

# Guide on producing CPI under lockdown



UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

# Guide on producing CPI under lockdown



**United Nations**

Geneva, 2021

©2021 United Nations  
All rights reserved worldwide

Requests to reproduce excerpts or to photocopy should be addressed to the Copyright Clearance Center at [copyright.com](http://copyright.com).

All other queries on rights and licenses, including subsidiary rights, should be addressed to: United Nations Publications, 405 East 42nd St, S-09FW001, New York, NY 10017, United States of America. Email: [permissions@un.org](mailto:permissions@un.org);  
website: <http://shop.un.org/>.

The findings, interpretations, and conclusions expressed herein are those of the author(s) and do not necessarily reflect the views of the United Nations or its officials or Member States.

United Nations publication issued by the United Nations Economic Commission for Europe.

This publication is issued in English.

ECE/CES/STAT/2021/3

eISBN: 978-92-1-005896-4  
ISSN: 0069-8458

## Preface

The lockdown that followed the outbreak of the Covid-19 pandemic in 2020 posed unprecedented challenges to produce the consumer price index (CPI) in many countries. Regular price collection for the CPI could not take place from sectors of the economy and retail stores that were closed. Price collection from outlets that remained open was in many cases prevented because price collectors were not allowed to travel or entering outlets to collect prices due to safety measures or social distancing. As a result, the lockdown caused drops in the number of collected prices beyond historic norms. To ensure the compilation of a reliable CPI, statistical offices therefore had to explore new data sources and new ways of collecting prices and implement methods to make up for missing price observations and changes in data sources and collection methods. The lockdown also created challenges for the dissemination and communication of the CPI as users requested additional information about the lockdown's impact on the CPI.

Based on countries' experiences with producing the CPI during the pandemic, this Guide provides recommendations and emerging best practices for CPI data collection, calculation methods and communication under lockdown conditions. Countries may refer to the Guide in future periods of lockdown or similar exceptional circumstances as they find it useful.



# Contents

<b>Preface</b> .....	<b>iii</b>
<b>Chapter 1 Introduction</b> .....	<b>1</b>
1.1 Background .....	1
1.2 Purpose of the Guide .....	2
1.3 Overview of chapters .....	2
1.4 More information .....	3
1.5 Acknowledgments .....	4
<b>Chapter 2 Data Collection</b> .....	<b>5</b>
2.1 Introduction .....	5
2.2 Organisation of price collection .....	5
2.3 Communication with outlets and respondents .....	6
2.4 Alternative modes of price collection .....	8
2.5 Maintaining the principles of CPI price recording .....	10
2.6 Issues to consider when changing mode of price collection .....	13
2.7 A permanent level shift in households' purchases on Internet .....	14
2.8 Recommencing regular price collection after lockdown .....	16
2.9 Making price collection more resilient in the long-term .....	18
2.10 Key points .....	19
<b>Chapter 3 Imputation</b> .....	<b>21</b>
3.1 Introduction .....	21
3.2 Imputation defined .....	21
3.3 Imputation methods .....	23
3.3.1 Imputation for available products .....	23
3.3.2 Imputation for unavailable products .....	25
3.3.3 Products with pronounced seasonal price variations .....	28
3.3.4 Seasonal products .....	29
3.3.5 Self-correction .....	29
3.4 Compiling the imputation rate .....	30
3.5 Change in data collection mode .....	32
3.6 Data validation .....	33
3.7 Choice of imputation method .....	33
3.8 Key points .....	36
<b>Chapter 4 Expenditure Weights</b> .....	<b>37</b>
4.1 Introduction .....	37
4.2 Impact on weights of a lockdown .....	37
4.3 Estimation of weights for future CPIs based on data from lockdown periods .....	38
4.4 Key points .....	40

<b>Chapter 5</b>	<b>Communication .....</b>	<b>41</b>
5.1	Introduction .....	41
5.2	Communication with users and stakeholders .....	41
5.3	Informing users about the quality and reliability of the CPI .....	42
5.3.1	Publication of heavily imputed index series.....	43
5.3.2	Ratio of imputed prices .....	44
5.3.3	Ratio of closed outlets .....	44
5.3.4	Indicators of data source and price collection method.....	44
5.3.5	Measures of statistical certainty.....	46
5.3.6	Sensitivity analysis .....	47
5.3.7	Decomposition.....	47
5.4	Experimental CPIs .....	49
5.4.1	Experimental CPI based on available products only.....	49
5.4.2	Experimental CPI based on estimated lockdown weights.....	49
5.4.3	Data sources for expenditure weights.....	50
5.4.4	Publication of special aggregates .....	51
5.5	Key points .....	52
<b>References</b>	<b>.....</b>	<b>53</b>

# Chapter 1 Introduction

## 1.1 Background

1.1 The outbreak of the Covid-19 pandemic in early 2020 and the measures implemented in countries to reduce its spread raised unprecedented challenges for national statistical offices around the world. In many countries different sectors of the economy were closed and there were restrictions on the movement of people and social distancing. Staff in statistical offices were not able or allowed to work or were asked to work remotely. This made it difficult to maintain data collection and the regular compilation and publication of official statistics. In many cases the usual statistical operations could not be carried out and the lack of data raised methodological and conceptual questions in the calculation of the statistics.

1.2 In most countries the consumer price index (CPI) is compiled and published monthly adhering to a tight production schedule to allow time for collection and validation of data and index compilation and ensuring publication on a pre-announced release date. The production of the CPI was therefore immediately and directly impacted by the pandemic and the resulting lockdowns, which created an abrupt surge in missing and potentially unusable input data for CPI compilation.

1.3 The regular price collection for the CPI was hampered because of temporary closure of outlets and difficulties with collecting prices from outlets that remained open. Some businesses and retail establishments were closed by government order, others were temporarily forced to close due to lack of supply or lack of demand, or because outlet employees were prevented from working. When outlets remained open, price collectors in many instances were not allowed to travel or not allowed to enter the outlets to collect prices for health and safety reasons, while statistical offices had to ensure that collection activities did not contravene emergency rules or place staff at risk. Outlets that report prices on questionnaires in some cases were not able to submit prices for the full set of sampled products because of disrupted supply chains preventing inventory from getting to retail locations or were delayed in reporting the prices.

1.4 The lockdown led to significant drops in the number of collected prices and demonstrated the vulnerability of depending on a single data source or a single data collection mode. To collect as many prices as possible statistical offices had to consider alternative data sources and different modes of price collection and how to integrate price data from different sources in the CPI compilation. The lockdown thus underlined the need of being flexible and innovative in responding to challenges to normal statistical operations.

1.5 Missing prices is a regular feature in the CPI compilation dealt with according to established procedures. However, the magnitude of missing observations because of the pandemic created problems that could not be dealt with by usual procedures. In some cases, it was not possible to collect prices at all because markets were closed. In many countries this happened for international and domestic airline travels, retail trade, personal services such as hairdressers etc., childcare and sports and cultural events. Statistical offices therefore needed to consider how to manage missing observations efficiently and apply suitable methods to adjust for the missing observations and ensuring the quality and reliability of the CPI.

1.6 The lockdown also created challenges for the dissemination and communication of the CPI. The CPI is used as the general measure of inflation and plays a key role in economic and monetary policy. It is also widely used for indexation of wages, pensions and social transfers and for indexation of contracts. Thus, the CPI is used by a large and diverse group of users for different purposes and attracts much interest

by households and the public in general. The pandemic underlined users' need for timely and relevant statistics and the challenges for statistical offices to meet the needs of different user groups.

1.7 The lockdown also raised questions about its effect on the CPI and the relevance and the quality of the index. Statistical offices therefore also had to communicate carefully and provide enough documentation and explanations to allow correct interpretation and use of the CPI and maintain the public trust in the index. Proactive communication with users is crucial for maintaining trust in official statistics during crises, and to manage and raise awareness of trade-offs among the quality criteria for statistics, e.g. timeliness, relevance and accuracy.

## 1.2 Purpose of the Guide

1.8 The purpose of this Guide is to provide recommendations and examples of emerging best practices for producing the CPI under lockdown and other exceptional circumstances. It aims to assist countries in maintaining the quality and international comparability of CPIs under periods of lockdown.

1.9 The *Consumer Price Index Manual: Concepts and Methods, 2020* (hereafter referred to as the *CPI Manual*) is the international statistical standard for compiling the CPI. It provides recommendations on sampling and data collection, data processing and calculation methods for the CPI under normal conditions. The manual does not give guidance on the production of the CPI under lockdown. The Intersecretariat Working Group on Price Statistics (IWGPS) therefore in May 2020 issued the *Consumer Price Index: Business Continuity Guidance*, which gives recommendations on data collection, index compilation and imputation methods and communication during the pandemic.

1.10 The recommendations of this Guide are in line with the *CPI Manual* and the IWGPS Business Continuity Guidance for CPI. The Guide complements these by presenting additional and more detailed guidance and examples of good practices for producing the CPI under lockdown, based on countries' experience with producing the CPI during the 2020-2021 pandemic.

1.11 Lockdown, in this guide, is understood as a situation where it is not possible to collect enough prices for the sample of goods and services to rely on the regular compilation methods for the CPI. It may cover the whole of a country or only certain areas and it may affect different markets of goods and services in different ways. A lockdown may be caused by pandemics or epidemics, natural disasters, conflicts, or unrests. Because of the drop in the number of price observations associated with a lockdown, suitable methods to ensure the quality of the CPI must be applied. This includes the utilisation of alternative data sources and modes of price collection and implementation of methods to adjust for missing observations or changes in data sources or mode of price collection.

1.12 Countries may refer to the Guide for establishing contingency plans that can be implemented in cases of lockdown and in developing multi-source and multi-mode data collection approaches and more resilient production systems for the CPI. Suitable procedures should be established in terms of contingency plans. Ideally, contingency plans in terms of back-up data collection alternatives and imputation methods should be integrated in the production procedures to be ready for implementation when required.

## 1.3 Overview of chapters

1.13 *Chapter 2 Data collection* provides practical advices on the conducting of field price collection under lockdown and communication with price collectors and outlets. It presents alternative data sources and alternative data collection methods, such as websites, newspapers advertisements, telephone, e-mail, in person price collection, scanner data, big data and their advantages and disadvantages. Use of multiple data sources and multiple data collection methods is recommended while NSOs must be aware of

associated problems with sample representativity, weighting and quality differences. The chapter provides guidance on the treatment of extra charges and delivery charges that may be introduced during the lockdown. The chapter advises on recommencement of regular price collection following the lifting of lockdown and provides long-term guidance on how to make price collection more resilient and moving towards a multi-source/multi-mode environment.

1.14 *Chapter 3 Imputation and index calculation* presents the most used imputation methods. It makes a distinction between products that are still available on the market and unavailable products and recommends imputation practices in these two cases. The chapter gives guidance on the treatment of products with pronounced seasonal price variation, adjustments that may be necessary when changing the mode of data collection, imputation principles, data validation and implementation of imputation rules and decision trees.

1.15 *Chapter 4 Expenditure weights*, firstly, discusses the possible impact on the weights of a lockdown. Secondly, it discusses how to derive weights for future CPIs based on consumption data from a lockdown period and present examples of different approaches.

1.16 *Chapter 5 Communication* stresses the importance of being transparent by publishing suitable documentation and explanations to facilitate correct interpretation and use of the CPI. The chapter gives recommendations for communicating with users and stakeholders under lockdown. It provides examples of additional information about the quality and reliability of the CPI that may be compiled to inform users, including measures of the impact of the lockdown such as imputation ratios. It also provides examples of experimental or analytical CPIs that may be compiled to supplement the headline CPI.

## 1.4 More information

1.17 Many statistical offices have posted documentation on CPI methods and practices under the pandemic on their websites. Some of these are referred to in the chapters. International organisations including Eurostat, ILO, IMF, OECD and UNECE, in spring 2020 established websites for sharing experiences in producing the CPI during the lockdown. The websites (see Box 1.1) include papers and presentations from national statistical offices, international organisations and researchers.

### Box 1.1 Websites with material on producing CPI under lockdown

The following websites contain selected papers and presentations on producing the CPI under lockdown by countries, international organisations and researchers.

**Eurostat:** <https://ec.europa.eu/eurostat/data/metadata/covid-19-support-for-statisticians>

**IMF:** <https://www.imf.org/en/Publications/SPROLLS/covid19-special-notes#stats>

**ILO:** <https://ilostat.ilo.org/topics/covid-19/>

**OECD:** <https://community.oecd.org/community/official-stats-workspace-covid19>

**UNECE:** <https://statswiki.unece.org/x/roKSE>

**UNECE/Ottawa Group Webinar on Producing CPI under Lockdown, October-November 2020:**  
<https://unece.org/statistics/events/webinars-producing-cpi-under-lockdown>

**UNECA Regional Seminar on data collection for the CPI in combat with Covid-19, May 2020:**  
<https://ecastats.unece.org/acswweb/FocusAreas/ESNA/CPI2020.aspx>

**Websites of national statistical offices:**

Many countries have published additional material and analysis on the website of the statistical office.

## 1.5 Acknowledgments

1.18 To facilitate continuing sharing of experiences among CPI compilers and more detailed discussions, UNECE in cooperation with the Ottawa Group on Price Indices organized a webinar on the production of CPI under lockdown, which took place in terms of four online sessions in October and November 2020. The webinar discussed theoretical issues and challenges and best practices in data collection, imputation methods and communication with users under lockdown. The webinar, which attracted global interest, was based on presentations by experts from countries, international organisations and researchers.

1.19 In October 2020, the Bureau of the Conference of European Statisticians requested the UNECE Steering Group on Consumer Price Indices to summarise recommendations and good practices for producing CPI under lockdown in a reference document that countries may refer to in similar situations in the future. As a follow-up to the decision of the Bureau, a group of experts established under the Steering Group was formed to summarise recommendations and good practices based on country experiences and in line with the *CPI Manual* and the Business continuity Guide of the IWGPS.

1.20 The Guide was drafted by Giorgi Tetrauli (National Statistics Office of Georgia), Federico Polidoro (National Institute of Statistics of Italy), Chris Jenkins (Office for National Statistics, United Kingdom), Rob Cage (Bureau of Labor Statistics, United States), Claude Lamoray (Eurostat) and Carsten Boldsen (editor, UNECE). Australian Bureau of Statistics and National Institute of Statistics and Economic Studies, France, contributed with country examples.

# Chapter 2 Data Collection

## 2.1 Introduction

2.1 Lockdown measures will present national statistical offices (NSOs) with numerous challenges for the CPI price collection, depending on the severity of the restrictions in place. In the most serious circumstances, routine price collection could fail due to:

- Price collectors being unable to travel to physically collect prices.
- Lockdown restrictions limiting access to physical outlets by price collectors, or dramatically reducing the time price collectors are welcome within an outlet.
- Sectors of the economy in complete shut-down, meaning there are no prices to collect.

2.2 This chapter presents recommendations and emerging best practices for price collection under lockdown, drawing on the experiences from the Covid-19 pandemic. It offers a set of guiding principles rather than established rules. Countries will be impacted differently by a lockdown and NSOs will have to be flexible to implement suitable practices to ensure the CPI can be produced to an appropriate standard.

2.3 Section 2.2 discusses the organization of the price collection under lockdown. Section 2.3 discusses the communication with outlets and respondents. Alternative price collection methods are presented in section 2.4. Section 2.5 stresses the need to compare like-with-like during the lockdown and be aware of possible quality changes and additional costs passed on to consumers because of the lockdown. Section 2.6 highlights issues to consider when changing the mode of price collection. Section 2.7 looks at the impacts of possible long-term growth in households' online purchases. Section 2.8 provides guidance for returning to the regular price collection when the lockdown is lifted. Section 2.9 gives long-term guidance to make price collection more resilient. Finally, section 2.10 presents key points from the chapter.

## 2.2 Organisation of price collection

2.4 Price collection for the CPI typically relies on two basic price collection approaches: *field price collection*, where price collectors visit outlets to collect prices, and *central price collection*, where prices are collected centrally by the NSO via questionnaires, Internet, electronic means, telephone or through other data sources, such as scanner data and administrative data. Price collection for the CPI is covered in detail in chapter 5 of the *CPI Manual*.

2.5 The NSO is central to the organization and control of both field and central price collection for CPI. For the field price collection, the NSO should be in regular contact with the field force and be ready to provide updated guidance and advice to account for any issues (lockdown or otherwise) that may occur.

2.6 In the anticipation of future lockdown restrictions, the NSO should develop contingency plans that set out the processes to follow in the event of any such shocks or pandemics occurring. Contingency plans should be regularly reviewed and updated to adapt price collection to reflect the restrictions in place. The contingency plans should be flexible, well communicated with both the price collection team and users of CPI and ultimately be designed to ensure the integrity of the CPI is maintained in periods of crisis. Chapter 5 of the *CPI Manual* provides useful advice on contingency planning.

2.7 In the event of a lockdown, the NSO should quickly determine the potential impact on CPI collection. There are likely to be two main impacts for consideration – areas of the basket where products remain available for consumption, but lockdown restrictions make it difficult to collect the prices, and

areas of the basket which are unavailable as lockdown restrictions mean that sectors of the economy are no longer open to consumers. The focus of the NSO should be to collect as many prices as possible for CPI via alternative collection modes, while maintaining the safety of price collectors and adhering to government restrictions. The extent to which this price collection is successful will vary. Chapter 3 considers the implications for imputation of both reduced price collection and for those price collections which cannot take place.

2.8 The NSO should have a clear understanding and definition of what constitutes an available or unavailable product during a lockdown period, for the benefit of price collection. During a lockdown scenario (or a similar shock event) a product should be defined as *available* if it is possible for a consumer to purchase it from either outlets that remain open, or via alternative means, such as online, telephone etc. A product should be defined as *unavailable* if it is no longer possible for a consumer to purchase it from either outlets or via alternative means, such as online, telephone etc. For both available and unavailable products an appropriate imputation must be made, as detailed in chapter 3 of this guide.

2.9 For available products (whose prices can still be collected even if through different modes), the NSO should adopt all the necessary measures to collect the prices and provide data collectors with recommendations and instructions. In particular, the NSO should authorize (and publicly communicate this decision) the use of data collection techniques that are new and/or alternative to those traditionally used.

