old and new issues in inflation measurement come to the fore as practical experience increases and economic circumstances and structures change.

Since the 2003 review of the ECB's monetary policy strategy, the quality of the HICP as an inflation measure for the euro area has been continuously improved by Eurostat and the statistical offices of EU Member States. First, HICPs have become more representative of actual changes in prices. One reason is that prices, which can fluctuate widely within a reporting month, as is the case with fuel and package holiday prices, for example, are now collected on several days within a reporting month (and not only on a single day). Another reason is that seasonal items like summer and winter clothing, seasonal food and holidays are now more adequately represented in HICP baskets. Significant progress has also been made in the area of quality adjustments. **Second, the comparability of the HICP across time and across countries has been improved** by, for example, harmonising the treatment of out-of-season items, developing common standards for sampling and product replacements, and making the annual updates of HICP weights at a more detailed level and in alignment with national accounts data. **Third, the HICP has become available in a timelier manner across euro area countries.** The provision of flash estimates has become compulsory for euro area Member States. The estimates are published for all Member States at the end of each reporting month for the HICP itself and now, for most countries, also for its main aggregates. **Fourth, data availability has been strongly improved** – mainly by introducing a more detailed level to the HICP's ECOICOP classification.

Despite the quality improvements in the HICP there is no clear evidence that the HICP measurement bias has substantially declined since the last strategy review. On the one hand, compared with 2003, annual weight-updating by all Member States and the introduction of explicit weighting at a more granular level of detail in the classification should have reduced the overall bias of the HICP. On the other, structural market changes such as the introduction of new and distinctly different outlet types (prevalence of e-commerce) and the increase in and short life cycle of existing product varieties, along with new and innovative products (certain electronic services, for example), create new challenges. Heterogeneous quality adjustment practices might still lead to biases impacting on differences across countries and even a negative bias in some cases. Overall, a knowledge gap

concerning the size of the overall measurement bias of the HICP still exists. This calls for further research by the ESCB and the research community at large, in close collaboration with the ESS.

Notwithstanding past improvements, the HICP needs to be further enhanced – especially by including OOH. As also supported by results from the ECB Listens events,[242] the inclusion of OOH is a key element for the enhancement of the HICP. This is desirable for reasons of both representativeness (as spending on housing represents a large part of household expenditure) and cross-country comparability (as the importance of rented accommodation and OOH varies markedly across euro area countries). The report evaluates two main options for the inclusion of OOH in the HICP: (i) the NA approach, which measures OOH based on observed transaction prices for the purchase of new dwellings, similarly to how other durable goods are treated in the HICP; and (ii) the RE approach, which imputes OOH costs on the basis of rents of other (comparable) dwellings.

There appears to be a valid case for including OOH in the HICP using the NA approach. First, the NA approach is conceptually consistent with the HICP framework, given its reliance on transactions rather than imputations. Second, it tracks conditions in the housing purchasing market more closely than other approaches. Eurostat already publishes quarterly OOHPIs based on the NA approach for euro area Member States; these provide a good basis for further work towards an "HICP-H" (HICP augmented with Eurostat's OOHPI). At the same time, OOHPIs do not match the quality of HICP statistics in terms of frequency (as they are quarterly) and of timeliness (publication lag of one quarter). This is problematic from a monetary policy perspective and may not be compatible with replacement of the HICP by the HICP-H as a measure to define price stability. Given the lack of appropriate data sources, it will most likely not be possible for the ESS to fully integrate OOH into the HICP, while meeting the required criteria, in the foreseeable future. Moreover, if the HICP-H included an asset component measured OOHPIs might move closely with (asset) prices in the housing market; the weight of OOHPI in HICP-H could tend to move pro-cyclically with the construction cycle. This could pose challenges for the conduct and communication of monetary policy.

In line with the allocation of responsibilities in the area of European statistics, implementing an official HICP-H index is in the remit of Eurostat and the NSIs, based on a European Parliament and European Council legal act (enhancement of HICP Framework Regulation ((EU) 2016/792 of 11 May 2016). Implementation will inevitably be step-wise. Preparing and implementing an official HICP-H would take several years in view of the statistical work required and the legislative process. According to a tentative roadmap designed by the work stream, in a first step "analytical" indicators for HICP-H constructed within the ESCB could be further improved and given more prominence in the ECB's (internal and potentially also published) analysis of inflation developments. A second step could be an interim period starting in early 2023 with the publication by Eurostat of an experimental HICP-H index. Publication of an official quarterly HICP augmented by OOH costs

---

[242] See also overview of ECB listens events.

could be envisaged to start in 2026. A time frame for an NA approach at a monthly frequency cannot currently be set, given the lack of underlying data sources.

In addition to integrating OOH into the HICP, further improvements, especially with respect to the harmonisation and standards of the HICP, are needed. For instance, the methodology of the HICP should be further harmonised across countries, especially regarding the treatment of product replacement, quality adjustment and sampling issues. In this context, a framework for calculating sampling errors for the HICP would be beneficial, to judge its statistical quality. The availability of more timely representative information about consumption structures would also be desirable, for example through the publication of preliminary HICP weights, before they are actually applied. More transparency around the integration of new data and new methods by the ESS would also be important. Providing access to full methodological details about the compilation of the HICPs and the underlying price and weight data would be very helpful for analyses and assessment within the Eurosystem. Many of the above-mentioned improvements in harmonisation and transparency could also help approximate and, importantly, reduce the HICP measurement bias in the euro area.

The paper finds that auxiliary inflation measures can play an important role in the ECB's economic and monetary analyses. This does not only hold for analytical series of OOH-augmented HICPs, but also for measures of underlying inflation or for a COLI.

The ECB should continue to monitor a broad set of underlying inflation measures. Measures of underlying inflation include price indices for sub-aggregates of the consumption basket, stripping out volatile components, and model-based indicators. The aim of underlying inflation measures is to help distinguish signal from noise in inflation developments and thereby gain a better sense of persistent developments. The monitoring of such measures by the ECB has proven useful in understanding and explaining HICP developments and in signalling how inflation is likely to evolve over the medium-term horizon. As the measures, taken individually, do not consistently give very precise or reliable signals, a broad range of measures of underlying inflation should be monitored. Regular and consistent references to measures of underlying inflation may be required in public communication because in recent years it has been used in the context of assessing the sustained adjustment path of inflation and state-dependent forward guidance. This suggests that underlying inflation measures play an important role as an operational device in achieving the price stability objective.

Establishing an experimental COLI could enhance the analysis of inflation in the euro area.[243] While COGIs, like the HICP, are designed to measure changes in the purchasing power of money, a COLI measures the change in minimum expenditure required by a household to purchase a basket of goods and services

---

[243] Further research on a COLI is already envisaged in the HICP Framework Regulation ((EU) 2016/792 of 11 May 2016), which states in recital 12: "In addition to the HICP, research on a harmonised cost of living index should be initiated."

that preserves a certain standard of living.[244] A COLI could also incorporate non-market prices in a transparent manner and allow for more frequent changes in consumption weights. Challenges for computing an experimental COLI for the euro area, which might require substantial resources, would be the need for more frequent collections of consumption weights and, in particular, the estimation of prices and quantities for the consumption of non-market goods.

Assigning an important role to auxiliary measures of inflation has notable implications for the ECB's communication. First, the methodology behind analytical series such as measures of underlying inflation, a COLI or an experimental HICP-H would need to be publicly available and the use of these series in monetary policy would need to be clearly communicated.[245] Second, there would have to be clarity around the difference between the HICP as the ECB's main measure for price stability and other auxiliary indicators that form part of the information set used to gauge the risks to price stability.

Especially during the long implementation period of integrating OOH into the HICP, careful communication would be necessary to avoid confusion about the ECB's quantitative measure of price stability. The ECB should clearly communicate that the existing HICP would remain the single gauge for formulating and assessing its price stability objective until such time as it was replaced by an official HICP-H produced and published by Eurostat. However, during the interim period more emphasis could be placed on an experimental HICP series encompassing OOH costs as an auxiliary indicator. Communication would also need to address the risk that the HICP-H may not completely fulfil public expectations with regard to the inclusion of housing costs in the HICP (i.e. the changes with respect to the HICP could be considered too small), especially in parts of the euro area where house prices have been rising significantly and for prolonged periods.

---

[244] One example for a COLI is the PCE price index, which is the index on which the Federal Reserve defines its target.

[245] Transparency requires exhaustive and detailed documentation on the methodology and availability of the most important "ingredients" – a knowledgeable researcher should be able to replicate an analytical measure.

References

Ahnert, H. and Kenny, G. (2004), "Quality Adjustment of European Price Statistics and the Role for Hedonics", Occasional Paper Series, No 15, ECB.

Albidgren, K. and Kuchler, A. (2019), "Revisiting the inflation perception conundrum", Working Paper, No 144, Danmarks Nationalbank.

Álvarez, L.J. and Cuenca, J.A. (2020), "Medium term inflation as underlying inflation measures", background slides, Banco de España.

Amann, J., Karádi, P. and Sánchez Bachiller, J. (2020), "Price setting in supermarkets on the two sides of the Atlantic", mimeo.

Arioli, R., Bates, C., Dieden, H.C., Duca, I., Friz, R., Gayer, C., Kenny, G., Meyler, A. and Pavlova, I. (2017), "EU consumers' quantitative inflation perceptions and expectations: an evaluation", Occasional Paper Series, No 186, ECB.