2.10 It should be remembered that field price collection for CPI necessitates human interaction. Typically, a field force of staff will visit outlets, marketplaces or service providing establishments to collect price data. At all times the health and safety of all those involved should be paramount and should never be compromised. Price collectors are particularly exposed to the risks of such a health emergency as Covid-19. They must be protected with sufficient safety measures when allowed to operate in the field and the importance of their activities in the lockdown period should be explicitly appreciated.

2.11 Online training for price collectors should be held and innovative tools to support the new modes of price collection provided, such as smartphones or PCs with Internet connection considering that part of their activities could be carried by teleworking. Data collectors should be quickly trained to use telephone to interview retailers and collecting prices from web sites of outlets they usually visit. The NSO may consider the development of standard scripts to use. Price collectors should be provided with (and trained in the use of) electronic sheets to send to retailers to collect price information via e-mails if this is an option available for price collection. Eventually, price collectors and CPI office staff may take the opportunity of their personal shopping to collect prices for the purpose of CPI, although this should be a last resort for the collection of price data.

## 2.3 Communication with outlets and respondents

2.12 Field price collection for the CPI (described in detail in the *CPI Manual*, chapter 5) is reliant on a good relationship between the NSO (or price collection field force) and the outlets where price collection takes place. Likewise, the NSO is reliant on the goodwill of respondents who may provide regular price data for use in the CPI.

2.13 A lockdown situation will quickly put pressure on the price collection process depending on the severity of restrictions being put in place within a country. It is essential that the NSO has a clear plan for communicating with both the outlets that are to be visited and the respondents who provide CPI data through survey questionnaires to ensure continuity of collection where possible.

2.14 The Covid-19 pandemic highlighted the importance of the NSO maintaining comprehensive and up to date contact information on the outlets and respondents in the CPI sample. At the onset of a lockdown, the NSO should identify those areas of the CPI basket that are likely to be impacted. To this end,

it is helpful to compile a complete list of outlets and respondents in the CPI sample that are either available for price collection, that are open but are likely to refuse price collection or that have temporarily closed. The list of outlets can be compiled based on available information updated by consulting government advice, trade and retailers' websites, price collectors (if price collection is outsourced, the third-party responsible for price collection) or other relevant sources.

2.15 The list of outlets should be reviewed frequently (even outside of lockdown situations) and where possible, supplemented with information regarding the potential availability of price data via alternative modes (such as a website, e-mail or telephone) should lockdown restrictions occur again. Hence, the NSO should continue to monitor the list with the aim to keeping it as up to date as possible. Where outlets are closed, but have advertised that they are to reopen, NSOs should aim to reintroduce them in the CPI collection at the earliest opportunity.

2.16 The list of outlets should be shared at the earliest opportunity with the price collectors to ensure an accurate list of remaining, available outlets to visit is maintained. It is likely to be the case that as price collection takes place during lockdown, the list of outlets may quickly become outdated as outlets shut at short notice, or outlets refuse price collector access due to safety restrictions or concerns. In this case, the price collector should feed this information back to the head office or the NSO.

2.17 Where an outlet has closed or refuses entry, the price collector should not attempt to visit an alternative outlet to collect prices. It is good practice for price collectors not to switch between outlets in the field, in the event of a sampled outlet being closed or refusing the price collector access. The prices for the outlet should be treated as temporarily missing until such time that access is resumed.

2.18 Despite an outlet being open, the NSO should be mindful that local price collection may not be welcomed during periods of lockdown or where strict social distancing guidelines are in place. The NSO should already have the location and contact details for each outlet in the CPI sample and should aim to confirm the availability of the outlet for price collection by making contact via telephone or email. Where possible, the NSO should attempt to agree a collection period that minimizes the impact on the outlet. Reassurances should also be provided to the outlet that a price collector will take full safety precautions, such as personal protective equipment, as necessary. The NSO may consider being more flexible with the timing of price collection than is usually the case, such as extending the collection period by a couple of days to ensure enough time to visit outlets at quieter periods.

2.19 When contacting respondents, the NSO should also attempt to collect auxiliary information to support planning for future possible lockdowns. This may include details regarding the availability of prices for that outlet via alternative means (e-mail, website or telephone) and if the outlet is closed, when the outlet will likely reopen.

2.20 The NSO should remain considerate of any additional burden the price collection may have on outlets both in terms of physical price collection during a period of lockdown or asking for the supply of price data via alternative methods.

2.21 Despite the reopening, outlets may still be reluctant to welcome price collectors for a period. The NSO should be respectful of this and arrange to collect prices at a convenient time or by alternative modes (if available) until normality is resumed. If no collection of prices is possible, despite the outlet being open then the NSO should continue to utilize the most suitable imputation approach (see chapter 3).

2.22 The NSO should be clear and transparent regarding the impact of lockdown (or similar events) on the quality of the CPI. As a minimum, alongside the published CPI, the NSO should identify all products in the CPI basket that were unavailable for collection or where alternative modes of price collection were used because of lockdown restrictions. Chapter 5 of this guide provides more detailed guidance on the communication of such issues.

2.23 The clear and early communication by the NSO of both the impact of a lockdown on CPI collection and the procedures taken to mitigate against this is not only essential to ensure transparency for users of CPI data, but also for those who are involved in the collection process, such as field force staff/central collection staff and outlets, where prices are collected. The NSO should aim to provide clear metrics alongside the published CPI to detail the impact of any lockdown restrictions on the amount of prices collected, and the mode of collection used.

### Box 2.1 Data collection in the UK and in Italy

---

In response to the Covid-19 pandemic and the emerging lockdown restrictions in the United Kingdom (UK), the Office for National Statistics (ONS) set out the implications of lockdown restrictions on the UK CPI and the plans to mitigate against this in an article<sup>1)</sup>, published as soon as possible on the ONS website. This provided users with a transparent and clear understanding of the approach ONS were taking to mitigate against the impact of Covid-19 on CPI price collection. Likewise, a follow-up article was published by the ONS to set out the clear steps ONS were taking to resume the local collection<sup>1</sup> of price data when the lockdown restrictions are relaxed.

In Italy, as the pandemic spread, the national statistical office (Istat) established a dedicated website (<https://www.istat.it/it/archivio/239854>) informing users and respondents about measures adopted to deal with the consequences of the pandemic on statistical activities. On this website, in April 2020, users and respondents were informed of changes in data collection methods because of the lockdown. Field price collection was replaced by Computer Assisted Telephone Interviewing (CATI) and online price collection.

1) <https://www.ons.gov.uk/economy/inflationandpriceindices/articles/coronavirusandtheeffectsonukprices/2020-05-06>

---

## 2.4 Alternative modes of price collection

2.24 The Covid-19 pandemic has highlighted the increasing potential for alternative data sources to support the production of CPI. More frequently, alternative data sources and price collection methods are being introduced for CPI such as the use of electronic point of sale (scanner) data or web scraping prices directly from outlets with online presence. Not only can alternative data sources provide NSOs with price data to supplement areas of the CPI basket where collection may be restricted by lockdown, it can also provide real-time intelligence on the impact of such restrictions and allow the NSO to plan for any potential shocks in the future collection of data for CPI.

2.25 The Covid-19 lockdown measures pushed NSOs to quickly use alternative sources and techniques to collect prices of products to overcome the restrictions to physical data collection. In some cases, alternative sources and collection techniques were already in use, in other cases the NSO had to quickly adapt and use these new modes of collection for the first time.

2.26 The NSO in all cases should have a clear definition for what is deemed an available product for price collection and should make every attempt to adapt data collection techniques to circumvent the need for a physical collection in the most severe restrictions. This may necessitate the move towards a multi-method collection during lockdown periods, utilizing such methods as collection over the internet, telephone or scanner data.

2.27 The main alternatives to traditional price collection in terms of sources and techniques available are collection via outlet websites (also queried through web scraping or crawling techniques), telephone calls, e-mails, newspapers advertisements, in person price collection, scanner data and other 'big data'

sources. Most of these other sources and methods to collect prices are discussed in chapter 5 of the *CPI manual* (5.175 – 5.206).

### Box 2.2 CPI data sources and collection modes

---

- Field price collection
  - Outlets reporting through survey questionnaires (electronic questionnaire they send by e-mail or an online questionnaire on the NSO data reporting portal)
  - Newspapers advertisements, catalogues etc.
  - e-mail
  - Telephone
  - Outlets' websites
  - Scanner data
  - Bulk download of data from the web
  - Administrative data
  - In person price collection (as a last resort and only for a limited number of items)
- 

#### Outlet websites

2.28 Under lockdown prices can be collected from outlet websites with two purposes: 1) To replace field price collection when price collectors cannot visit physical outlets but the product is available for sale, and 2) To collect prices of retailers that are allowed to offer exclusively online their products in response to the lockdown.

2.29 In the first case (replacement of the physical price collection but where the products are available), prices collected online should, as far as possible, represent the price offered in the physical outlet. A phone contact with the outlet to check this is recommended. One aspect to be aware of is that online prices may refer to a national price and this may miss out any regional variation that is shown in physical outlets across the country. The NSO should ensure such cases are investigated.

2.30 In the second case (collecting prices of retailers only offer their products online), the recorded prices should represent the transaction prices that would be proposed in the physical outlet. In this case, concerning the online collection of prices for goods, possible charges relating to a delivery service should be considered in the price collected (see section 2.5).

2.31 If appropriate and the NSO has sufficient capability, websites could be queried by automatic procedures to scrape or 'crawl' the data (usually referred to as web scraping). Where this is possible as an option to collect prices, the NSO should be aware of the outlet website terms and conditions regarding automated web scraping. To avoid activation of robot exclusion protocols, NSOs should inform and ask permission from outlets and agree to the most suitable web scraping technique.

#### Telephone, newspapers advertisements

2.32 Making recourse to telephone, e-mails, newspapers advertisements, in person price collection (where this means collecting prices when the staff of the NSO or field collection team visit an outlet for personal shopping) is mainly aimed at replacing price collectors who are temporarily not able to visit physical shops. While these modes of collection provide the NSO with the opportunity to fulfil the collection of prices in time of a lockdown, the NSO needs to ensure the prices collected reflect the products being priced pre-lockdown. Hence, the NSO may need to invest additional resources in the quality assurance of prices collected via these modes.

2.33 The main disadvantages of price collection by telephone, e-mails and newspaper advertising mainly consist of the difficulties in assessing the continuity of the price collection in terms of the characteristics of the product offer. In the case of telephone and e-mail this issue could be partially overcome by interviewing the retailer asking carefully the price of each specific product included in the sample, (but it should be noted this is burdensome both on the outlet and NSO) or sending a questionnaire where the characteristics of the product offer are specified. The use of newspaper advertising does not allow interaction with the respondent and should only be used when the information provided is sufficient to identify the individual product offers.

### Scanner data

2.34 Scanner data are discussed in detail in chapter 10 of the *CPI manual*. Utilizing scanner data provides the NSO with a solution to deal with the issues arising from lockdown restrictions but it is not easy to implement quickly, is reliant on developed expertise and technology and such data are not immediately suitable for use in CPI compilation. If the NSO is already developing the systems and processes for collection of scanner data, then the lockdown period could represent the opportunity to speed up the process of their integration in the regular CPI compilation. If scanner data are already used in the CPI, the possibility to expand their coverage to other groups of products is another potential solution to deal with the issues of data collection in lockdown period.

2.35 Other 'big data' sources represent alternatives to traditional CPI data collection techniques during periods of lockdown. The use of third-party web scraping tools to implement bulk download of prices from the web may provide an opportunity to maintain the price data for CPI. In this context the cooperation with such 'big data' providers could be considered, but may come at additional cost and would need to be covered by guarantees regarding the transparency of collection and the compliance of the methodological choices adopted with the statistical requirements of CPI production.

### In person price collection

2.36 A 'last resort' option to mitigate problems of data collection under lockdown, if the above mentioned are not feasible, is price collection during personal shopping. This collection could be carried out by staff from the NSO or the price collection team. In this case, the staff member could visit an outlet already in the sample and take the opportunity of his/her personal shopping to collect prices of items selected. This solution is feasible only if prices of a few items have to be collected in specific outlets and the NSO should be mindful of the need to quality assure those prices collected by staff who are not familiar with price collection for the CPI. It should be stressed that such instances of price collection should only be carried out as a last resort.

## 2.5 Maintaining the principles of CPI price recording

2.37 Regardless of the mode of price collection, in response to a lockdown, the prices of goods and services are likely to be impacted by some of the measures introduced to adhere to government restrictions or health and safety requirements. Likewise, outlets and service providers may need to adapt their products to ensure they are still available for consumption during the lockdown. Examples may include the need to provide personal protective equipment, catering services being provided as takeaway only, or delivery charges now being added to the cost of a good that can no longer be purchased in person.

2.38 Every effort should be made to collect the same quality of good or service although the NSO should be aware that there will be difficulties presented given the issues mentioned. As a result, there may need to be a slight loosening of the collection rules regarding comparable and non-comparable products depending on the level of restrictions and the impact of the lockdown. The availability of resources and

time will dictate how each NSO can handle such issues. The following is intended to provide a general sense of how specific issues may be treated for the collection of prices under lockdown.

2.39 At this stage, it is worth revisiting the underlying principle of a fixed basket price index. The *CPI Manual* explains the importance of the ‘matched-model methods’ (MMM) in chapter 6. The measurement of changes in consumer prices is complicated by the appearance and disappearance of new and old products, as well as changes in the quality of existing ones. A period of lockdown, particularly in response to a pandemic such as Covid-19, is likely to put strain on the ‘MMM’ approach as the economy reacts to the impact of lockdown and restrictions. The *CPI Manual* (6.6) states:

“The NSO typically go to great lengths to ensure measured price changes are not influenced by changes in the quality of items. By measuring the price change of a fixed, constant quality basket of goods and services, NSOs use the MMM.”

2.40 This means the price collector should attempt to price the exact same good or service each period to meet the requirement of comparing like with like. Clearly, a lockdown will present challenges. Some of the changes in products during and post lockdown will be subtle, while other changes will clearly be changes in quality.

2.41 The NSO will need to provide regular support and guidance for the price collectors, aiming to ensure that the quality of a product, based on the discernible price determining characteristics, should be as close as possible in the pre, during and post lockdown collection periods. Any doubt regarding comparable or non-comparable products should be raised immediately with the NSO, so the lockdown could lead to an increased number of collection queries for which the NSO should be prepared.

### Extra charges and delivery charges

2.42 Prices recorded for the CPI should be purchasers’ prices. The purchaser’s price is the amount actually paid by the consumer. It includes value added tax (VAT) and any other indirect taxes paid by the purchaser. It excludes taxes deductible by the purchaser and subsidies.

2.43 In some cases, the consumer has no choice but to pay an extra charge in addition to the advertised price to acquire the product. In cases where most customers will pay the extra charge, the charge should be added to the advertised price to determine the price for the CPI, provided the underlying product is deemed as sufficiently comparable to the one supplied without extra charges before the lockdown.

2.44 In the event where extra charges are passed onto the consumer to account for, e.g., pandemic related extras like face masks or sanitization requirements, these charges should be included in the price and recorded as a price increase where it is not possible to obtain the product without the additional charge. From a conceptual perspective, increases in production costs that are passed onto the consumer should lead to increases in the CPI. However, if the lockdown related extra charges are optional, the charges should be omitted from the collected price. Box 2.3 provides an example of the treatment of optional and non-optional extra charges.

2.45 The delivery of a product may include a *delivery charge*, whether the product has been purchased by visiting the outlet, over the phone or online. For CPI compilation, charges that are directly connected to the delivery of the product and not separately invoiced, should be included in the price. If the delivery charge is separately invoiced or relates to the purchase of several items, it should not be included in the price for the CPI. Instead, it should be included under transport services. The *CPI manual* provides more details on the treatment of extra charges and delivery charges (see paragraphs 5.18-5.19, 5.196, 11.57 and 11.78- 11.79).

## Box 2.3 Treatment of extra charges

---

### An optional extra charge

A gentlemen's hair cut prior to a pandemic cost \$15.00. This is a basic dry cut, no further styling. In response to the pandemic, government guidance states, if asked by the consumer, the barber should provide protective face masks and hand sanitization for all customers. The cost of the face mask and hand sanitizer is \$4.00, so if the consumer includes these optional extras the price increases to \$19.00. The price collector should collect the price without the optional extras, so the service is comparable with the pre pandemic service

### A non-optional extra charge

A gentlemen's hair cut prior to a pandemic cost \$15.00. This is a basic dry cut, no further styling. In response to the pandemic, government guidance states barbers are required to provide protective face masks and hand sanitization for all customers and provide additional sanitization of equipment before each cut. These additional safety requirements increase the production cost for the barber by \$2.00 per customer, so the price of a basic cut increases to \$17.00. The price increase should be taken into account for the CPI collection.

---

2.46 In response to lockdown restrictions, the government may intervene to manage the setting of price levels. For example, a period of lockdown restrictions may lead to temporary (or longer-term) loss of employment. As a result, the government may introduce temporary rent controls to limit the increase in price of private rents. Similarly, the lockdown may lead to panic-buying by consumers wishing to stockpile high demand products, such as paracetamol or sanitization goods, leading to price control or introduction of products by government at nominal prices. In these cases, the NSO should continue to collect and include these products in the CPI on the assumption that the recorded price reflects the amount actually paid by households for consumption. The NSO should closely monitor these situations and ensure any updated price development is captured once any interventions have ceased.

### Restaurants and cafes

2.47 The collection of price data for restaurants and cafes is likely to be impacted during a lockdown. Depending on the severity of restrictions, such establishments may only offer limited 'dine in' services or offer a reduced menu for takeaway or delivery. The NSO will need to distinguish if the outlet is open for business before making any further decision regarding what price data can be collected and via what mode (field collection or via alternative data sources). As a guide:

- If a restaurant/cafe is not open for consumers (either for 'dine in' or takeaway) the sample for that outlet should be treated as temporarily unavailable, even though the price list may be available online (see chapter 3 for imputation options).
- If a restaurant/cafe is open for consumers, but additional costs are now added to the price of a meal to cover any pandemic related measures, such as face masks, health and safety requirements etc. then these additional costs should be included in the price collected.
- If a restaurant/cafe is offering takeaway only, the price collector will need to determine if the available menu is comparable to the pre-lockdown menu. If the lockdown menu is deemed comparable, the price should be collected and used in the CPI. If the lockdown menu is considered non-comparable then standard replacement and quality adjustment methods can be applied (see chapter 6 of the *CPI Manual*).