Aucremanne, L., Collin, M. and Stragier, T. (2007), "Assessing the Gap between Observed and Perceived Inflation in the Euro Area: Is the Credibility of the HICP at Stake?", Working Paper Research, No 112, Nationale Bank van België/Banque Nationale de Belgique.

Australian Bureau of Statistics (2017), Measuring price change of attached dwellings in the CPI.

Bańbura, M. and Bobeica, E. (2020), "PCCI – a data-rich measure of underlying inflation in the euro area", Statistics Paper Series, No 38, ECB.

Bertolotto, M. (2019), "Online Price Index with Product Replacement: The Closest-Match Approach", paper presented at the 16th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro.

Białek, J. and Bobel, A. (2019), "Comparison of Price Index Methods for CPI Measurement Using Scanner Data", paper presented at the 16th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro.

Binder, C. (2075), "Measuring uncertainty based on rounding: new method and application to inflation expectations", Journal of Monetary Economics, Vol. 90, pp. 1-12.

Bishop, G. (2013), "Geometric Indexes and Substitution Bias in the CPI", paper presented by van M. Kints at the 13th meeting of the Ottawa Group on Price Indices, Copenhagen.

Blaudow, C. and Burg, F. (2018), "Dynamic pricing as a challenge for consumer price statistics", EURONA, No 1/2018, pp. 80-93.

Blaudow, C. and Seeger, D. (2019), "Fortschritte beim Einsatz von Web Scraping in der amtlichen Verbraucherpreisstatistik – ein Werkstattbericht" [Progress in employing web scraping in official consumer price statistics – a workshop report], WISTA – Wirtschaft und Statistik, No 4, pp. 19-30.

Bureau of Labor Statistics (2011), "Differences between the Consumer Price Index and the Personal Consumption Expenditures Price Index", Vol. 2, No 3, May.

Boskin, M., Dulberger, E., Gordon, R., Griliches, Z. and Jorgenson, D. (1996), "Final Report of the Advisory Commission to Study the Consumer Price Index", U.S. Government Printing Office.

Boettcher, I. (2015), "Automatic data collection on the Internet (web scraping)", paper presented at the 14th Meeting of the Ottawa Group on Price Indices, Tokyo.

Brachinger, H.-W., Schips, B. and Stier, W. (2000), Expertise zur Relevanz des „Boskin-Reports" für den schweizerischen Landesindex der Konsumentenpreise [Expert report regarding the relevance of the Boskin report for the Swiss national index of consumer prices], Swiss Federal Office for Statistics.

Braun, R. and Lein, S. (2021), "Sources of Bias in Inflation Rates and Implications for Inflation Dynamics", Discussion Paper Series, No 15663, Centre for Economic Policy Research.

Bruine de Bruin, W., Vanderklaauw, W., Downs, J.S., Fischhoff, B., Topa, G. and Armantier, O. (2010), "Expectations of Inflation: The Role of Demographic Variables, Expectation Formation, and Financial Literacy", Journal of Consumer Affairs, Vol. 44, No 2, pp. 381-402.

Bryan, M. and Meyer, B. (2010), "Are some prices in the CPI more forward-looking than others? We think so", Economic Commentary, No 2, Federal Reserve Bank of Cleveland.

Bureau of Labor Statistics (2020), "Variance Estimates for the Consumer Price Indexes", accessed on 28 May 2020.

Burke, M.A. and Manz, M. (2014), "Economic Literacy and Inflation Expectations: Evidence from a Laboratory Experiment", Journal of Money, Credit and Banking, Vol. 46, No 7, pp. 1421-1456.

Camba-Mendez, G. (2003), "The definition of price stability: choosing a price measure", in Issing, O. (ed.), Background Studies for the ECB's Evaluation of its Monetary Policy Strategy, ECB, pp. 32-42.

Camba-Mendez, G., Gaspar, V. and Wynne, M. (2001), "Measurement Issues in European Consumer Price Indices and the Conceptual Framework of the HICP", Summary and Conclusions of the CEPR/ECB Workshop on Issues in the Measurement of Price Indices, ECB, November.

Carvalho, V.M., Hansen, S., Ortiz, Á., Ramón García, J., Tomasa, R., Rodriguez Mora, S. and Ruiz, J. (2020), "Tracking the COVID-19 Crisis with High-Resolution

Transaction Data", CEPR Discussion Papers, No 14642, Centre for Economic Policy Research.

Cavallo, A. (2017), "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers", American Economic Review, Vol. 107, No 1, pp. 283-303.

Cavallo, A. (2020), "Inflation with Covid Consumption Baskets", NBER Working Paper Series, No 27352, National Bureau of Economic Research.

Chessa, A.G. (2016), "A new methodology for processing scanner data in the Dutch CPI", EURONA –Eurostat Review on National Accounts and Macroeconomic Indicators, No 1/2016, pp. 49-69.