- In some cases, where the restaurant/cafe is providing a takeaway/delivery menu only, the price will include an additional charge for the delivery. If delivery costs are clearly displayed and available, they should be *excluded* from the recorded price. If delivery charges are not separately identifiable, the full price should be collected and included in the CPI.

### Box 2.4 Charge for restaurants' delivery of a menu

---

In response to lockdown restrictions, a restaurant closes its 'dine in' facility but offers a limited delivery menu. The products on the limited menu are comparable to the pre-lockdown products that were collected. Before the lockdown restrictions were in place, a dine in meal of cheeseburger and chips was priced at \$18. This same meal is now available for delivery. The menu price is now \$19 plus an additional \$3 delivery charge. The price collector records the price as \$19, ignoring the additional delivery costs.

---

### Travels, package holidays and accommodation services

2.48 The travel and transport sector, package holidays and accommodation services will present many difficulties for the collection of suitable prices during a period of lockdown. In the most severe restrictions, many services may be suspended and therefore unavailable for consumption. A general rule to follow in this instance, where services are not available for consumption is to mark the prices as unavailable and impute the price using the methods outlined in chapter 3. Section 2.8 provides more detail about the treatment of travels, package holidays and accommodation services.

## 2.6 Issues to consider when changing mode of price collection

2.49 Taking into account the aforementioned alternative data sources and alternative collection options which could realistically be implemented in a period of lockdown, it is crucial (and potentially complicated) that the NSO understands the limitations and issues in running this mix of data sources, retail trade channels and data collection techniques. The main ones concern the quality differences that may exist, the replacements between product offers, the sample representativeness and the weighting schemes connected with the progressive change in the survey design.

2.50 Some of the questions that need to be considered are:

- Eating a meal at home that is delivered by a restaurant - is it the same as eating a meal in the restaurant, with the table service etc.?
- Following a gym class online at home, is it the same as physically attending the gym (where people have use of changing rooms and showers etc.)?
- Choosing an online product - is it the same as choosing a product in the outlet where the consumer could ask for assistance and clarification?
- If a good is delivered to your home, is the price the consumer pays comparable with the price he/she will pay in the physical shop where he/she must go to buy that specific good?
- Are prices compiled using scanner data for a group of products comparable with prices collected via traditional data collection for the same group of products?

2.51 Shifting data source or collection mode raises issues of comparability of the prices of the collected products. The NSO needs to pay particular attention when the new data source or data collection modes means a potential change in the quality of the product previously selected in the sample. Broadly speaking, if the new data source or collection mode is implemented to replace field price collection, a quality change

should not arise, if the NSO are sure that the product offer is the same as it was pre lockdown. When the new data source or collection mode imply a change in the distributional channel (from the physical outlet to the web outlet) a quality change may emerge and if possible and appropriate the NSO should adjust the collected price to take account of this change.

2.52 The change in the selected distributional channel is evident when the sample unit has to be replaced if it is not able to adapt to the lockdown restrictions (for instance a clothing or footwear outlet that is not able to move to online trade) and it closes without a perspective of reopening. The replacement of this sample unit with a new one in general produces an issue of quality change unless the new sample unit belongs to same category of outlets and sells the same variety of products. When replacing a physical outlet with an online outlet the NSO should adjust for quality changes. Even if the same product offer is available, direct comparison cannot be carried out and at least an implicit quality adjustment of the price should be carried out.

2.53 The sample is also affected in terms of time, which is one of the crucial dimensions of sampling when measuring inflation. Moving from the traditional data collection (that in general means collecting prices once or twice a month) to a new data source such as scanner data or bulk download of online prices, does not allow a direct comparison of prices for sampling reasons. Also, in this case at least some implicit adjustment should be carried out to compare the new price derived from the new data source with the previous one collected by traditional methods.

2.54 In some cases, it may be necessary to update the sample to keep it representative. For example, if online sale for a group of products was disregarded before the lockdown period as it was negligible, it may be necessary to consider it in the sample by replacing or complementing the physical shops with the new channel. For instance, if the sale of appliances on the web increases, the sample may need to be updated with price observations from the web to remain representative.

2.55 Changes of the sample design should produce a revision of the sample weights associated to the sample units and to different retail trade channels. If it is known that the replacing outlet has a turnover clearly higher than the previous one (because similar outlets were not available), it should have a higher sampling weight. If the prices that are available online start to be considered, this channel must be considered with its own weight in the compilation of indices at the elementary level.

2.56 The sample product varieties should be kept as representative as possible. If new data sources or new data collection modes have been introduced during a lockdown, the NSO may take the opportunity to consider implementing them for the CPI compilation permanently. However, the NSO should be mindful of the wider difficulties and implications of a lockdown and remain flexible. Under lockdown, the conceptual ideal for a CPI may be difficult to achieve so the NSO should attempt to make the best use of the available resources.

## **2.7 A permanent level shift in households' purchases on Internet**

2.57 Household may purchase goods and services over Internet where the goods or services are then delivered online or physically. It includes both domestic and international, cross-border transactions. The potential savings from purchases over Internet are substantial in terms of reduction in travel and search costs. Because of the potential cost reductions, purchases over Internet may be expected to continue growing, facilitated by digitalization and the availability of online tools and apps to select and purchase goods and services. In many countries, the movement towards Internet purchases was accelerated by the lockdown under the Covid-19 pandemic. While purchases over Internet to some extent may revert to pre-lockdown level it can still be expected to increase in the longer term reflecting more permanent changes in households' consumption behaviour.

2.58 Under lockdown households may not be able to purchase products in physical outlets (or they can purchase them with restrictions in terms of opening times or with difficulties due to social distancing requirements, such as long queues) and therefore make wider recourse to online purchases. This is particularly true for technological products, but also for clothing, footwear and for meals, food and beverages offered by restaurants, pizzeria, cafes, and the like. Moreover, purchases over Internet becomes increasingly more popular as it adapts to the increased demand such as for grocery products, and, for an example, allow to arrange online delivery for extended periods.

2.59 The expected structural changes in households' consumption behaviour should be considered in CPI sample design and thus in the organization of data collection. When the sample for the CPI is designed to incorporate changes in consumption pattern (post lockdown), the NSO should also consider tailoring the price collection to represent it. It is helpful to monitor the development in retail trade channels. When, e.g., purchases on Internet reaches a certain level, the NSO should decide if price collection from the web should be established. Hence, it is useful to monitor the development in retail trade channels to complement field price collection with alternative data sources and collection modes, including web prices, scanner data and administrative data. Prices collected from the web will give an indication of the online price development and can be used to complement to complement field price collection.

2.60 A matrix approach can be used to design the data collection, as illustrated in Table 2.1. For each product group (at whichever level of COICOP or other classification used) it may be possible to estimate the distribution of total turnover on retail trade channels based on available sources. The distribution of turnover on trade channels can then be used to design the sample and data collection. A redesign may have consequences in terms of the weighting of prices collected from different channels, which the CPI compiler must consider.

**Table 2.1 Turnover by retail trade channel and COICOP divisions (percentage)**

COICOP EXPENDITURE DIVISION	Large scale retail trade channels	Traditional distribution	Service provider establishment or companies	Internet purchase	Others	Total
01 Food and non-alcoholic beverages	60	20	0	20	0	100
02 Alcoholic beverages, tobacco	50	30	0	15	5	100
03 Clothing and footwear	30	40		30	0	100
04 Housing, water, electricity, gas, other fuels	10	10	50	20	10	100
05 Furnishings, household equipment and routine household maintenance	40	30	10	20	0	100
06 Health	20	30	40	10	0	100
07 Transport	10	30	30	30	0	100
08 Communication	30	20	20	30	0	100
09 Recreation and culture	20	20	30	20	10	100
10 Education	0	0	100	0	0	100
11 Restaurants and hotels	0	0	60	40	0	100
12 Miscellaneous goods and services	20	20	20	30	10	100

## 2.8 Recommencing regular price collection after lockdown

2.61 Having considered the impact on CPI under lockdown we should now consider how the NSO returns to a more regular price collection once restrictions are lifted. The NSO should prepare for coming out of a lockdown period and the implications this may have on price collection. In particular, the lockdown may impact supply chains for a longer period leading to temporary unavailability of certain products beyond the lifting of lockdown conditions. Likewise, a sustained period of lockdown may lead to the permanent closure of outlets. The NSO should have clear and well communicated advice available to deal with such situations.

2.62 The impact of lockdown on price collection for CPI may lead to some substantial short-term changes in well-established methods, such as the collection of price data via alternative means (telephone, email, online or scanner data). It is highly likely that lockdown restrictions will be gradually eased rather than removed in one go. Therefore, the NSO should plan for how best to re-introduce the pre-lockdown collection of observed prices and how this should be communicated to key users and stakeholders.

2.63 Central to any decision by the NSO to resume regular price collection will be government guidance on the lifting of lockdown restrictions, the impact of any lifting of restrictions on outlets and the availability of price collectors to resume work. The NSO should continue to keep in close contact with both sampled outlets and the field price collection staff, remembering that safety and adhering to any restrictions are top priority. As with all aspects of managing the impact of lockdown restrictions on CPI, the NSO should remain flexible and ready to adapt to emerging issues as they arise. Even the best laid out plan may need to be adapted to reflect any short-term changes in restrictions or the availability of price collectors.

2.64 A return to regular price collection will introduce several issues that potentially need to be overcome. A period of lockdown, such as the one caused by the Covid-19 pandemic may lead to a shift in the availability of products that consumers can purchase. Likewise, as lockdown restriction ease, there may be changes in the quality of goods and services when comparing with goods and services available before lockdown. Clear guidance for price collectors should be developed and regularly maintained by the NSO to ensure these issues are accurately captured.

2.65 When reintroducing regular field price collection, the change between the last pre-lockdown price and the first post-lockdown price must be correctly captured, irrespective of any temporary changes in data sources, collection mode or imputations made during the lockdown period. This means that any effects from replacement products or imputations during the lockdown will be discarded. The NSO should aim to collect updated prices for all those goods and services observed in the last pre-lockdown collection.

2.66 Because of this principle, there may be some unusual movements in the index calculations as the reintroduction of the pre-lockdown sample replaces imputations, online/alternative data source prices or observed prices with low sample sizes. This effect may be felt for a few periods, particularly where there is a phased loosening of restrictions which may delay the reinstatement of pre-lockdown price chains. Unfortunately, it is not possible to implement an approach that does not impact the CPI calculations in some way. The NSO should ensure a full description of the plan for reinstating regular field price collection and the impact of doing so on the CPI be shared with key users and stakeholders<sup>1</sup>.

2.67 In a return to regular field price collection, attempts should be made to maximize the collection of observed prices from outlets. However, there may be issues to overcome. Depending on the length of the lockdown, price collectors may find it difficult to find the products previously priced before restrictions, price collectors may need to queue for access to outlets due to social distancing and the collection process may be less efficient than in normal times. With this in mind, there should be flexibility in the amount of

---

<sup>1</sup> For example, the Office for National Statistics, UK, published an article on resuming field price collection, available on: <https://www.ons.gov.uk/releases/resumingfieldbasedpricecollectionforconsumerpricestatistics>

time allocated for price collection. NSO's may wish to allow an extended collection period to overcome such issues, but regularly review this with the goal of returning to 'normal' as soon as possible.

2.68 Lockdown restrictions may vary from one to several months or years. The longer the lockdown period, the more likely there are to be changes in the availability and quality of products that were available pre-lockdown. Guidelines for handling comparable and non-comparable replacements are provided in chapter 6 of the *CPI Manual*. However, the price collector should pay particular attention to such issues following a period of lockdown. Where possible, price collectors should make every effort to re-price the original, pre-lockdown product and only if this is unavailable should replacements be introduced. The decision on whether a replacement is comparable or not may need additional consideration following lockdown, and NSOs should provide more detailed guidance for price collectors where possible.

2.69 Following on from this, the identification and treatment of quality changes will be of increased importance for post-lockdown price collection. There may be a shift in the quality of products when comparing the pre and post lockdown collections, and the NSO should ensure there is consistency in how quality changes are treated. The *CPI Manual* provides full details of accounting for quality changes in chapter 6. Under lockdown, the following should be noted:

- There is unlikely to be a 'one rule fits all' solution for dealing with changes in quality. The NSO should be mindful of the principles that underpin the compilation of a CPI. A change in price should not be used as the basis for deciding if two products are comparable or not. The price collector should consider any differences in the product characteristics or conditions of supply under lockdown to make any decision on quality change.
- Given that the impact of lockdown restrictions may lead to quite frequent changes to the quality and availability of products that were available before lockdown, it may be increasingly difficult to identify and quantify quality changes from the perspective of the CPI. In such cases, where the change in quality cannot be explicitly accounted for, it should be implicitly considered by regular methods such as bridged overlap or class mean imputation.

2.70 Temporary extra charges during the lockdown period must be treated in a consistent manner. Such charges may be particularly apparent in personal services, such as haircuts, where costs for personal protective equipment or sanitization material is added to the price. As explained in section 2.5, for the lockdown period, additional charges should be reflected as a price increase in the CPI when the consumer must pay this additional charge to consume the good or service. When the lockdown is lifted and prices recorded without extra charges, reciprocal treatment should be implemented to ensure the temporary charges will not have long-term effects in the CPI.

2.71 As a result of the lockdown consumers may increase their purchase of products that were not previously included in the sample. For example, in response to the Covid-19 pandemic, protection and sanitization products gained importance. The NSO should consider such cases and determine if shifts in consumption of products are likely to be sustained. If the shift in consumption implies that the current sample is not representative, the NSO can account for this through the selection of replacement products or in extreme circumstances by resampling. Resampling can be carried out mid-cycle if the current set of products are no longer representative. Resampling mid-cycle should only be carried out when deemed essential. Prices for the new sample must be available for two consecutive periods to allow for the chaining of the new sample into the previous basket. Also, the new sample of products must be consistent with the other products in the existing elementary aggregate.

2.72 In some cases, a service may be gradually introduced after the lockdown. This could mean that the price collection sample is based on both observed prices and a proportion that is still missing and needs to be imputed. In such cases, if enough prices can be observed for services that are consumed, then

preference should be given to the imputation of the missing prices with the price change of the available services. The NSO should closely monitor the availability of products and how representative these are. If the number of observed prices is deemed to be insufficient by the NSO, then imputation methods, as detailed in chapter 3 should be used.

2.73 An unfortunate implication of a period of sustained lockdown is that some outlets will be forced to permanently cease trading. As a result, when price collection resumes the outlet will no longer be available for price collection and this could lead to a permanent loss of sample size and quality in the CPI. Where possible, the NSO should identify store closures, either through using (amongst others) reports in the media, trade websites or through filed price collectors in advance, and follow the usual processes for sample replacement as described in chapter 4 of the *CPI Manual*.

## Travels, package holidays and accommodations service

2.74 After a period of lockdown, there are likely to be specific issues with recommencing price collection for services such as air travel, package holidays and accommodation services. In July 2020, Eurostat issued guidance related to the resumption of price collection in these areas (Eurostat 2020b). It notes that as prices become available, the market situation is likely to be different from the one before the onset of lockdown restrictions. The guidance can be summarised as follows:

2.75 Air travel is likely to be one such service where the reintroduction of regular collection may cause issues for the NSO. As a general rule, observed prices should be included in the CPI in the month in which the service resumes. For services such as air travel, prices are typically collected in advance (such as 1-month, 3-months and 6-months). Where prices have been collected in advance, for example before the onset of lockdown restrictions, they should be included as long as the actual service is supplied.

2.76 If the price collection scheme for services such as air travel is considered to no longer be representative of current (post lockdown) consumption pattern, and if the changes are considered to persist, the price collection should be reviewed by the NSO. Any changes in the price collection, such as collection one month in advance instead of three or if changes in travel dates are associated with changes in quality then, in line with CPI principles, they should be subject to quality adjustments.

2.77 Where such changes have occurred, the ideal solution would be to explicitly value the quality difference between the old and new products to adjust the price. However, this is likely to be difficult. Direct comparison should be used when differences between the old and new products are considered not to be substantial or difficult to quantify. In each case, the NSO should consider the most suitable approach. However, quantifying the change in quality attached to such changes in booking specifications are likely to be difficult and as such, these changes may have to be ignored by NSOs.

## 2.9 Making price collection more resilient in the long-term

2.78 The lockdown following the Covid-19 pandemic demonstrated that relying on only one or few data sources or collection modes exposes the CPI compilation to risks when these are no longer available or not available to the same extent. This was the case for field price collection that in many instances could not be carried out. NSOs' efforts to investigate and utilise new data sources and collection modes was crucial to ensure enough prices could be collected for the compilation of the CPI. Without the use of alternative data sources and collection modes, the compilation of the CPI would have been seriously hampered and additional imputations would have been required. However, experiences also showed that changing data sources and collection mode require carefully planning and enough time since it involves changing operational and administrative procedures and data processing steps.

2.79 To develop more resilient production systems, CPI compilation may have to be increasingly based on a multi-source and multi-mode approach to data collection. This involves investigating and utilising a range of available data sources and methods in which to collect data for the CPI, including field price collection, internet portals for respondents, scanner data, Internet (manual collection from websites and targeted and bulk web scraping), administrative data and other providers of relevant data (e.g. market research institutes), reducing the dependency of any single data source or collection mode.

2.80 Moving toward the use of electronic data sources, including scanner data and price collection from Internet, requires investments in suitable software, training of staff and adaption of data processing and calculation steps. Use of multiple sources and multiple data collection modes will require a redesign of the sample and weighting schemes and the way in which collected prices are aggregated into price indices. However, implementation of electronic data sources and collection modes will be crucial in establishing more resilient and efficient data collection.