Chessa, A.G. (2019a), "MARS: A method for defining products and linking barcodes of item relaunches", paper presented at the 16th meeting of the Ottawa Group on Price Indices, Rio de Janeiro.

Chessa, A.G. (2019b), "A comparison of index extension methods for multilateral methods", paper presented at the 16th meeting of the Ottawa Group on Price Indices, Rio de Janeiro.

Chessa, A.G., Verburg, J. and Willenborg, L. (2017), "A Comparison of Price Index Methods for Scanner Data", paper presented at the 15th meeting of the Ottawa Group on Price Indices, Eltville.

Christoffel, K., Coenen, G. and Warne, A. (2008), "The new area-wide model of the euro area. A micro-founded open-economy model for forecasting and policy analysis", Working Paper Series, No 944, ECB.

Chung, H., Doyle, B.M., Hebden, J. and Siemer, M. (2020), "Considerations Regarding Inflation Ranges", Finance and Economics Discussion Series, No 2020-075, Board of Governors of the Federal Reserve System.

Ciccarelli, C. and Osbat, C. (2017), "Low inflation in the euro area: Causes and consequences", Occasional Paper Series, No 181, ECB.

Coenen, G., Montes-Galdón, C. and Smets, F. (2020), "Effects of state-dependent forward guidance, large-scale asset purchases and fiscal stimulus in a low-interest-rate environment", Working Paper Series, No 2352, ECB.

Coibion, O., Gorodnichenko, Y. and Hong, G.H. (2015), "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications", American Economic Review, Vol. 105, No 3, pp. 993-1029.

Commission Implementing Regulation (EU) 2020/1148 of 31 July 2020 laying down the methodological and technical specifications in accordance with Regulation (EU) 2016/792 of the European Parliament and of the Council as regards harmonised indices of consumer prices and the house price index, Official Journal L 252, 4.8.2020, pp. 12-23.

Commission Regulation (EU) No 1114/2010 of 1 December 2010 laying down detailed rules for the implementation of Council Regulation (EC) No 2494/95 as regards minimum standards for the quality of HICP weightings and repealing Commission Regulation (EC) No 2454/97, Official Journal L 316, 2.12.2010, pp. 4-6.

Conflitti, C. (2020), "Alternative measures of underlying inflation in the euro area", Questioni di Economia e Finanza, Occasional Papers No 593, Banca d'Italia.

Costa, D.L. and Kahn, M.E. (2003), "The Rising Price of Nonmarket Goods", American Economic Review, Vol. 93, No 2, pp. 227-232.

Croushore, D. (2019), "Revisions to PCE Inflation Measures: Implications for Monetary Policy", International Journal of Central Banking, Vol. 15, No 4, pp. 241-265.

Cunningham, A. (1996), "Measurement Bias in Price Indices: an Application to the UK's RPI", Working Paper No 47, Bank of England, Monetary Analysis Division.

D'Acunto, F., Hoang, D., Paloviita, M. and Weber, M. (2019), "Cognitive Abilities and Inflation Expectations", AEA Papers and Proceedings, Vol. 109, pp. 562-66.

D'Acunto, F., Malmendier, U., Ospina, J. and Weber, M. (2021), "Exposure to Grocery Prices and Inflation Expectations", Journal of Political economy, forthcoming.

D'Acunto, F., Malmendier, U. and Weber, M. (2020), "Gender Roles and the Gender Expectations Gap", Chicago Booth Research Paper, No 20-03, University of Chicago.

Dalén, J. and Ohlsson, E. (1995), "Variance Estimation in the Swedish Consumer Price Index", Journal of Business & Economic Statistics, Vol 13, No 3, pp. 347-356.

Dalén, J. and Tarassiouk, O. (2013), "Replacements, quality adjustments and sales prices", paper presented at the 13th meeting of the Ottawa Group on Price Indices, Copenhagen.

de Haan, J. and van der Grient, H. (2009), "Eliminating Chain Drift in Price Indexes Based on Scanner Data", paper for the 11th Meeting of the Ottawa Group on Price Indices, Neuchâtel.

de Haan, J., Hendricks, R. and Scholz, M. (2016), "A Comparison of Weighted Time-Product Dummy and Time Dummy Hedonic Indexes", Graz Economics Papers, No 2016-13, University of Graz.

Del Giovane, P., Fabiani, S. and Sabbatini, R. (2009), "What's behind 'inflation perceptions'? A survey-based analysis of Italian consumers", Giornale degli Economisti e Annali di Economia, Nuova Serie, Vol. 68 (Anno 122), No 1, pp. 25-52.

DESTATIS (2009), "Handbook on the application of quality adjustment methods in the Harmonised Index of Consumer Prices", Statistics and Science, Vol. 13, Federal Statistical Office of Germany.

Deutsche Bundesbank (2013), "The contribution of the accommodation services sub-index to explaining the current deviations between CPI and HICP inflation rates", Monthly Report, April, p. 7.

Deutsche Bundesbank (2019a), "The revision of the sub-index for package holidays and its impact on the HICP and core inflation", Monthly Report, March, pp. 8-9.

Deutsche Bundesbank (2019b), "Dampening special effect in the HICP in July 2019", Monthly Report, August, pp. 57-59.

Deutsche Bundesbank (2019c), "The relevance of survey expectations for the Deutsche Bundesbank", Monthly Report, December, pp. 53-71.

Deutsche Bundesbank (2021), "One-off effects relating to COVID-19 in the HICP in 2021", Monthly Report, February, pp. 63-66.

Dietrich, A., Eiglsperger, M., Mehrhoff, J. and Wieland, E. (2021), "Chain-linking over December and methodological changes in the HICP – view from a central bank perspective", Statistics Paper Series, No 40, ECB.

Diev, P., Kalantzis, Y., and Lalliard, A. (2019), "Why have strong wage dynamics not pushed up inflation in the euro area?", Banque de France Bulletin, Vol. 225, No 6, September-October.

Diewert, E. (2002), "Harmonized Indexes of Consumer Prices: Their Conceptual Foundations", Working Paper Series, No 130, ECB.

Diewert, W.E. (1976), "Exact and superlative index numbers", Journal of Econometrics, Vol. 4, No 2, pp. 115-145.

Döhring, B. and Mordonu, A. (2007), "What drives inflation perceptions? A dynamic panel data analysis", European Economy – Economic Papers 2008–2015, No 284, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Dunn, A.C, Hood, K.K. and Driessen, A. (2020), "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending Using Card Transaction Data", BEA Working Paper Series, No WP2020-5, U.S. Bureau of Economic Analysis.

Ehrmann, M., Ferrucci, G., Lenza, M. and O'Brien, D. (2018), "Measures of underlying inflation for the euro area", Economic Bulletin, Issue 4, ECB.

Eiglsperger, M. (2019a), "New features in the Harmonised Index of Consumer Prices: analytical groups, scanner data and web-scraping", Economic Bulletin, Issue 2, ECB.

Eiglsperger, M. (2019b), "A new method for the package holiday price index in Germany and its impact on HICP inflation rates", Economic Bulletin, Issue 2, ECB.

Eiglsperger, M., Ganoulis, I., Roma, M. and Vlad, A. (2020), "Owner-Occupied Housing", Strategy Review Background note, ECB.

European Central Bank (1999), "The stability-oriented monetary policy strategy of the Eurosystem", Monthly Bulletin, January, pp. 39-50.

European Central Bank (2003), Press release of 8 May 2003 on the ECB's monetary policy strategy, Monthly Bulletin, May, p. 8.

European Central Bank (2005), "The Harmonised Index of Consumer Prices: Concept, properties and experience to date", Monthly Bulletin, July, pp. 55-68.

European Central Bank (2010), "Update on developments in general economic statistics for the euro area", Monthly Bulletin, February, pp. 95-107.

European Central Bank (2011a), "Methodological changes in the compilation of the HICP and their impact on recent data", Monthly Bulletin, April, pp. 38-42.

European Central Bank (2011b), The monetary policy of the ECB.

European Central Bank (2012), "New standards for HICP weights", Monthly Bulletin, April, pp. 36-39.

European Central Bank (2014), "Potential measurement issues in consumer price indices", Monthly Bulletin, April, pp. 40-42.

European Central Bank (2016a), "Harmonised Index of Consumer Prices – Easter effects and improved seasonal adjustment", Economic Bulletin, Issue 3, pp. 17-20.

European Central Bank (2016b), "What accounts for the recent decoupling between the euro area GDP deflator and HICP excluding energy and food", Economic Bulletin, Issue 6, pp. 46-48.

European Central Bank (2020a), "Inflation Concept and its implementation", Strategy review background paper.

European Central Bank (2020b), "NSI web scraping survey – results", presentation for the meeting of the ESCB WG General Economic Statistics.

European Commission (2018), Report from the Commission to the European Parliament and the Council on the suitability of the owner-occupied housing (OOH) price index for integration into the harmonised index of consumer prices (HICP) coverage.

Eurostat (2017a), Harmonised Index of Consumer Prices – Practical Guide for Processing Supermarket Scanner Data.

Eurostat (2017b), Technical manual on Owner-Occupied Housing and House Price Indices.

Eurostat (2018), "Harmonised Index of Consumer Prices (HICP) – Methodological Manual".

Eurostat (2020a), "Guidance note on the compilation of HICP weights in case of large changes in consumer expenditures".

Eurostat (2020b), "Practical guidelines on web scraping for the HICP".