## 2.10 Key points

- During a lockdown scenario (or a similar shock event) a product should be defined as *available* if it is possible for a consumer to purchase it from either outlets that remain open, or via alternative means, such as online, telephone etc. A product should be defined as *unavailable* if it is no longer possible for a consumer to purchase it from either outlets or via alternative means, such as online, telephone etc.
- Assuring the health and safety of both the field collection staff and those personnel working in the sampled outlets (where prices are collected) should be the highest priority for NSO. Regular field price collection should only resume once it is safe to do so.
- It is good practice for price collectors not to switch between outlets in the field, in the event of a sampled outlet being closed or refusing the price collector access. The prices for the outlet should be treated as temporarily missing until such time that access is resumed.
- Extra charges that are introduced because of the lockdown, for instance for personal protective measures, in general should be included in the price recorded for the CPI. However, if extra charges are clearly optional, they should be excluded from the recorded price. Delivery charges that are directly connected to the delivery of the product and not separately invoiced, should be included in the price for the CPI. After lockdown, when such charges are lifted, a reciprocal treatment must take place to ensure the long-term rate of change of the CPI over the lockdown period is not affected by temporary charges during the lockdown.
- The change between the last pre-lockdown price and the first post-lockdown price must be correctly captured in the index, independently of any price collection or imputation applied during the lockdown period. This means that replacement products or imputations made during the period of lockdown will be discarded. The NSO should aim to collect updated prices for all those goods and services observed in the last pre-lockdown collection.
- Changing data sources and data collection modes usually will involve operational and management decisions and take time. Alternative data sources and collection modes, therefore, should be considered as part of a CPI contingency plan.
- To develop more resilient and efficient production systems in the longer-term, NSOs would need to apply a multiple source and multiple collection mode approach, utilizing available data sources and different methods and techniques for the collection of data.



# Chapter 3 Imputation

## 3.1 Introduction

3.1 The lockdown that followed the 2020-2021 pandemic resulted in an abrupt surge in missing and potentially unusable input data for CPI calculations. In order to calculate and publish the CPIs, considerable strain was exerted on the established imputation methods. Could established procedures fully accommodate the sheer quantum of missing data? Could novel situations brought about by the lockdown be adequately addressed by existing, default imputation mechanisms? Were interventions or new procedures required? In effect, the lockdown called the prudence of long-standing imputation procedures into question.

3.2 This chapter provides best practices recommendations for CPI imputation during periods of large-scale market interruption. Section 3.2 outlines cases where imputation is needed and discusses definitions of imputations. Section 3.3. presents imputation methods and practical examples distinguishing between imputations for available and unavailable products. The section also provides guidance on the treatment of products with pronounced seasonal price variation and. Section 3.4 suggests a classification for deriving imputation rates. Section 3.5 discusses adjustment for change in data collection mode and Section 3.6 provides a few hints for data validation under lockdown. Section 3.7 discusses the choice of imputation methods, illustrated by examples of decision trees. Section 3.8 summarises key points of the chapter.

## 3.2 Imputation defined

3.3 By and large, national statistical offices (NSOs) traditionally have relied upon sampling and survey data to collect the information needed to compile the CPI. Over the past decade, NSOs have partially transitioned to automated capture of high volume, high frequency data (e.g., scanner data, web-scraped prices) to meet input data needs more efficiently. Neither traditional nor alternative collection methods provide the input data with perfection. Both processes result in situations where a required data element is missing (e.g., due to nonresponse) or is unusable (e.g., due to non-sampling errors). For these reasons, NSOs have developed imputation methods and techniques to ensure the calculation of each monthly CPI, while mitigating the impact any missing or unusable input data might have on total measurement error.

3.4 Imputation methods are designed to handle a variety of phenomena governing the availability of consumer goods and services in the marketplace and situations commonly afflicting CPI business operations. These include:

- Item temporarily missing
- Item permanently missing
- Seasonal item
- Collection frequency less than monthly
- Outlet temporarily closed
- Outlet permanently closed
- Outlier observation
- Erroneous (incorrect) observation
- Essential change in quality of observed item

3.5 The term ‘imputation’ can take on several different meanings:

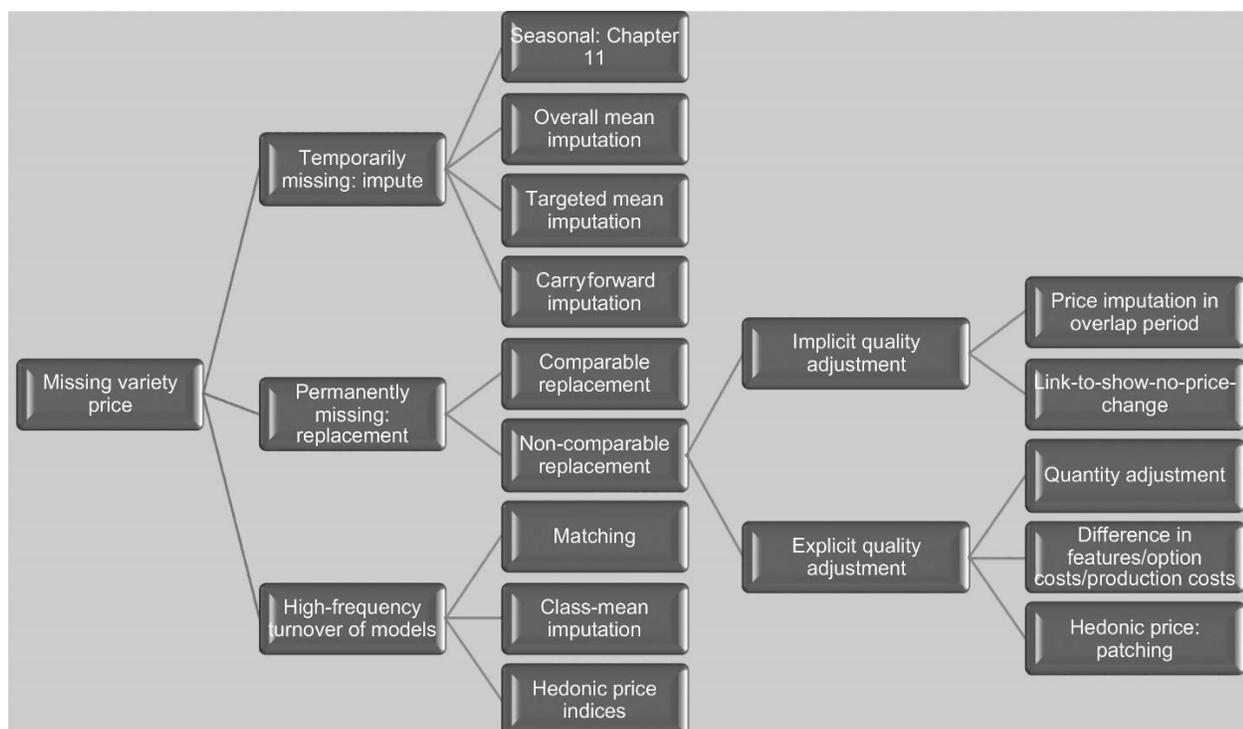
- A procedure of entering a value for a specific data item where response is missing or unusable.<sup>1</sup>
- Replacing either missing or invalid data with accepted data. Normally performed in accordance with predetermined decision rules.<sup>2</sup>
- Estimation of an implied value in a situation where an actual market transaction does not exist, e.g. imputed market rental value of an owned home.<sup>3</sup>

3.6 The *CPI Manual* provides extensive explanation of, and application guidance for, imputation methods in the context of a CPI. The manual does not provide a definition of the broad term ‘imputation’ but the manual’s Glossary of Main Terms defines the specific term ‘imputed price’ as:

“The price assigned to a variety for which the price is missing in a particular period. The term “imputed price” may also refer to the price assigned to a variety that is not sold on the market, such as a good or service produced for own consumption, including housing services produced by owner-occupiers, or one received as payment in kind or as a free transfer from a government or non-profit institution.”

3.7 Figure 6.3 of the *CPI Manual*, reproduced below, provides a comprehensive summary illustration to the treatment of missing prices for CPI calculation.

Figure 3.1 **Methods for imputation of missing prices**



<sup>1</sup> UNECE (2000)

<sup>2</sup> Statistics Canada (2011)

<sup>3</sup> System of National Accounts (2008), paragraph 3.75

3.8 The choice of imputation method turns on the objective – what is the goal of imputation in the first place? Under the match-model approach, the immediate objective is to select the best proxy price of the prior period priced variety to minimize measurement error. This could be the price of a like-kind, close-match substitute. Or it could be an estimate of the price based on the modelled price movement, usually an average, of highly similar, homogenous products. Or, perhaps the objective is not focused on estimating the best proxy price of just one sampled variety; the priority might be to minimize measurement error of an aggregate estimate. For example, one might choose to impute to the mean to minimize variance, even when the best proxy for a particular observation is a different price along the sampling distribution, e.g. the missing observation is characteristic of behaviour in the tails. Thus, the choice of proxy can vary due to the reason why the price of the matched variety is missing or unusable.

### 3.3 Imputation methods

3.9 To discuss imputations during a lockdown it is useful to distinguish between *imputations for available products* and *imputations for unavailable products*. For available products, although the product continues to be transacted, prices can be missing. For example, the specific variety included in the CPI is temporarily unavailable at the sampled outlet, or the sampled outlet is closed altogether, but the variety and similar varieties of that product continue to be sold, perhaps in other outlets. Prices can also be missing because price collection could not be carried out as usual although it continues to be possible for households to purchase the relevant products.

3.10 The situation with unavailable products arises when products are no longer available for sale and there are no transactions of the product on the market. This may happen if markets or outlets are closed, so that certain goods are not sold anymore, or services are not available for consumption. Such kind of lockdown measure may be implemented at a sub-national level, or even at the national level. Transactions that typically are part of household consumption expenditure cease to exist. Consequently, entire aggregates of the CPI basket could be missing.

#### 3.3.1 Imputation for available products

3.11 As stated above, it may happen that a price sampled for the CPI is missing and must be estimated. For example, the specific variety to be priced may be temporarily out of stock and therefore a price for that variety cannot be observed. Prices can also be missing if price collection failed, although the specific variety continues to be sold. For example, during lockdowns, statistical agencies may reduce price collection in the field, and it is not always possible to obtain prices through other means. Where a price has not been collected, an imputed price can be obtained by multiplying the previously collected price with an estimated price change. The objective consists in estimating a price change that would have been observed had the product been on offer in the outlet or had the prices data been properly collected.

3.12 It is recommended to impute missing prices by the price changes of comparable products, by the elementary price index to which the missing observation belongs, or the nearest available higher-level price index. For example, it may be possible to collect prices of the same or similar products in other outlets or other areas of the country, on which to base the imputation. If an entire elementary index is missing, it should be imputed based on the elementary indices of comparable products or the nearest available higher-level price index.

3.13 It is possible to automate imputations by formalising the rules for the imputations and integrating the rules in the calculation system. While this can save time, the compiler still needs to oversee and control the impact of the imputations. In cases where it is not obvious how a missing price should be imputed and when individual price observations carry a large weight, manual control is necessary.

### Box 3.1 Imputation for available products

Prices for available products should be imputed by:

- The price changes of comparable products
- The change of the elementary price index
- The change of a comparable elementary price index or nearest available higher-level price index
- Carry forward

3.14 Carrying forward the last observed price is easy to implement and explain to users. However, it is not neutral on the results. The more prices that are carry forward, the more the aggregate monthly price change converges to zero. Carrying forward therefore should only be used in duly justified circumstances.

3.15 The example in Table 3.1 (taken from UNECLAC 2020) illustrates imputation within an elementary aggregate. Prices for national lager beer are only observed in supermarkets. The average price change for the three first varieties is used to impute a price for the three missing varieties.

3.16 Imputing with the average change in the prices of the varieties that remain included in the elementary aggregate is equivalent to omitting the missing observation from the calculation. This has practical and variance minimizing benefits, especially for replicate-based variance estimators. This approach does decrease the statistical power of the pool of available data by decreasing the effective sample size and could result in the introduction of a bias.

Table 3.1 Example of imputing within an elementary aggregate

Item code	Level	Outlet code	Description	Specification	Price (t-1)	Price (t)	Relative (t/t-1)	Geometric mean	Variety-Outlet index	Variety index
02.13	Class		Beer							
02.13.0	Subclass		Beer							
02.13.0.1	Product		National beer							
02.13.0.1.1	Variety		National lager							101.09
02.13.0.1.1	Variety-Outlet	101-1	Supermarket 1	Bavaria, bottle	1 090	1 100	100.92		100.92	
02.13.0.1.1	Variety-Outlet	101-2	Supermarket 2	Pilsen 6.0, bottle	770	775	100.65		100.65	
02.13.0.1.1	Variety-Outlet	101-3	Supermarket 3	Imperial, 710 ml can	1 180	1 200	101.69	101.09	101.69	
02.13.0.1.1	Variety-Outlet	103-1	Off-licence 1	Pilsen 6.0, bottle	818	827			101.09	
02.13.0.1.1	Variety-Outlet	103-2	Off-licence 2	Imperial, bottle	1 190	1 203			101.09	
02.13.0.1.1	Variety-Outlet	103-3	Off-licence 3	Bavaria, bottle	1 030	1 041			101.09	
02.13.0.1.2	Variety		National dark beer							
02.13.0.2	Product		Imported beer							
02.13.0.2.1	Variety		Imported lager							
02.13.0.2.2	Variety		Imported dark beer							

3.17 The number of missing prices may be unusually high under lockdown. In practice, it may be useful to set a *threshold* for how many prices should be available before deciding on imputation methods. For instance, a threshold could be that minimum 20% of prices should be collected. If minimum 20% of prices are collected, the index can be calculated as usual. In many cases this implies that the missing prices are imputed with the average price change of the collected prices. Elementary indices for which less than, say, 20% of the prices are collected should be reviewed and their reliability assessed. If the collected prices are still found to provide a reliable estimate of the elementary index, the missing prices may be imputed by the price development of the observed prices.

3.18 However, if only few prices are collected and/or the collected price do not provide reliable information about the average price development of the elementary index to which they belong, they may be left out of the index calculation. The elementary index may then be imputed by the development of a comparable elementary index or the nearest available higher-level price index. The imputed prices and their impact on higher-level price indices may also be identified during the validation phase of the CPI production process (see section 3.6).

3.19 Suppose that a subclass (or a more detailed micro-class) is sub-divided by region or by outlet-type. It may happen that within an elementary aggregate (or stratum) all prices are missing. Another, ‘similar’, elementary aggregate could then be used as a basis for imputation. For example, missing prices in one region could be imputed with the price changes observed in a nearby region, or missing prices for a given outlet-type could be imputed with the price changes observed in another outlet-type. In practice, decisions must be made on what is meant by a ‘similar’ elementary aggregate. Such rules could be designed by studying past correlations of price changes. For example, if price changes for a product across different regions are highly correlated (because for example of national pricing strategies), it makes sense to impute missing prices with the prices data collected in other regions of the country.

3.20 If all of this fails and prices for a subclass are missing in all outlets and in all regions, an index from the next level index hierarchy should be used as a basis for imputation. This strategy is based on the idea that products and product groups can be considered the closest substitutes for the non-available products; and they are therefore next to each other in the classification structure used for the CPI. In the example in Table 3.2 from the IWGPS guidance note, a class consists of three subclasses. The third subclass is missing and is imputed using the combined price change of the first two subclasses ( $105.00 * 1.0139 = 106.46$ ).

Table 3.2 Example of imputing a missing subclass

	Weight	Dec 2019 = 100			Monthly change March-April 2020
		Dec. 2019	March 2020	April 2020	
Subclass 1	0.50%	100.00	103.00	105.00	1.94%
Subclass 2	0.70%	100.00	101.00	102.00	0.99%
<b>Subclasses 1 and 2</b>			<b>101.83</b>	<b>103.25</b>	<b>1.39%</b>
Subclass 3	1.20%	100.00	105.00	106.46	1.39%
<b>Class</b>	<b>2.40%</b>	<b>100.00</b>	<b>103.42</b>	<b>104.86</b>	<b>1.39%</b>

3.21 In general, the donor price change should come from products as similar as possible to the missing product. However, there are trade-offs. A more targeted imputation is likely to ensure a more representative price movement for the missing variety, but it may be less reliable if based on a smaller number of price observations. The average price change derived from a larger number of prices or elementary aggregates may be more reliable, but less representative for the specific variety that is missing.

### 3.3.2 Imputation for unavailable products

3.22 The imputation for unavailable products raises the question of a target price of a product that is not transacted anymore. From a theoretical perspective, Diewert and Fox (2020) argue the prices of unavailable products should be estimated by the ‘reservation price’ that would lead to zero demand. However, the measurement of reservation prices is not straightforward. Reservation prices are typically not implemented as part of the regular CPI production routines.

## Box 3.2 Imputation for unavailable products

---

Prices for unavailable products should be imputed by:

- The change of a comparable elementary price index
  - The change of the nearest available higher-level price index
  - The change of the overall CPI
  - Carry forward
- 

3.23 Where an entire aggregate of the CPI is missing, it is important that the imputation method is implemented in a transparent manner. In general, the imputation methods are based on a selection of available, observed, price changes. This type of procedure ensures that the imputations remain consistent with the observed prices data. For example, suppose that the prices for the missing elementary aggregate A are imputed with the average price change of the elementary aggregate B. The combined price change of elementary aggregates A and B is then equivalent to omitting the missing elementary aggregate A from the calculation. The recommendation of the IWGPS is to use one of the following imputation methods.

### Impute with a comparable elementary price index.

3.24 A similar elementary aggregate can be selected, and the price change observed for this be used as a basis for imputation. This method relies on the assumption that the expenditure has been substituted to that similar product group. For example, the price change of drinks in restaurants could be measured with the price change of drinks bought in supermarkets. This procedure ensures that the combined price index of drinks in restaurants and drinks in supermarkets is consistent with a price index without imputations and only consisting of drinks sold in supermarkets. In the absence of relevant expenditure data, judgements must be made in selecting a comparable index.

### Impute with the nearest available higher-level price index

3.25 The CPI baskets are based on hierarchical product classifications, such as COICOP. If an index is missing (for example a subclass index), one could use the next level index in the classification hierarchy (for example the class index) as the basis for imputation, as illustrated in Table 3.2.

### Impute with the overall CPI

3.26 Imputation by the overall CPI corresponds to leaving the missing product out of the CPI calculation. This method can be justified in cases where markets have been closed and where there are zero or only negligible transactions.