Eurostat (2020c), "Recommendation on the Bridged Overlap method", document PSWG12.2020/08 – Annex for the meeting of the Price Statistics Working Group, 25-26 November 2020.

Eurostat (2021), HICP reference metadata, accessed on 2021-01-20.

Eusepi, S., Hobijn, B. and Tambalotti, A. (2011), "CONDI: A Cost-of-Nominal-Distortions Index", American Economic Journal: Macroeconomics, Vol. 3, No 3, pp. 53-91.

Ferrucci, G. (2018), "The use of measures of underlying inflation at selected central banks", Economic Bulletin, Issue 4, Box 1, ECB.

Fluch, M., Fritzer, F. and Rumler, F. (2013), "Inflation und Preiswahrnehmung: eine Bestandsaufnahme für Österreich", Statistiken – Daten und Analysen, Q4/13, Oesterreichische Nationalbank, pp. 67-85.

Fritzer, F. and Rumler, F. (2015), "Determinants of Inflation Perceptions and Expectations: an Empirical Analysis for Austria", Monetary Policy & the Economy, Issue 1, Oesterreichische Nationalbank, pp. 11-26.

Gábor-Tóth, E. and Vermeulen, P. (2018), "Elementary Index Bias: Evidence from a Large Scanner Dataset", German Economic Review, Vol. 20, No 4, pp. 618-656.

Fröhling, A., O'Brien, D. and Schaefer, S. (2021), "A new measure for domestic inflation in the Euro Area", mimeo.

Goldhammer, B., Henkel, L. and Osiewicz, M. (2019), "Bias related to the Bridged-overlap- and Link-to-Show-No-Price-Change Method", poster presented at the 16th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro.

Goldhammer, B., Kouvavas, O., Beka, J. and Krasnopjorovs, O. (2021), Innovative products and e-commerce - survey results", paper presented at the HICP workshop, 29-30 June., mimeo.

Gonçalves, E., Henkel, L., Kouvavas, O., Porqueddu, M. and Trezzi, R. (2021), "2021 HICP weights and their implications for the measurement of inflation", Economic Bulletin, Issue 2, ECB.

Goodhart, C. (2001), "What Weight Should be Given to Asset Prices in the Measurement of Inflation?", The Economic Journal, Vol. 111, Issue 472, pp. 335-356.

Gorodnichenko, Y., Sheremirov, V. and Talavera, O. (2018), "Price Setting in Online Markets: Does IT Click?", Journal of the European Economic Association, Vol. 16, Issue 6, pp. 1764-1811.

Greenless, J. and McClelland, R. (2008), "New Evidence on Outlet Substitution Effects in Consumer Price Index Data", Working Paper Series, No 421, U.S. Bureau of Labor Statistics.

Greenless, J. and Williams, E. (2009), "Reconsideration of Weighting and Updating Procedures in the US CPI", paper presented at the 11th meeting of the Ottawa Group, Neuchâtel.

Griffioen, R. and ten Bosch, O. (2016), "On the Use of Internet Data for the Dutch CPI", paper presented at the UNECE Meeting of the Group of Experts on Consumer Price Indices, Geneva.

Groshen, E., Moyer, B., Aizcorbe, A., Bradley, R. and Friedman, D. (2017), "How Government Statistics Adjust for Potential Biases from Quality Change and New Goods in an Age of Digital Technologies: A View from the Trenches", Journal of Economic Perspectives, Vol. 31, No. 2, pp. 187-210.

Guðnason, R. (2003), "How do we measure inflation? Some measurement problems", paper presented at the 7th meeting of the Ottawa Group on Price Indices, Paris, May.

Hacioglu, S., Känzig, D.R. and Surico, P. (2020), "Consumption in the time of COVID-19: evidence from UK transaction data", CEPR Discussion Paper Series, Centre for Economic Policy Research, No 14733.

Handbury, J., Watanabe, T. and Weinstein, D.E. (2013), "How Much Do Official Price Indexes Tell Us About Inflation?", NBER Working Paper, No 19504, National Bureau of Economic Research.

Hayman, C. (2006), "Outlet Substitution Bias in the Canadian Consumer Price Index: A Case Study", Analytical Series, Research Paper No 18, Statistics Canada, Prices Division.

Herzberg, J., Knetsch, T. and Schwind, P. (2020), "Back-calculation of the OOHPI and the HICP including owner-occupied housing for Germany to the start of the HICP series", mimeo, Deutsche Bundesbank, Directorate General Statistics.

Herzberg, J., Knetsch, T., Schwind, P. and Weinand, S. (2021a), "Quantifying bias and inaccuracy of upper-level aggregation in HICPs for Germany and the euro area", Discussion Paper, No 6/2021, Deutsche Bundesbank.

Herzberg, J., Knetsch, T., Popova, D., Schwind, P. and Weinand, S. (2021b), "Empirical findings on upper-level aggregation issues in the HICP", mimeo, Deutsche Bundesbank, Directorate General Statistics.

Hoeberichts, M.M. and Zekaite, Z. (2020), "Rent controls in Europe", mimeo, De Nederlandsche Bank/Central Bank of Ireland.

Hoffmann, J. (1998), "Problems of inflation measurement in Germany", Discussion Paper, No 1/98, Deutsche Bundesbank, Economic Research Group.

Hohnerlein, E.M. (2019), "Pension indexation for retirees revisited – Normative patterns and legal standards", Global Social Policy, Vol. 19, No 3, pp. 246-265.

Huynh, K., Lao, H., Sabourin, P. and Welte, A. (2020), "What do high-frequency expenditure network data reveal about spending and inflation during COVID-19?", Staff Analytical Note, No 2020-20 (English), Bank of Canada.

Instituto Nacional de Estatística (2009), "Final report on land price separation study".

International Labour Organization, International Monetary Fund, Organisation for Economic Co-operation and Development, Statistical Office of the European Communities, United Nations, The International Bank for Reconstruction and Development, The World Bank (2004): "Consumer price index manual: Theory and practice".

International Labour Office, International Monetary Fund, Organisation for Economic Co-operation and Development, European Union, United Nations, World Bank (2020): "Consumer price index manual: concepts and methods".

Issing, O. (ed.) (2003), Background Studies for the ECB's evaluation of its Monetary Policy Strategy, ECB.

Ivancic, L., Diewert, W.E. and Fox, K.J. (2011), "Scanner Data, Time Aggregation and the Construction of Price Indexes", Journal of Econometrics, Vol. 161, No 1, pp. 24-35.

Jaravel, X. and O'Connell, M. (2020), "Inflation Spike and Falling Product Variety During the Great Lockdown", CEPR Discussion Papers, No 14880, Centre for Economic Policy Research.

Kalantzis, Y., Knetsch, T., Wintr, L., Roma, M., Hoeberichts, M., Eiglsperger, M. and Lünnemann, P. (2021), "Roadmap to a full integration of OOH in the HICP according to the net acquisition approach via implementing an interim parallel system of monthly HICP and quarterly HICP-H", background note, MPC/STC Work Stream "Inflation Measurement".

Kalantzis, Y., Kouvavas, O. and Zekaite, Z. (2021), "Roadmap for OOH Rental Equivalence", background note, MPC/STC Work Stream "Inflation Measurement".

Kalantzis, Y., Lalliard, A. and Vertier, P. (2020a), "OOHPI indices", mimeo, Banque de France, Directorate General Economics Statistics and International Relations.

Kalantzis, Y., Lalliard, A. and Vertier, P. (2020b), "How sensitive is HIPC-H to the weight of OOH?", mimeo, Banque de France, Directorate General Economics Statistics and International Relations.

Kaplan, G. and Menzio, G. (2015), "The Morphology of Price Dispersion", International Economic Review, Vol. 56, No 4, pp. 1165-1206.

Kaplan, G. and Schulhofer-Wohl, S. (2017), "Inflation at the household level", Journal of Monetary Economics, Vol. 91, pp. 19-38.

Keating, J. and Murtagh, M. (2018), "Quality adjustment in the Irish CPI", paper presented at the meeting of the Group of Experts on Consumer Price Indices, Geneva, May.

Kouvavas, O., Trezzi, R., Eiglsperger, M., Goldhammer, B. and Gonçalves, E. (2020), "Consumption patterns and inflation measurement issues during the COVID-19 pandemic", Economic Bulletin, Issue 7, ECB.

Kryvtsov, O. (2016), "Is there a quality bias in the Canadian CPI? Evidence from microdata", The Canadian Journal of Economics, Vol. 49, No 4, pp. 1401-1424.

Lalliard, A. and Kalantzis, Y. (2020), "The fine core index of underlying inflation", background slides, Banque de France.

Lebow, D. and Rudd, J. (2003), "Measurement Error in the Consumer Price Index: Where Do We Stand?", Journal of Economic Literature, Vol. XLI, pp. 159-201.

Leifer, H.-A. (1999), "Zur Berechnung von Preismeßzahlen auf der "Elementarebene" des Harmonisierten Verbraucherpreisindex (HVPI)" [About the calculation of "Elementary Aggregate Indices" of the HICP], Allgemeines Statistisches Archiv, Vol. 83, pp. 338-349.

Lequiller, F. (1997), "Does the French consumer price index underestimate inflation?", Série des documents de travail de la Direction des Etudes et Synthèses Économiques, No G 9714, Insee.

Luciani, M. and Trezzi, R. (2019), "Comparing Two Measures of Core Inflation: PCE Excluding Food & Energy vs. the Trimmed Mean PCE Index", Feds Notes.