3.27 Let  $I_i^{t-1,t}$  denote the month-on-month index of elementary aggregate  $i$  to which a weight  $w_i^{t-1}$  is attached.<sup>1</sup> We suppose that the weights over all elementary aggregates sum to unity. The month-on-month overall index is derived as the weighted average of the month-on-month indices for the elementary aggregates:

$$I^{t-1,t} = \sum_{i \in N} w_i^{t-1} I_i^{t-1,t} \quad (3.1)$$

---

<sup>1</sup> The weights correspond to the weights used in the price reference period that are price-updated to period t-1 and normalized:  $w_i^{t-1} = \frac{w_i^0 I_i^{0,t-1}}{\sum_{i \in N} w_i^0 I_i^{0,t-1}} = \frac{w_i^0 I_i^{0,t-1}}{I^{0,t-1}}$

3.28 Let us suppose that the set of elementary aggregates  $N$  can be split into ‘available’ elementary aggregates  $A$  and ‘missing’ elementary aggregates  $M$ . Within a missing elementary aggregate, all the prices must be imputed. We denote by  $\hat{p}_j^t$  the imputed price of a product  $j$  in period  $t$  and by  $p_j^{t-1}$  the observed price of the same product  $j$  in the previous period  $t-1$  in such a missing elementary aggregate. We then have the following imputation rule for the missing prices based on the average price change of all available elementary aggregates:

$$\hat{p}_j^t = p_j^{t-1} \frac{\sum_{i \in A} w_i^{t-1} I_i^{t-1,t}}{\sum_{i \in A} w_i^{t-1}} \quad (3.2)$$

3.29 Assume that the formula for the missing elementary index is such that if all prices in period  $t-1$  are multiplied with the same factor to impute period  $t$  prices, the index changes by that factor between  $t-1$  and  $t$ . It follows that for all missing elementary aggregates  $i \in M$  we must have:

$$I_i^{t-1,t} = \frac{\sum_{i \in A} w_i^{t-1} I_i^{t-1,t}}{\sum_{i \in A} w_i^{t-1}} \quad (3.3)$$

Then it can be shown that:

$$I^{t-1,t} = \sum_{i \in N} w_i^{t-1} I_i^{t-1,t} = \sum_{i \in A} w_i^{t-1} I_i^{t-1,t} + \sum_{i \in M} w_i^{t-1} I_i^{t-1,t} = \frac{\sum_{i \in A} w_i^{t-1} I_i^{t-1,t}}{\sum_{i \in A} w_i^{t-1}} \quad (3.4)$$

3.30 This result shows that with this specific imputation method, the index boils down to an index without any imputations. In other words, imputing with all available sub-indices corresponds to an implicit reweighting, where the weight of the missing sub-indices is set to zero and the weight of the available sub-indices is rescaled to sum to unity. This implied weighting structure can still differ a lot from the actually observed expenditure pattern during the lockdown periods.<sup>1</sup>

3.31 The imputation with the overall CPI may be the preferred solution. There can be various implementations of this method. For example, only reliable sub-indices, i.e. sub-indices that are not subject to a significant imputation, could enter the calculations of the imputed price change. The regional dimension may also play a role. For example, the imputation could be based on the all-items of a specific region, or on the all-items of the entire country. In reality, the implementation and interpretation of this method is further complicated by the fact that the set of ‘available’ and ‘missing’ sub-indices is changing over time. Table 3.3 shows some examples of products (in this case services) that were imputed with respect to the all-items during lockdowns in some countries.

**Table 3.3 Country examples of products imputed with the overall CPI**

Austria (May 2020)	Spain (May 2020)	Australia (June quarter 2020)
<ul style="list-style-type: none"> <li>Recreational and sporting services – Attendance</li> <li>Cinemas, theatres, concerts</li> <li>Museums, libraries, zoological gardens</li> <li>Secondary education</li> <li>Restaurants, cafés and dancing establishments</li> <li>Canteens</li> </ul>	<ul style="list-style-type: none"> <li>Driving lessons, tests, licences and road worthiness tests</li> <li>Passenger transport by sea</li> <li>Recreational and sporting services - Attendance</li> <li>Recreational and sporting services - Participation</li> <li>Cinemas, theatres, concerts</li> <li>Museums, libraries, zoological gardens</li> </ul>	<ul style="list-style-type: none"> <li>Urban transport fares</li> <li>Domestic holiday travel and accommodation</li> <li>International holiday travel and accommodation</li> <li>Sports participation</li> <li>Other recreational, sporting and cultural services</li> </ul>

<sup>1</sup> See also Diewert and Fox (2020) for the impact of ‘inflation-adjusted carry forward’ prices

### 3.3.3 Products with pronounced seasonal price variations

3.32 The prices of some products are known to follow pronounced seasonal patterns. Examples are flights, package holidays, or accommodation services. Such seasonal patterns may not be reflected using the month-on-month price change of available products. The seasonal pattern can be approximated, for example, by imputing with the annual rate of change of available products<sup>1</sup>. Another strategy to respect seasonal patterns consists in imputing with the monthly price change of the seasonal product observed one year ago. By respecting the seasonal pattern of the series, the impact of the imputations on the annual rate of change of the overall CPI is minimized. A consequence of repeating past seasonal behaviour is that the monthly price changes during the lockdown will reflect market developments seen under normal circumstances.

3.33 The impact of possible seasonal patterns of the series on imputation is best shown by the example in Table 3.4 (presented in the IWGPS guidance note).

Table 3.4 **Imputation for products with a seasonal pattern**

	2019		2020	
	March	April	March	April
All-items excl. accommodation services	100.0	100.5	102.0	103.5
Monthly rate of change				1.5%
Annual rate of change				3.0%
<b>Option 1: Imputing with the monthly rate of change of all items excl. accommodation services</b>				
Accommodation services	100.0	112.0	101.0	$101.0 \cdot 103.5 / 102.0 =$ <b>102.5</b>
Monthly rate of change				1.5%
Annual rate of change				-8.5%
<b>Option 2a: Imputing with the annual rate of change of all items excl. accommodation services</b>				
Accommodation services	100.0	112.0	101.0	$112.0 \cdot 103.5 / 100.5 =$ <b>115.3</b>
Monthly rate of change				14.2%
Annual rate of change				3.0%
<b>Option 2b: Imputing with the monthly rate of change of accommodation services one year ago</b>				
Accommodation services	100.0	112.0	101.0	$101.0 \cdot 112.0 / 100.0 =$ <b>113.1</b>
Monthly rate of change				12.0%
Annual rate of change				1.0%

3.34 In the European Union, the decision was made to adopt an imputation that preserves the integrity of the seasonal pattern (see Eurostat 2020a). This choice had significant consequences on results. An index derived solely from using a carry forward imputation mechanism would have resulted in significant lower inflation rates in the Euro Area (see Lamboray et al. 2020).

<sup>1</sup> See Appendix in Statistics Sweden (2020) for a formalization of this method

### 3.3.4 Seasonal products

3.35 To the extent possible, seasonal products should be treated as they normally are. Prices of out-of-season products should be imputed using one of the imputation methods described in chapter 11 of the *CPI Manual*. Compilers should be aware that a lockdown may change the month in which a seasonal product is re-introduced in the index. For instance, a given type of fresh fruit may usually be entered in the CPI calculation in June, but because of a lockdown it may only be possible to collect prices from, say, August. Such cases may have an impact on the CPI that differ from the usual seasonal pattern.

### 3.3.5 Self-correction

3.36 Under normal circumstances, prices for temporally missing varieties are imputed only during a limited period (for example 2 to 6 months) before they are substituted with a replacement variety. Under lockdown, a more flexible approach is warranted regarding the length of imputations. Prices of products may be imputed for much longer. Moreover, with a continuous tightening and relaxation of restrictions, imputation patterns can become quite irregular and complex. For example, after a first lockdown period, prices may become available for some time before being unavailable again thereafter.

3.37 However, irrespective of the length of the imputation period, an important principle is that the imputation of a given sampled item should be ‘self-correcting’. This means that the long-term change from the last observed price before the lockdown to the first observed price after the lockdown must be correctly reflected in the index, irrespective of how prices in the interim are imputed. Self-correction ensures that the index reaches the correct level over the imputation period.

3.38 Box 3.3 illustrates self-correction with an example. Given the uncertainties of the imputations, this can sometimes lead to an unusual price change between the last imputed price of the lockdown period and the first observed price of the post-lockdown period. However, it may be difficult to explain such technical price changes to users.

3.39 Self-correction must also be ensured in the treatment of extra charges and fees that temporarily have been introduced during a lockdown. Suppose before the lockdown a burger was priced at a restaurant for \$10. Then, during lockdown, the price was obtained online from the restaurant’s website, for \$15 including a \$4 delivery fee. Under the match-model approach, the \$4 delivery fee would be discarded and leaving a price of \$11. If the mode change was considered comparable, then no further quality adjustment should be made and the price of \$11 recorded for the index calculation. Upon return to pricing the burger at the restaurant when the lockdown is lifted, then this mode change must be treated reciprocally as comparable without quality adjustment.

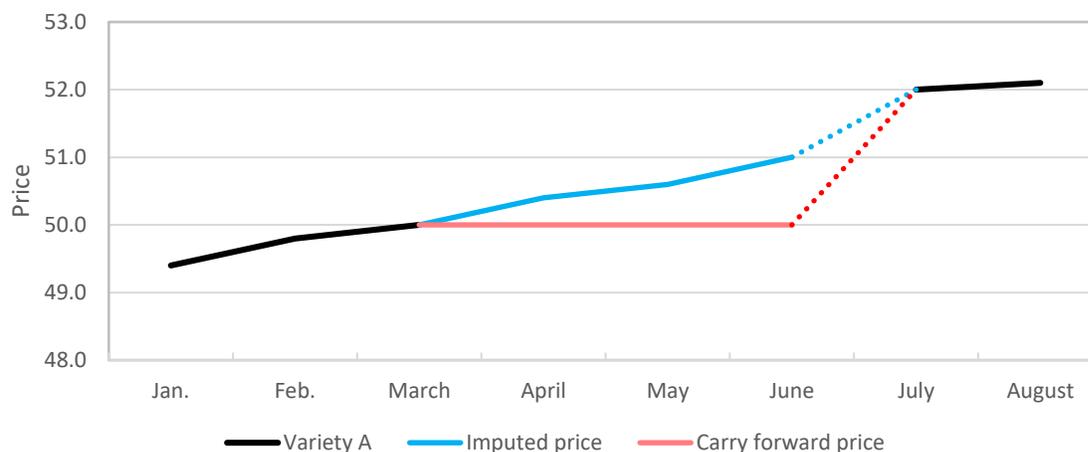
3.40 In practice, there are additional complications. For example, it may happen that prices are imputed until the end of the planned pricing cycle of a product in the index. At that moment, the imputations simply stop and there is no self-correcting adjustment. It may also happen that the selection of a variety is not possible at the planned period of initialisation and, consequently, no price can be collected. One solution consists in inserting a notional price for that variety in the missing periods based on the imputed price movements once price collection is successful<sup>1</sup>.

---

<sup>1</sup> See section 6 of ONS (2021) for a discussion on this issue in the UK.

### Box 3.3 Self-correction

Variety A disappears from the market in April. The last observed price in March is 50. It reappears in July at a price of 52. From March to July this is an increase of 4%, or an index of 104 ( $52/50 \times 100$ ).



	Jan.	Feb.	March	April	May	June	July	August
Variety A	49.4	49.8	50.0				52.0	52.1
Imputed price			50.0	50.4	50.6	51.0		
Carry forward price			50.0	50.0	50.0	50.0		

In the example the missing price is imputed with a suitable price index, which leads to an imputed price of 51 for variety A in June. The change from the last imputed price in June to the observed price in July (52), an increase of 1.96%, must be included in the calculation to ensure the index reaches the correct level in July. When all monthly price relatives are chained into a long-term index, the index will show the correct development from March to August ( $50.4/50 \times 50.6/50.4 \times 51/50.6 \times 52/51 = 52/50 = 1.04$ ).

If the price is carried forward, the same principle applies. The change from the last carry forward price in June (50) to the observed price in July (52), an increase of 4%, must be included in the calculation to ensure the correct level of the long-term price index in July ( $50/50 \times 50/50 \times 50/50 \times 52/50 = 52/50 = 1.04$ ).

## 3.4 Compiling the imputation rate

3.41 When compiling 'imputation rates' for CPI price data, a definition of what constitutes an 'imputed' price is needed. Under one approach, only imputation of temporarily missing prices are included in the calculation of the imputation rate. Under an alternative view, imputation of permanently missing items could also be included. This distinction has consequences. Under the first approach, e.g., the imputation rate for CPI in the U.S. was 12.2 percent in the six-month period from April to September 2019. In a broader sense including as imputation also all situations where the price of the previous month match-model was not collected, e.g. because of bimonthly pricing or because a substitute replacement item was selected, one arrives at an imputation rate of 40 percent. During the first part of the pandemic period (April to September 2020), the rate of imputation at the narrower definition was 24% compared to 52% at the broader definition in the U.S. Moreover, if a change in the mode of collection (e.g., from in store to internet) is also considered an imputation, then the imputation rate would leap up to 90% for the U.S. during the lockdown. Hence, the definition matters to public perception, and in adjudicating whether sufficient prices have been observed to publish aggregate indexes (see chapter 5).

3.42 Under the traditional match-model approach to index estimation, Table 3.5 itemizes eight possible different classification of a CPI sample categorized by price collection outcome that are used by the Bureau of Labor Statistics (BLS), with hypothesized sample sizes for illustration.

3.43 Response rate is typically defined as  $(a+b+c+d+e)/(a+b+c+d+e+f) = 87.5\%$ . The subset of the sample where collection is not attempted is usually excluded from the ratio, because behavioural scientists are interested in the response action to the active survey collection. The bilateral classification of the sample into response-nonresponse outcomes (87.5%-12.5%) is slightly different than the bilateral classification of the sample into observed-imputed data.

**Table 3.5 Classification of sample used for CPI calculation**

<b>Total sample targeted for calculation</b>		<b>100,000</b>
Observed (collected) prices		
a	Same item as previous month, price used	65,000
b	Same item as previous month, outlier price not used	250
c	Substitute item - comparable, price used	2,000
d	Substitute item - quality adjusted, price used	1,500
e	Substitute item - noncomparable, price not used	1,000
Unobserved (not collected) prices		
f	Collection attempted, price unobserved	10,000
g	Collection unattempted by design, off-cycle,	20,000
h	Collection unattempted by design, permanent	250

3.44 Under a very broad interpretation, the imputation rate could be defined as  $(b+c+d+e+f+g+h)/(a+b+c+d+e+f+g+h) = 35\%$ . This is not closely tethered to the nonresponse rate and could be misleading to users. A narrower definition of the imputation rate is recommended as  $(b+e+f+h)/(a+b+c+d+e+f+h) = 14.4\%$ , which is more in sync with the nonresponse rate of 12.5%. The narrower definition treats substitute items, for which a usable, if adjusted, price was obtained, as a non-impute. In the abstract, one could posit the price of the substitute item, while not precisely the price of the same unique item as required for calculation, is an exact proxy of what the price of the same item would have been, such that it equates to the targeted observed price and not an 'imputed' price.

3.45 In a similar fashion, if the price of an item is intended to be collected in-store, but a data collector is prevented from visiting the store to observe the price, and instead observes the price online for the exact same item from the exact same retail company, this can be viewed as merely a change in the mode of collection, or a trivial change in the attributes of the unique item, essentially classifying the price as a match-model price. The extension of this argument, then, is that a comparable item, although a variant item and not a genuine match-model, is by definition a perfect match for quality purposes. Hence, its price should be regarded as the observed price of the match, and not an impute.

3.46 To estimate the share of consumption expenditure that is imputed, the imputation rates for the elementary aggregates should be weighted by the weights of the elementary aggregates. This approach was followed by Eurostat<sup>1</sup> and European countries when disseminating imputation rates for higher-level price indices up to the overall index. The calculation of imputation rates is illustrated in Table 3.6.

<sup>1</sup> Imputation rates can be found at the following website <https://ec.europa.eu/eurostat/web/hicp/methodology>

Table 3.6 Calculation of imputation rates

	No. of prices (Total)	No. of imputed prices	Weight	Imputation rate by expenditure weight	Imputation rate by number of prices
Total	420	120	100%	$(60\%*13\%+40\%*92\%)/100\%=$ <b>45%</b>	$(20+100)/(300+120)=$ <b>29%</b>
Sub-index 1	300	20	60%	$(20\%*40\%+40\%*0\%)/60\%=$ <b>13%</b>	$(20+0)/300=$ <b>7%</b>
Elem. aggr. 1.1	50	20	20%	40%	$20/50=40\%$
Elem. aggr. 1.2	250	0	40%	0%	$0/250=0\%$
Sub-index 2	120	100	40%	$(15\%*80\%+25\%*100\%)/40\%=$ <b>92%</b>	$(80+20)/120=$ <b>83%</b>
Elem. aggr. 2.1	100	80	15%	80%	$80/100=80\%$
Elem. aggr. 2.2	20	20	25%	100%	$20/20=100\%$

### 3.5 Change in data collection mode

3.47 Alternative collection modes can be used in case of failure of the regular price collection in the field. For example, prices previously collected in outlets are now obtained from the web, over the phone, or from transaction data that may be available to the statistical agency.

3.48 There are different ways to interpret and treat such a change in data collection mode.

#### Directly compare the prices of the different collection modes

3.49 This approach relies on the assumption that the price collected under the new collection mode is a reasonable estimate of the price that would have been observed in the same period under the previous collection mode. The way that the product is supplied could differ. There could be, for example, shipping costs or take-out fees that must be deducted from the price observed with the new mode. Alternatively, one could also argue that some of these extra charges should be taken into account and recorded as a price increase on the grounds that it is not possible to obtain the product without the payment of the additional charge.

#### Make a quality adjustment between the prices of the different collection modes

3.50 One could argue that the product sold in an outlet is of different quality than a product sold on the web. In addition to possible changes in the specifications of the product that is priced, changes in the way that the product is supplied contributes to changes in the quality of the product. There is a need to decide if price differences observed from the different collection modes should be treated as genuine price changes or as a change in quality. Explicit quality adjustment methods remain difficult to apply in such circumstances. Therefore, implicit quality adjustments methods are applied in practice.

#### Resampling

3.51 During a lockdown there may be a shift towards online purchases. Products typically not covered by CPI (e.g. face masks and disinfectant) suddenly gain importance. New outlets (for example web stores) and new products (either in already selected or in newly selected outlets) could be introduced into the index through re-sampling. Ideally, sample rotation could be implemented without further delay to

capture the changes in consumption. In practice, it may only be possible to update samples following the normal sample rotation schedule which could even be delayed due to logistical problems caused by the lockdown.

## 3.6 Data validation

3.52 Analysis of outliers may be helpful to eliminate possible measurement errors and/or non-representative price quotations. This is especially important when the total number of registered prices is low, due to unavailability of items or outlet closure. Under such circumstances even one outlier may distort the total index significantly. Outlier analysis can be performed using median price ratios and standard deviations, quartiles, or more advanced methods (see *CPI Manual*, chapter 5 for more details).

3.53 For example, in the U.S. BLS, a semi-automated data review process automatically flags certain sampled items for manual review, if the price change of the item is outside established tolerances levels, set by historical standard error data. Upon review, if a non-sampling error is detected, e.g., a misclassification of the product into the wrong item category, or a keying error, and the error cannot be corrected in time for publication, then the analyst can remove the item from index calculation and it will be imputed. If the price and all price adjuncts are verified as correct, the price is usually accepted for index calculation. Price changes are winsorized such that each month-on-month price relative ( $r$ ) for a sampled product is bounded between  $0.05 \leq r \leq 20$ .

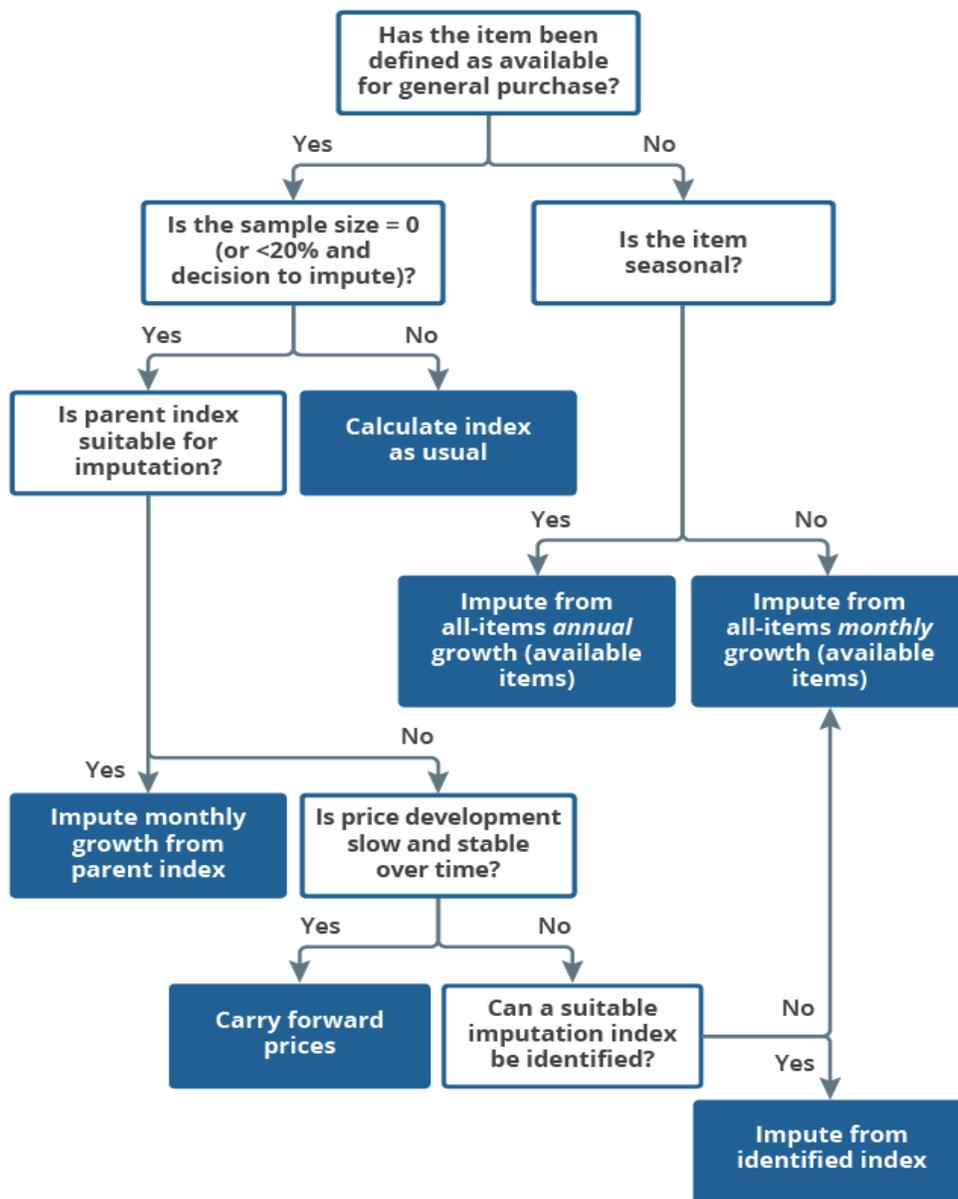
3.54 However, outlier analysis should be conducted with caution to avoid unjustified price deletion. Possible interventions NSOs may take in response to outlier detection include (1) to accept the outlier price as is, (2) modify the price based on auxiliary information that might suggest a non-sampling error has contributed to a wrong price, or (3) to convert the outlier price to missing so that it can be imputed.

## 3.7 Choice of imputation method

3.55 Given the variety of possible imputation methods, decisions must be made about which imputation method is used for which product. Decision rules or principles can be formalized to make the imputation process more organized and transparent. However, even if the decision process is structured and to some extent automated, judgements must still be made by the NSO at the different stages, for example for the selection of a specific donor index.

3.56 In response to the lockdown, the ONS developed an imputation policy as represented by the decision tree in Figure 3.2. For available items a threshold for the minimum size of the sample helps to decide if missing prices should be imputed as usual, or if special lockdown procedures for imputation should be applied. The threshold is calculated at the first suitable level of aggregation, which may be at the level of elementary indices or below that. If the number of prices is too low, or if all the prices of the elementary aggregate are missing, then the prices are estimated using an appropriate imputation method. Depending on circumstances, imputation can be based on the parent index, or carry forward may be applied. If none of these strategies is suitable, any other donor index, including the all-items index, may be used instead. Items that are not available anymore for general purchase are imputed using either the annual (for seasonal items) or the monthly (for non-seasonal items) rate of change observed for the available items.

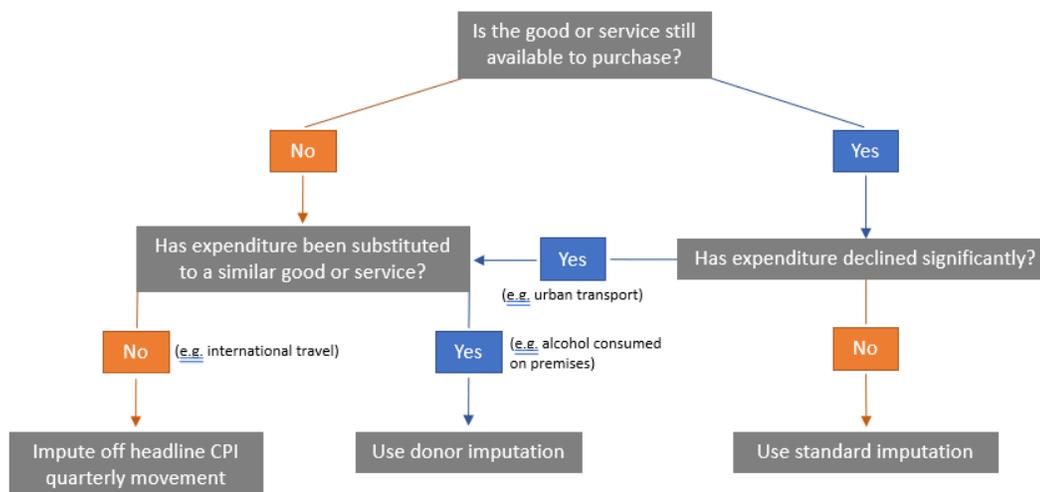
Figure 3.2 Decision tree (Office for National Statistics, United Kingdom)



3.57 It is up to the statistical office to set suitable thresholds on the share of available prices within an aggregate when making decisions on imputation methods. The threshold should be high enough to ensure that the resulting indices have the necessary quality. A common threshold for all elementary indices is easy to implement and monitor but requires careful assessment of the resulting list of products or elementary aggregates that do not meet the minimum requirement. In addition to the proportion of prices that are collected, also the divergence in expected price changes and the weight of the elementary index should be considered. It may be easier to accept a larger share of missing observations for elementary indices that consists of products with expected similar price movements and for elementary indices with relative low weights. In all cases it is helpful to consider the potential impact of imputed prices on the CPI.

3.58 Australian Bureau of Statistics developed the decision tree in Figure 3.3 for the Australian CPI. In this example, standard imputation methods are used for products that are still purchased in sufficient amounts. In the other cases, the imputations are based on a substitute product where possible and otherwise on the overall CPI. For example, it could be assumed that households have substituted alcohol consumed on premises to takeaway alcohol. In such a case, the price change of takeaway alcohol (the ‘donor’) is used to estimate the price change of alcohol consumed in the restaurants. When it is not possible to identify a donor index, imputation is applied with respect to the headline CPI.

Figure 3.3 Decision tree (Australian Bureau of Statistics)



3.59 Another example of a decision tree is presented in Destatis (2021) for the German CPI. Other statistical agencies have developed similar decision support tools or rules not necessarily formalized in a decision tree. Box 3.4 includes an example by the Bureau of Labor Statistics, United States.

#### Box 3.4 Imputation decision model of the Bureau of Labor Statistics (BLS)

The BLS uses the following hierarchical model for deciding on imputation:

- (1) If at least one price quote is available within the item-area cell, impute from within the same elementary item-area cell – usually via a cell-mean or from a similar response from the sample.
- (2) If no source sample is available in the home cell, impute from the exact same item in a different but similar city. Identification of ‘similar’ items and areas are predetermined in a concordance parameter file.
- (3) if all imputation source pools are empty in (2) then impute via carry-forward.

During the Covid-19 crisis, the large majority of cases were handled using rule (1), whereas the remaining cases managed to be treated using rule (2). Carry forward (rule 3) happened to be used only on very rare occasions. These decision rules favoured imputations of the same item across different regions, instead of moving up in the index hierarchy. Carry forward was seen as a method of last resort.

### 3.8 Key points

- The imputation strategy for available products should follow a bottom-up approach considering the reliability of the imputations. The imputations could be based on the average price change of all, or of a subset of available items within the same elementary aggregate. As a next option, the imputation could be based on the price change of a close elementary aggregate. Otherwise, the imputation could be derived from the price change of the parent index in the index hierarchy.
- Specific imputation methods could be considered for unavailable products that are not transacted during the lockdown. Options include imputing with a comparable elementary price index, or with the nearest available higher-level price index, or with the overall CPI. Imputation with the overall CPI corresponds to omitting the elementary aggregate or product group that is being imputed from the calculation of the CPI.
- Imputation methods must be self-correcting to ensure that imputations do not introduce a permanent bias into the index after the lockdown. Similarly, a change in the collection mode or in the specification of the product during the lockdown and a subsequent return to the initial collection mode or product specification after the lockdown should be treated in a symmetric manner to avoid introducing long-term effects in the CPI.
- NSOs should develop consistent metrics to monitor the extent of imputations in the CPI. The definition of imputation rate for public disclosure should fit with related quality metrics such as official response rate definitions and published response rates.
- Statistical agencies should develop and communicate to users a comprehensive imputation policy. Imputation methods and rules should be implemented in the CPI production systems so that they can be applied at any moment. In addition to standard non-response encountered in a CPI, such a policy should consider the possibility of major disruptions such as those seen during the COVID-19 crisis.

# Chapter 4 Expenditure Weights

## 4.1 Introduction

4.1 While the focus of imputation is typically on the price side of the CPI calculus, it is important not to neglect the impact of missing or unusable expenditure data. The global 2020-2021 pandemic caused abrupt changes in spending behaviour as consumers reacted by stockpiling many grocery items and household supplies. In other cases, consumers were prevented from purchasing many services due to local, regional, and national lockdown orders. In some cases, these sudden changes persisted for several months and remain sustained throughout the pandemic, resulting in level-shifts in spending time-series data. This shock to an otherwise historically relative stable pattern of consumer expenditures could undermine the integrity of a CPI based on lagged consumption expenditure data.

4.2 Section 4.2 briefly discusses the possible impact of a lockdown on the CPI through changes in the expenditure weights. Section 4.3 discusses the challenges of deriving weights for future CPIs based on data from the lockdown period and presents country examples of different approaches. Session 4 summarizes key points.

## 4.2 Impact on weights of a lockdown

4.3 Most NSOs calculate the CPI and all higher-level price indices as the expenditure weighted average of the elementary aggregate indices. The weight reference period will usually refer to a year, or an average of several years, that precedes the price reference period in which the weights are introduced in the CPI. Because of this lag, the weight data used as input to CPI calculations had already been collected and processed prior to the emergence of the pandemic.

4.4 However, the changes in households' consumption pattern brought about by the pandemic were perceived to be significantly larger than normal, questioning the relevance of the headline CPI based on historical weights. For instance, Reinsdorf, M. (2020) finds that due to mainly under-weighting of rising food prices and over-weighting of falling transport prices, the CPI will underestimate consumer price inflation in the early months of the pandemic. The issue, then, was whether an NSO should employ some sort of weight adjustment to realign the lagged spending weights with current conditions.

4.5 Most CPIs are based on annual weights. A CPI based on annual weights shows the monthly change in the price of buying the same annual fixed basket. This is conceptually clear but requires use of estimated/imputed prices for products that are not available throughout the year. It is recommended to keep expenditure weights fixed adhering to the regular schedule for updating weights. Changing expenditure weights within a year is not consistent with the fixed basket approach and will make it difficult to interpret the CPI since changes may be caused by changes both in prices and in expenditure weights.

4.6 Ad hoc changes of the weights may also be questioned and may harm the credibility of the CPI. While it is recommended to keep expenditure weights constant it should be noted that imputing missing elementary indices with the all-items CPI corresponds to leaving these out of the CPI compilation, i.e., setting their weights equal to zero and rescaling the weights of the remaining elementary indices to sum to unity (see section 3.3.2).

4.7 Under normal conditions, expenditure weights are relative stable. From year to year changes in the weights usually have only little effect on the overall CPI. Changes in the weights only impact the overall

CPI to the extent that the price development of products differ from the average price development. In the hypothetical case where all elementary indices increase by the same rate of change, the weights will have no influence on the overall CPI at all. Even though this is an unlikely scenario, it highlights the fact that the CPI can be statistically robust to relatively large changes in expenditure weights alone. To impact the overall CPI, changes in weights must be accompanied by changes in relative prices.

4.8 Assume, for illustration, that we have two sets of weights,  $w_A^i$  and  $w_B^i$ , for the same set of elementary price indices  $P^i$  (for example updated weights and the previously used weights). We can then calculate two higher-level price indices,  $P^A$  and  $P^B$ , by aggregating the elementary indices with the two sets of weights. Subtracting the two indices shows the effect of the weights on the higher-level index:

$$\begin{aligned}
 P_{0:t}^A - P_{0:t}^B &= \sum w_A^i \cdot P_{0:t}^i - \sum w_B^i \cdot P_{0:t}^i \\
 &= \sum (w_A^i - w_B^i) P_{0:t}^i \\
 &= \sum (w_A^i - w_B^i) (P_{0:t}^i - \bar{P}_{0:t}) + \bar{P}_{0:t} \sum (w_A^i - w_B^i) \\
 &= \sum (w_A^i - w_B^i) (P_{0:t}^i - \bar{P}_{0:t})
 \end{aligned} \tag{4.1}$$

where  $\sum w_A^i = \sum w_B^i = 1$  and  $\bar{P}_{0:t}$  is the average price change from 0 to  $t$  ( $P_{0:t}^A$  or  $P_{0:t}^B$ ).

4.9 The calculation in (4.1) illustrates that differences in weights only affects the overall CPI to the extent price changes differ. If an elementary index is identical or close to the average price development, changes in its weights will have no influence on the overall CPI. On the other hand, this is a hint to be careful with the treatment of products with particular price developments, e.g., products with pronounced seasonal price variations, since these potentially may have large effects on the overall CPI.

### 4.3 Estimation of weights for future CPIs based on data from lockdown periods

4.10 Since expenditure weights are periodically updated, eventually the period encapsulating the lockdown will be used as weight reference period for future CPI. Would it be appropriate to select the lockdown period as a base-period to represent future CPI calculations? This section provides examples of how the Australian Bureau of Statistics, Statistics Canada and Office for National Statistics, United Kingdom responded to this issue and corresponding guidance issued by Eurostat.

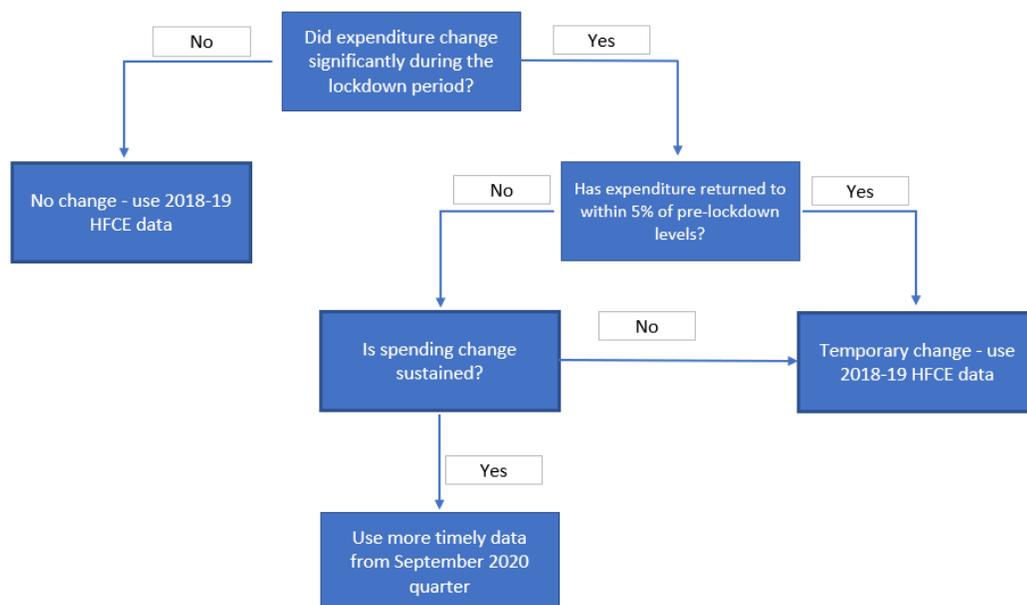
4.11 Because of a lockdown, weights may be unusual, which can justify a special treatment. The *CPI Manual* recognizes problems with unusual weight reference periods (paragraphs 3.72-3.75). It is preferable to choose a “normal” consumption period as the basis for weights and avoiding periods in which there are special factors of a temporary nature at work. To arrive at “normal” weights, the manual suggests adjusting the data to overcome irregularities by e.g. smoothing erratic data or taking averages of more than one year. If the 2020 weights will be used to compile the CPI for, say, 2022-2025 it can also be argued that the weights should be as representative as possible for the average consumption expenditure pattern of this period.