Lünnemann, P. and Wintr, L. (2021), "Properties of HICP augmented with owner-occupied housing in the euro area: backcasting and scenario analysis", mimeo, Banque centrale du Luxembourg.

Malmendier, U. and Nagel, S. (2016), "Learning from Inflation Experiences", The Quarterly Journal of Economics, Vol. 131, No 1, pp. 53-87.

Mehrhoff, J. (2019), "Introduction – The Value Chain of Scanner and Web Scraped Data", Economie et Statistique / Economics and Statistics, No 509, Insee, pp. 5-11.

Moulton, B. (2018), "The Measurement of Output, Prices, and Productivity: What's Changed Since the Boskin Commission?", Brookings Institution.

Nakov, A. (2020), "Costs of Nominal Distortions Index", mimeo, ECB.

National Institute of Statistics and Economic Studies (2020), "How to compute a Consumer Price Index in the context of the Covid-19 crisis?", methodological note.

Neves, P. and Sarmento, L. (1997), "The Substitution Bias of the Consumer Price Index", Economic Bulletin, Banco de Portugal, pp. 25-33.

Office for National Statistics (2020), Re-weighted consumer prices basket – adjusting for consumption changes during lockdown, July.

Office for National Statistics (2019), "Consumer Prices Indices Technical Manual"

Regulation (EU) 2016/792 of the European Parliament and of the Council of 11 May 2016 on harmonised indices of consumer prices and the house price index, and repealing Council Regulation (EC) No 2494/95, Official Journal L 135, 24.5.2016, pp. 11-38.

Reinsdorf, M. (2020), "COVID-19 and the CPI: Is Inflation Underestimated?", IMF Working Paper, No 20/224, International Monetary Fund.

Rubbo, E. (2020), "Networks, Phillips curves, and monetary policy", manuscript, Harvard University.

Rubene, I. (2020), "Recent developments in euro area food prices", Economic Bulletin, Issue 5, ECB.

Rumler, F. and Valderrama, M.T. (2020), "Inflation literacy and inflation expectations: Evidence from Austrian household survey data", Economic Modelling, Vol. 87(C), pp. 8-23.

Sabourin, P. (2012), "Measurement Bias in the Canadian Consumer Price Index: An Update", Bank of Canada Review, Summer 2012, pp. 1-12.

Schultze, C. (2003), "The Consumer Price Index: Conceptual Issues and Practical Suggestions", Journal of Economic Perspectives, Vol. 17, No 1, pp. 3-22.

Schultze, C. and Mackie, C. (eds.) (2002), At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes, National Academy Press.

Sims, C.A. (2010), "Rational Inattention and Monetary Economics", Handbook of Monetary Economics, Vol. 3, Chapter 4, pp. 155-181.

Surico, P., Känzig, D. and Hacioglu, S. (2020), "Consumption in the time of Covid-19: Evidence from UK transaction data", CEPR Discussion Papers Series, No 14733, Centre of Economic Policy Research.

The Federal Reserve Boards (2019), Monetary Policy and the Economic Outlook.

van der Cruijsen, C.A.B., Jansen, D.-J. and de Haan, J. (2015), "How Much Does the General Public Know about the ECB's Monetary Policy? Evidence from a Survey of Dutch Households", International Journal of Central Banking, Vol. 11, No 4, pp. 169-218.

Van Loon, K. and Roels, D. (2018), "Integrating big data in the Belgian CPI", paper presented at the meeting of the Group of Experts on Consumer Price Indices, Geneva, May.

van Overbeek, F. and Hoeberichts, M. (2020), "Can the costs of home ownership be included in headline and core inflation?", mimeo, De Nederlandsche Bank.

von Auer, L. (2016), "Processing Scanner Data by an Augmented GUV Index", internal report, ECB.

von Auer, L. (2019), "The nature of chain drift: Implications for Scanner Data Price Indices", paper presented at the 16th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro, May.

von der Lippe, P. (2007), Index Theory and Price Statistics, Peter Lang Publishing.

Wynne, M. (2005), "An Estimate of the Measurement Bias in the HICP", Working Paper, No 0509, Federal Reserve Bank of Dallas, Research Department.

Zafar, J.-D. (2019), "Distinguishing price and quality effects for electronic goods: how information collected online can help estimate hedonic models", presentation for the meeting on the measurement of prices and volumes in a digitalised world, OECD.

Zekaite, Z. (2020a), "Underlying inflation measure based on time-varying volatility and persistence", internal note, Central Bank of Ireland.

Zekaite, Z. (2020b), "What drives consumers' inflation perceptions in the euro area?", Economic Letters, 06/EL/20, Central Bank of Ireland.

Zimmer, H. and Jonckheere, J. (2020), "Consumer prices in light of the COVID-19 crisis", NBB Economic Review, Nationale Bank van België/Banque Nationale de Belgique.