4.12 It may be necessary to carefully examine the weights of goods and services that have been particularly affected by the pandemic, for example for such markets that have been closed. In these cases, additional adjustments of the weights may be justified. Possible adjustment techniques include using 2019 data instead of 2020, composite estimating or averaging 2020 data with prior period data to smooth out outlier changes or forecasting 2020 data from historical time series.

## Australian Bureau of Statistics

4.13 The Australian CPI has been annually re-weighted since 2018, predominately using household expenditure data lagged by 18 months. In response to Covid-19, the Australian Bureau of Statistics (ABS) used a range of more timely retail trade data for approximately 20 per cent of the CPI basket for the 2020 annual re-weight.<sup>1</sup> For the remaining 80 per cent of the basket, traditional updating techniques were used. In response to the pandemic, ABS developed a decision tree to help determine how the weights should be derived, presented in Figure 4.1.

Figure 4.1 **Weight derivation (Australian Bureau of Statistics)**



## Statistics Canada

4.14 In Canada, basket weights are derived primarily from Statistics Canada's Survey of Household Spending (SHS) and are updated every two years. Weights were updated in February 2019 using 2017 SHS data. Statistics Canada decided to delay the 2021 update, planned for February 2021, to July 2021 to assess how actual spending in 2020 may deviate from the 2019 SHS estimates.

4.15 Partnering with the Bank of Canada, Statistics Canada obtained temporary access to aggregated spending data from March, April, and May 2020 representing nearly all CPI product categories, and used these data to simulate an experimental index with more contemporaneous weights compared to the official Canadian CPI. The experiment provided insight into how CPIs can be affected by sudden, extreme shifts in purchasing patterns and the specific impact of Covid-19. The evidence suggested inflation was slightly higher than the headline CPI in April and May.<sup>2</sup> Following the ABS framework, Statistics Canada plans to adjust for item categories that had large changes in 2020. Statistics Canada also plans to execute another weight update in 2022, which usually would not be done, to remove any adjustments made for 2020 that are no longer needed.

<sup>1</sup> See <https://www.abs.gov.au/statistics/research/2020-annual-re-weight-australian-consumer-price-index>

<sup>2</sup> See <https://www150.statcan.gc.ca/n1/pub/62f0014m/62f0014m2020010-eng.htm>

## Office for National Statistics, United Kingdom

4.16 The Office for National Statistics followed Eurostat guidance (Eurostat 2020c) and proceeded with their weight update in 2021 using national accounts data as the source of the weights. The update for 2021 would normally use 2019 data as a proxy for 2020. However, based on Eurostat guidance, ONS used early estimates of 2020 Q1-Q3 to adjust 2019 data to be more reflective of 2020 spending. When there was no evidence of a significant change, ONS used the unadjusted 2019 data (see ONS 2021).

## EUROSTAT

4.17 The HICP is an annually chained Laspeyres-type index. Each year, the weights are updated at the beginning of the year and are kept fixed throughout the year. The weights are based on national accounts data from t-2, and the corresponding expenditure shares are reviewed and updated to make them representative of year t-1. In the context of the pandemic, the legal and conceptual framework of the HICP has been kept unchanged. In line with this framework, guidance for the compilation of HICP weights in cases of large changes in consumer expenditures was issued by Eurostat in December 2020 (Eurostat 2020c). The main principle of the HICP guidance is that the 2021 HICP weights should be based on the best estimates possible for household consumption expenditure patterns in 2020. As a result, the weights used in 2021 include the significant effect that the Covid-19 pandemic has on household consumption patterns.

4.18 In practice, Eurostat outlines an eight-step process to follow for the estimation of 2020 household consumption expenditures by ECOICOP for use in calculating 2021 indexes. Generally, this guidance directs NSOs to use the commensurate 2019 to 2020 change in quarterly consumption observed at least for the three first quarters of 2020 to update 2019 expenditure data.

## 4.4 Key points

- NSOs should avoid making instantaneous changes to index aggregation formula and changes in the frequency of weight updates as a reaction to a lockdown or economic shock.
- If a planned weight update involves spending data for the lockdown period, and those data have higher measurement error than recent history or otherwise could be considered as non-representative, NSOs may modify the planned weight update by skipping the weight update for the impacted year, averaging the impacted year with prior year data, or delaying the weight update until additional analysis can be performed on the spending data.
- For annually updated CPIs, it is possible to derive weights based on the most recent expenditure data. There can be trade-offs between the quality and the timeliness of data sources available for deriving weights. Data sources permitting, shifts in consumption expenditure patterns caused by the lockdown can be incorporated in the index with the annual update of the weights. The NSO may or may not decide to further adjust some of the expenditure data covering the lockdown period (for example by excluding or averaging) considering the measurement objective of the CPI. The weights can be updated again the following year.
- NSOs could develop and produce ex-post experimental or research versions of the CPI that use a superlative index formula to aggregate component indexes into the all-items CPI, or that otherwise use current period spending weights, in order to compare and contrast to the headline CPI that was published using lagged weights.

# Chapter 5 Communication

## 5.1 Introduction

5.1 Because of the widespread use of the CPI for both indexation purposes and as a measure of inflation it is crucial to maintain the public trust in the statistics. To this end, NSOs must ensure transparency which means that information about data sources and methods should be made publicly available to the users of the CPI. Enough documentation and explanations must be provided to facilitate correct interpretation and use of the statistics. Careful communication is particularly important in periods of lockdown to inform users of possible impacts on the statistics and maintaining public trust in the CPI.

5.2 In this chapter, section 5.2 gives broad recommendations on the communication with users and stakeholders under lockdown. Section 5.3 provides more details about measures of quality and reliability of the CPI, including the ratio of imputed prices, sensitivity analysis and decomposition, illustrated with practical examples. Section 5.4 gives examples of experimental CPIs that may be compiled to compliment the headline CPI, data sources for weights and special CPI aggregates that may be useful to inform users. Section 5.5 summarizes key points of the chapter.

## 5.2 Communication with users and stakeholders

5.3 The CPI should be published following best practices for the dissemination of official statistics and the recommendations of the *CPI Manual* chapter 14 on Publication.

5.4 During a lockdown, users will continue to need CPI series at a detailed level. NSOs, therefore, should continue to publish all index series that are usually disseminated whenever possible. Imputed series should be flagged and clearly communicated to users and accompanied by additional information about index quality. This provides users with a full set of index series and helps to ensure transparency.

5.5 Because of the lockdown there may have been considerable changes in the sample of goods and services. New price collection methods and new data sources may have been implemented and new imputation methods may have been applied. Users should be informed about changes in data sources and methods that may affect the statistics. Hence, changes in the sample, data sources, and compilation methods should be documented, published and explained to users. The price collection period is of key importance for the CPI. If this has been changed users should be informed. Additional information about the impact of the lockdown on the quality of the CPI may also be published (see section 5.3 below).

5.6 The monthly (or quarterly) production process may be delayed because of the lockdown, which may make it difficult to meet the pre-announced publication date of the CPI. Eventually, it may not be possible to release the CPI at the announced date. The NSO needs to evaluate if there is a risk of missing the publication date. If the release must be postponed users should be informed about this in advance and a new release date should be communicated. NSOs are encouraged, if possible, to estimate when the publication schedule is expected to return to its pre-defined dates.

5.7 If the release of the CPI is postponed some NSOs may consider publishing a preliminary CPI. However, in this case the statistics should be clearly marked as preliminary. If the CPI is subject to particular uncertainty this should also be highlighted. Users should be informed when the final CPI will be released and that this may deviate from the preliminary CPI.

5.8 The CPI is used for a variety of purposes, including indexation of wages, pensions and social benefits and as a measure of inflation in monetary and economic policy. When communicating the CPI the different group of users should be considered. User groups may have different needs for information and documentation so various information should be provided to meet these needs. Examples of such information are given below.

5.9 Suitable information should be made available together with the release of the CPI. Documentation and information may be provided through the official website of NSOs, through social media and press-conferences. It may also be possible to create a separate webpage that entails the relevant documentation and explanations in relation to the lockdown and its effect on the CPI and FAQs. This is especially user-friendly and timesaving for those who need quick access to specific information.

### Reaching out to users and stakeholders

5.10 Online press conferences are useful communication ways to give more detailed information to users. There is also the possibility of organising online events/webinars for specific user groups, for instance researchers, the government sector or others to focus on the issues most relevant to these groups. Timely, comprehensive and transparent information provided to users and stakeholders are important to ensure public trust in official statistics. Extensive communication with user groups during the Covid-19 pandemic was carried out, for example, by statistical offices of Australia<sup>1</sup>, Germany<sup>2</sup>, France<sup>3</sup>, United Kingdom<sup>4</sup> and United States US<sup>5</sup>.

5.11 It is useful to be proactive in communication with users. For instance, the NSO may reach out to key users such as central banks, ministries of finance and labour in advance and inform them about expected changes. This could also include economists, researchers and users that use the CPI for indexation purposes, including governments and labour organizations. In cases where the CPI is used for indexation this is likely to create additional interest in the possible effect on the CPI of the lockdown since this may have significant consequences for groups that have their income or pension indexed with the CPI. At the same time, it is highly important to ensure equal availability of information for all user groups, in line with the UN Fundamental Principles of Official Statistics and the principle of “Impartiality and Objectivity” of the European Statistics Code of Practice.

5.12 Two-way communication with user groups may help to establish relationships. Contact to key users may also be used for consultation and soliciting inputs and views that may be helpful to the NSO when deciding on methods and communication. Consultations may take place bilaterally or the NSO may organize (online) briefings for selected users. Important changes in data sources or methods should be communicated to the public well in advance and would also in many countries involve consultation with user groups and advisory boards on CPI.

## 5.3 Informing users about the quality and reliability of the CPI

5.13 Changes in the sample of goods and services or in compilation methods caused by the lockdown are likely to affect the quality of the CPI. Countries are recommended to compile and release measures to inform users about how much the lockdown may have impacted the quality of the CPI.

---

<sup>1</sup> [https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2020/Importance\\_of\\_communication.pdf](https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2020/Importance_of_communication.pdf)

<sup>2</sup> <https://www.destatis.de/EN/Themes/Economy/Prices/Consumer-Price-Index/Methods/corona-cpi-hicp.html>

<sup>3</sup> [https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2020/UNECE\\_webinar\\_Insee\\_202011.pdf](https://unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2020/UNECE_webinar_Insee_202011.pdf)

<sup>4</sup> <https://www.ons.gov.uk/economy/inflationandpriceindices/articles/coronavirusandtheeffectsonukprices/2020-05-06>

<sup>5</sup> <https://www.bls.gov/covid19/effects-of-covid-19-pandemic-on-consumer-price-index.htm>

5.14 Statistical agencies should develop and communicate to users a comprehensive policy on publication adequacy. The policy should directly and objectively adjudicate a publication decision for each index and data product planned for publication, evaluating whether sufficient input data have been directly collected to achieve public confidence in the published measure. The objective adequacy rules developed by each NSO should be sample size or variance based, and robust to contemplate small to significant amounts of nonresponse and imputed prices. Adequacy rules should not be subjectively changed in response to market shocks that may temporarily alter the amount of collected input data.

### 5.3.1 Publication of heavily imputed index series

5.15 An increase in the number of imputations immediately impacts the reliability of the resulting published indices. Should indices be suppressed from publication if a large percentage of the index is estimated without directly observed prices? If so, what constitutes a 'large' percentage? This section highlights the publication practices of example NSOs regarding the impact of imputation on the quality of officially released statistics.

#### Bureau of Labor Statistics, United States

5.16 The U.S. CPI employs on monthly basis an adequacy criterion to determine whether a particular aggregate index should be published. The rule is three-part:

- (1) First, each elementary cell is assigned a dummy variable, 1=Non-imputed, 0=imputed. A cell is deemed 'non-imputed' if at least one sampled item in the cell has a directly observed price. This includes prices observed for match-model products, comparable substitutes, and quality-adjusted substitutes.
- (2) Second, a weighted average of the binary 0,1 variables for each elementary cell defining the aggregate is computed, using the aggregation spending weights assigned to each elementary cell. The result is an 'average' non-impute rate for the aggregate,  $0 \leq \alpha \leq 1$ .
- (3) If  $\alpha > 0.5$  then the aggregate index is deemed 'adequate' for publication. That is, if more than 50% of the weighted inputs are non-missing, then the aggregate estimate can be published. If  $\alpha \leq 0.5$ , then the index estimate is suppressed from publication and footnoted accordingly in published news releases and tables.

5.17 At first glance, the BLS adequacy rule seems to suggest at least 50% of the sample must be collected to publish results. But that is not the case. Closer inspection reveals that just one collected price in each elementary building block is required. This equates to roughly 5% of the total sample size. Moreover, if the collected prices were distributed in the largest weighted categories, e.g. rent, gasoline, food away from home, then the U.S. CPI could be deemed adequate for publication with as low as 2.5% of the total sample collected.

#### Eurostat

5.18 Guidance from Eurostat (Eurostat 2020a) on this issue is two-part. First, if there is more than 50% imputation in a published index, countries are to flag or denote the index as such. Moreover, for each flagged index, additional information on the imputation method and the imputation ratio are published each month. This guidance permits the publication of all sub-indices. For each sub-index, the European Union aggregate is compiled as the weighted average of the country indices. A European Union index is flagged if composed of more than 50% (in terms of weight) of country indices flagged as imputed.

## Office for National Statistics, United Kingdom

5.19 The United Kingdom adopts a similar, but slightly varied policy relative to the Eurostat guidance. In extreme lockdown cases, very small samples at the low-level strata classification level have been rejected and the item index imputed. A flag is included at the item level which indicates whether the item is fully imputed or partially imputed. At higher COICOP levels, series are flagged which are based on less than 50% of the usual sample size.

### 5.2.2 Ratio of imputed prices

5.20 In addition to flagging indices with a high ratio of imputed prices the ratio of imputed prices may also be published. This ratio can be a simple count of the number of imputed prices out of the total number of prices (more details are given in section 3.4). The ratio can also take the weight of the individual price relatives into account. Aggregating the weighted price relatives or indices would give an indication of how much of the weighting basis of the overall CPI has been imputed, and how large impact potentially this may have on the overall CPI.

### 5.3.3 Ratio of closed outlets

5.21 The ratio of closed outlets will be an indication of how much different markets have been impacted by the lockdown and hence how much different sub-indices may have been impacted. If compiled monthly, it will also be an indicator of the severity of the lockdown.

### 5.3.4 Indicators of data source and price collection method

5.22 It is also possible to compile ratios or other types of indicators to reflect how the total number of observed prices is distributed on different data sources or collection methods. For instance, the share of prices collected in the field by price collectors visiting the outlets, from the web, from survey questionnaire or by phone or e-mail. While information on data source and collection methods may not be of interest to all users it may be relevant so some, and for the index compilers it may help to monitor shifts in data sources and collection methods over time.

5.23 Examples of such auxiliary data can be found on websites of statistical agencies, including United States (figure 5.1), Canada (figure 5.2) and Belgium.<sup>1</sup>

Figure 5.1 Report on price collection mode under lockdown (Bureau of Labor Statistics)

Month	2019			2020		
	Personal visit	Telephone	Online	Personal visit	Telephone	Online
March	74	10	16	44	14	42
April	74	10	16	0	18	82
May	74	10	16	0	17	83
June	73	10	17	0	17	83
July	74	10	16	0	17	83
August	73	10	17	0	17	83
September	73	10	17	0	17	83
October	73	10	17	0	16	84
November	73	9	18	0	16	84
December	72	10	18	0	16	84
Average, Feb 2019–Feb 2020	73	10	17			

Source: <https://www.bls.gov/covid19/consumer-price-index-covid19-impacts-december-2020.htm>

<sup>1</sup> <https://statbel.fgov.be/en/news/impact-covid-19-index-calculation-and-measurement-inflation>

Figure 5.2 Report on item availability and outlet closure during lockdown (Statistics Canada)

Consumer Price Index product and outlet availability, field, scanner and web scraped data collection modes, April 2020 and April 2019

	Basket weight at link month (%)	Average rate of out of stock product offerings (%)		Average rate of temporarily closed outlets (%)		Average rate of prices received as scheduled (%)	
		April					
		2020	2019	2020	2019	2020	2019
<b>Total</b>	<b>100</b>	<b>13.9</b>	<b>4.4</b>	<b>3.2</b>	<b>0.3</b>	<b>82.9</b>	<b>95.3</b>
Food and non-alcoholic beverages	16.48	11.3	1.8	1.3	0.1	87.5	98.1
Shelter	27.36	9.0	2.4	5.2	0.9	85.7	96.7
Household operations, furnishings and equipment	12.8	20.2	6.2	1.9	0.2	77.9	93.6
Clothing and footwear	5.17	15.1	14.1	9.0	0.4	75.9	85.4
Transportation	19.95	1.8	0.9	7.6	1.4	90.6	97.7
Health and personal care	4.79	28.6	3.0	6.4	0.1	65	96.9
Recreation, education and reading	10.24	14.8	5.7	5.7	0.5	79.5	93.7
Alcoholic beverages, tobacco products and recreational cannabis	3.21	9.3	1.9	1.1	2.8	89.6	95.3

Source: <https://www150.statcan.gc.ca/n1/pub/62f0014m/62f0014m2020006-eng.htm>

5.24 Measures of the impact of a lockdown on the CPI can be published as supplementary information. Additional information and examples of imputation ratios and use of imputation methods are available on Eurostat's website<sup>1</sup>. Another example of auxiliary information is presented in Table 5.1.

<sup>1</sup> <https://ec.europa.eu/eurostat/documents/272892/272974/Overview-flagged-HICP-sub-indices-Jan-2021.xlsx/>

Table 5.1 Example of auxiliary CPI information

COICOP divisions	Number of price observations under normal circumstances	Number of price observations in lockdown month	Share of imputed prices	Share of closed outlets	Flag (share of imputed prices >50%)
<b>Overall CPI</b>	<b>16065</b>	<b>9146</b>	<b>43.1%</b>	<b>39.4%</b>	-
01 Food and non-alcoholic beverages	5561	4292	22.8%	13.9%	-
02 Alcoholic beverages and tobacco	387	318	17.8%	12.9%	-
03 Clothing and footwear	2297	145	93.7%	96.4%	Flag
04 Housing, water, electricity, gas and other fuels	540	308	43.0%	33.0%	-
05 Furnishings, household equipment and routine household maintenance	1855	764	58.8%	48.5%	Flag
06 Health	1069	855	20.0%	17.9%	-
07 Transport	916	634	30.8%	27.3%	-
08 Communication	288	171	40.6%	44.6%	-
09 Recreation and culture	1392	442	68.2%	77.4%	Flag
10 Education	186	146	21.5%	19.4%	-
11 Restaurants and hotels	282	78	72.3%	69.5%	Flag
12 Miscellaneous goods and services	1292	993	23.1%	20.6%	-

### 5.3.5 Measures of statistical certainty

5.25 When products are collected through probability sampling the statistical certainty of an index can be calculated by use of variance estimation of the price relatives. However, while probability sampling may be used for some product groups, it may not be used for others which make it difficult to derive a measure for the statistical uncertainty of the overall CPI<sup>1</sup>.

5.26 Analysis of index variance may also be used for decisions about flagging or not a particular price index, in parallel with information about the ratio of imputed prices. A threshold can be estimated by index compilers above which the index data will be considered as of low quality. When analysing the variance NSOs must consider differences between outlet types and products. E.g., threshold for food and beverages can be set higher than the one for services. Also, smaller outlets can have very different pricing strategies, compared to larger outlet chains and open-air markets.

5.27 Attention should be paid also to the quality of regional indices if compiled and disseminated. Some regions may be affected more heavily by the lockdown than others and, as a result, their index quality will deteriorate more than other regions' indices or the national CPI.

<sup>1</sup>When non-probability sampling is used, additional analyses are required (see *CPI Manual*, chapter 12, paragraphs 12.13-12.16).

### 5.3.6 Sensitivity analysis

5.28 Users may ask how and how much the lockdown has impacted the CPI. To this end, it may be useful to inform users about the importance (the weight) of goods or services for which markets have been closed temporarily, and what potential impact this may have on the overall CPI. This could also include information about how missing goods or services are treated in the CPI, where these have a significant weight or have been attracting media interest, for instance childcare institutions, air travels or hotels/restaurants.

5.29 When the CPI is based on annual weights it can be helpful to explain to users what implications this has, i.e. that (temporary) changes in the consumption pattern during the lockdown is not reflected in the CPI and that the CPI may differ from households' perceived inflation.

5.30 Also changes in the price collection period may impact the CPI. While in general the regular price collection period should be maintained, the NSO may have extended the period with some days to allow more prices to be collected. As a result, figures of a particular month may represent inflation of a period longer than one month, while the next or previous figure will represent inflation of less than one month. Such effects will be of temporary nature and eliminated when returning to the regular price collection period.

### 5.3.7 Decomposition

5.31 Users of the CPI may be interested in how much different sub-indices contribute to the change in the overall CPI. This can be compiled by decomposing the change in the overall index into its constituent parts. Such decomposition showing, for instance, the contribution of each aggregation level of COICOP (or another classification used on national level) to the overall CPI is common in many countries as part of the regular publication of the CPI. Decomposition and the construction of a decomposition table are explained in the *CPI Manual* (9.104-9.114).

5.32 Under normal conditions, each sub-index contributes to the rate of change of the overall CPI according to the weight of the sub-index, irrespective of whether some of the prices entering the sub-index are actually observed or imputed. However, during a lockdown some elementary aggregates or even higher-level aggregates may not be available. Indices for unavailable elementary aggregates may have been imputed by other available elementary price indices or by the overall CPI (corresponding to leaving the elementary index out of the calculation of the CPI). In such cases, in order to calculate the actual contribution of each elementary index or each sub-index, the weights must be rescaled to sum to unity based only on products that are available on the market, i.e., those elementary indices or sub-indices for which prices have been observed. This will result in contributions that are assigned only to products available for purchase (but still summing up to the total inflation rate), which can be easily explained to the public. An example is presented in Table 5.2 for illustration.

Table 5.2: Index decomposition

	COICOP Division	Group weights % (1)	Index January/ December (2)	Available indices (3)	Normalized weights (4)	Contribution to the monthly index in January (Dec. = 100) (% points)**
1	Food and non-alcoholic beverages	32.2%	100.2	100.2	39.3%	0.08
2	Alcoholic beverages and tobacco	6.4%	100.0	100.0	7.8%	0.00
3	Clothing and footwear	3.8%	99.75*			
4	Housing, water, electricity, gas and other fuels	9.0%	100.9	100.9	11.0%	0.10
5	Furnishings, household equipment and routine household maintenance	5.3%	99.75*			
6	Health	7.8%	103.1	103.1	9.5%	0.29
7	Transport	13.0%	95.3	95.3	15.9%	-0.74
8	Communication	3.6%	102.4	102.4	4.4%	0.11
9	Recreation and culture	3.7%	99.75*			
10	Education	4.6%	97.2	97.2	5.6%	-0.15
11	Restaurants and hotels	5.4%	99.75*			
12	Miscellaneous goods and services	5.3%	101.0	101.0	6.5%	0.06
	<b>Overall CPI</b>	<b>100.0%</b>	<b>99.75</b>		<b>100.0%</b>	<b>-0.25</b>

\* Since these divisions are unavailable in the current period, the overall CPI is used for imputation.

\*\* Contributions are calculated based only on indices of available divisions and their normalised weights (4).

As an example, the contribution of “Food and non-alcoholic beverages” to the change in the overall CPI is calculated as:

$$\begin{aligned}
 &= \frac{(\text{Index of Division 1 in January} - \text{Index of Division 1 in December})}{\text{Total index in December}} \times \\
 &\quad \times \frac{\text{Weight of Division 1}}{\text{Sum of weights of available divisions}} \times 100 = \\
 &= \frac{(100.2 - 100)}{100} \times \frac{32.2\%}{81.9\%} \times 100 = 0.08 \text{ percentage points}
 \end{aligned}$$

## 5.4 Experimental CPIs

5.33 While the general recommendation is to continue compiling the regular headline CPI based on existing annual weights and imputed missing prices, it may be useful to compile supplementary experimental or analytical CPI series. These may not meet the quality standards of the regular CPI but may still help to picture the development and meet user needs. Analytical series may be compiled simultaneously with the official CPIs or with a certain time lag. Experimental CPIs should not replace the regular CPI and be clearly marked as experimental or analytical if published to avoid any confusion with the headline CPI.

### 5.4.1 Experimental CPI based on available products only

5.34 The simplest approach is to compile an experimental CPI based on available products only. This can be done simply by omitting products or elementary aggregates for which there were no transactions during the lockdown. By rescaling the expenditure weights of the available products or elementary aggregates to sum to unity, an experimental “Lockdown CPI” can be calculated based on available expenditure data and price index series.

5.35 Rescaling the expenditure weights corresponds to assuming households maintain fixed relative expenditure shares for the products or elementary aggregates that are available for purchase. In practice, this may not be the case. For instance, depending on the restrictions that are implemented, there may be significant shifts in households’ purchase of food products or transport services, households may hoard particular goods etc. The effects of such shifts will not be captured by an index based on the existing weighting structure of available products.

### 5.4.2 Experimental CPI based on estimated lockdown weights

5.36 If weights during the lockdown period are available or can be estimated, it may be possible to compile an experimental CPI reflecting the price development of the lockdown basket. These weights may be more volatile and less reliable than annual weights derived from the regular household budget survey but may still serve as estimates. Timely short-term (monthly or quarterly) expenditure weights estimates may also provide insight into consumer behaviour, which information may be published or shared with interested users.

5.37 This approach may be implemented in two ways. (1) The experimental CPI can be based on weights that are kept constant throughout the lockdown period. (2) The experimental CPI can be based on monthly weights, or quarterly weights, if a quarterly index is compiled. However, it should be considered that an index based on monthly (or quarterly) weights is subject to bias due to frequent chain-linking.

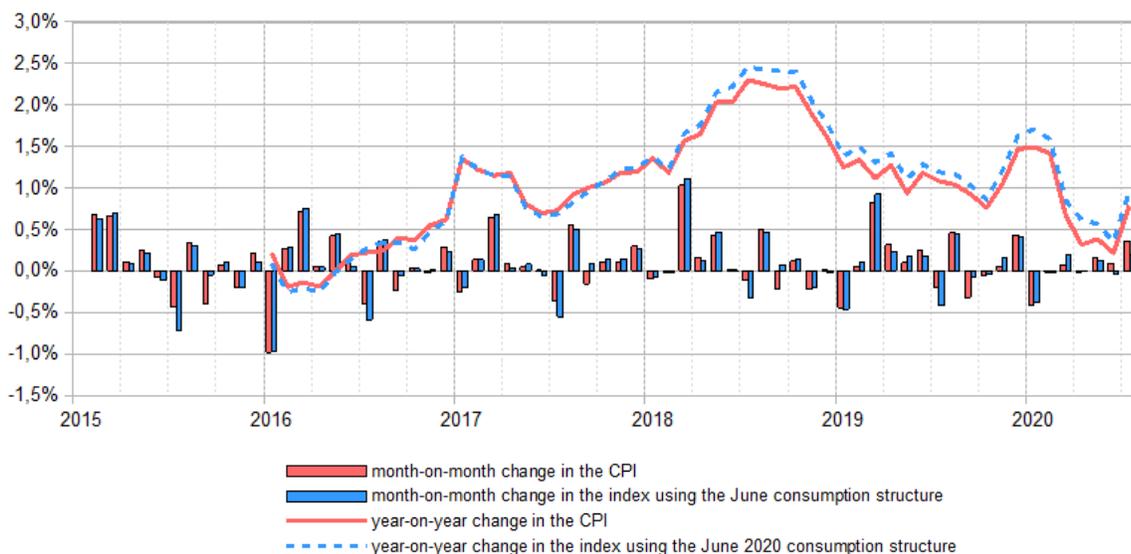
5.38 Available data sources will often be incomplete compared to providing a full set of expenditure weights for the CPI. Hence, compilers are advised to investigate possible alternative data sources that can be used to complement or improve weights estimates. Consumer behaviour should also be taken into consideration. Some consumers may choose to spend less in response to economic uncertainties under lockdown while they may be less prone to reduce consumption of food products, as opposed to durables. Therefore, CPI compilers are advised to consult with area experts.

5.39 An experimental CPI can be compared with the regular CPI based on historical weights. Below is an example by the statistical office of France.

## National Institute of Statistics and Economic Studies (INSEE), France

5.40 INSEE in 2020 calculated an alternative price index for France based on estimated expenditure weights for June 2020, reproduced in Figure 5.3 (INSEE 2020a). The estimated weights were mainly derived from credit card transactions and scanner data. Comparing the alternative price index with the CPI (which was based on weights from 2019) the alternative index in June increased by 0.0% against 0.1 of the CPI. The alternative index showed an annual rate of change of 0.4% against 0.2% for the CPI.

Figure 5.3 The French CPI compared with an alternative price index



### 5.4.3 Data sources for expenditure weights

#### Continuous household expenditure survey

5.41 Some countries conduct a continuous survey of household expenditures, which can provide valuable information on change in the expenditure pattern. However, rolling sample surveys are usually based on smaller sample sizes compared to annual surveys and therefore also associated with larger statistical uncertainty, which should be taken into consideration.

#### Scanner data

5.42 Countries that already have integrated scanner data in their CPI surveys can use this data to obtain up-to-date information on consumer expenditures. Since scanner data are usually available right after the end of reference period (or even during the reference month), this will enable NSOs to calculate current weights under the lockdown. However, scanner data is not available in all countries and in countries where scanner data are available, they will usually only be available for a subset of products of the CPI; for instance, some outlets or open-air markets may not have scanner equipment. Hence, for some products scanner data cannot be obtained, although the products are available for purchase. Therefore, scanner data can mainly be used only as auxiliary information.

#### Credit and debit card information

5.43 Using this data for weights calculation requires close communication with financial institutions to obtain transaction data about households, which may not always be possible due to legal restrictions.

Besides, some purchases are conducted using cash instead of bank cards. Therefore, using this type of data has similar limitations as scanner data. However, during the Covid-19 pandemic, consumers tended to use bank cards more frequently, making this type of data more complete than in usual circumstances. For an example of detailed country experiences, see INSEE (2020b).

5.44 It must be considered that lockdown periods may be characterized by panic buying of essential products. This may result in large changes in consumption pattern from month to month. Therefore, it is recommended to frequently update the experimental weights using all available information and, e.g., using the average of several months if needed. Comparing the weights of the regular CPI with the estimated lockdown weights can provide a useful indication of the potential impact of the change in consumption pattern and which products are particularly affected. Carvalho et al. (2020) provides an analysis of using transaction data from both credit cards and point of sales terminals to track the development in consumption pattern in Spain in the first period of the Covid-19 pandemic.

#### 5.4.4 Publication of special aggregates

5.45 Special aggregate indices can be derived as part of the regular CPI compilation by grouping the relevant set of elementary indices or sub-indices. Such aggregates can be compiled and published at almost no additional costs. Depending on the situation in a country and user needs, numerous indicators can be derived and disseminated based on the existing CPI data:

- CPI for food products and beverages
- CPI for essential products<sup>1</sup>
- CPI for automotive fuels (due to increased expenditures for operation of personal transport equipment where public transport services were suspended)
- CPI for water supply, electricity, gas and other fuels
- CPI for online purchases
- CPI for education (prices may change due to massive switch to online learning)
- CPI for telecommunication services, etc.

5.46 A lockdown may be accompanied by economic and social programs aimed at supporting business activities and households. This might include reduced or deferred taxes, subsidies on public transport fees, childcare, health care or other services. CPI compilers should pay attention to such programs as they may affect consumer prices and thus must be reflected in the CPI. An example is government subsidies on childcare in Australia<sup>2</sup>. While the *CPI Manual* provides recommendations on incorporating such kind of subsidies in the CPI (paragraphs 2.110, 2.112, 8.26, 11.295, 11.298, 11.280; table 11.14), NSOs can produce additional indicators such as CPI without subsidies for analytical purpose, if these subsidies bring large fluctuations to the index and cause questions among users.

5.47 However, production of alternative indicators should be treated with cautions. As stated in the *CPI Manual*, index compilers should avoid confusing the users with different inflation measures. It should be always clearly stated which indicator measures the headline inflation and which is an analytical, auxiliary index. Besides, the purpose and use of each indicator should be explained, to avoid misuse and misunderstanding of the published data. Finally, all additional CPI data should be in compliance with existing methods and recommendations, should not be in contradiction to and should not replace the headline inflation measure.

---

<sup>1</sup> <http://www.statssa.gov.za/?p=13253>. See also Reinsdorf, M. (2020).

<sup>2</sup> <https://www.abs.gov.au/articles/methods-changes-during-covid-19-period#measuring-the-consumer-price-index-during-a-time-of-covid-19>

## 5.5 Key points

- During lockdown NSOs should continue disseminating CPI data at the common level of disaggregation, according to the pre-defined publication calendar, whenever possible.
- Proactive communication with users and stakeholders is helpful to maintain transparency and credibility of disseminated figures. Changes in data sources, data collection modes and compilation methods that may happen during lockdown should be documented and made public.
- When data quality is heavily impacted by lockdown, it is advisable to publish auxiliary information about the CPI figures, such as: ratio of imputed prices, ratio of closed outlets, measures of statistical uncertainty, decompositions of contributions to the overall CPI etc.
- Various analytical indicators can be produced to accompany the headline inflation figures. These may be derived by estimating lockdown basket weights or constructing various aggregates from available data.

# References

- Destatis, Federal Statistical Office of Germany. 2021. *Impact of the corona crisis on the CPI/HICP price collection*. <https://www.destatis.de/EN/Themes/Economy/Prices/Consumer-Price-Index/Methods/corona-cpi-hicp.html>
- Diewert, W. E. and Kevin. J. Fox. 2020. *Measuring real consumption and CPI bias under lockdown conditions*. NBER Working paper 27144. <https://statswiki.unece.org/x/roKSE>
- Carvalho, Vasco M., et al. 2020. *Tracking the Covid-19 crisis with high-resolution transaction data*. <https://statswiki.unece.org/x/roKSE>
- Eurostat. 2020a. *Guidance on the compilation of the HICP in the context of the COVID-19 crisis*. Available at: [https://ec.europa.eu/eurostat/documents/10186/10693286/HICP\\_guidance.pdf](https://ec.europa.eu/eurostat/documents/10186/10693286/HICP_guidance.pdf)
- Eurostat. 2020b. *Guidance note on HICP issues emerging from the lifting of lockdown measures*. [https://ec.europa.eu/eurostat/documents/10186/10693286/HICP\\_lifting\\_lockdown\\_measures\\_guidance.pdf](https://ec.europa.eu/eurostat/documents/10186/10693286/HICP_lifting_lockdown_measures_guidance.pdf)
- Eurostat. 2020c. *Guidance on the compilation of HICP weights in case of large changes in consumer expenditures*. Available at: <https://ec.europa.eu/eurostat/documents/10186/10693286/Guidance-on-the-compilation-of-HICP-weights-in-case-of-large-changes-in-consumer-expenditures.pdf>
- ILO, IMF, OECD, European Union, United Nations, World Bank. 2020. *Consumer Price Index Manual: Concepts and Methods*. [www.imf.org/cpi](http://www.imf.org/cpi).
- INSEE. 2020a. *How to compute a Consumer Price Index in the context of the Covid-19 crisis?* [https://www.insee.fr/en/statistiques/documentation/Data%20quality%20Covid19\\_IR\\_06\\_2020.pdf](https://www.insee.fr/en/statistiques/documentation/Data%20quality%20Covid19_IR_06_2020.pdf)
- INSEE 2020b. *Household Consumption*. French Economic Outlook. [https://www.insee.fr/en/statistiques/fichier/4620512/point-conj170620\\_consoM.pdf](https://www.insee.fr/en/statistiques/fichier/4620512/point-conj170620_consoM.pdf)
- Link, Koen and Antonio Chessa. 2020. *Impact of imputation methods on the CPI and HICP in view of the COVID-19 crisis*. <https://statswiki.unece.org/x/roKSE>
- Intersecretariat Working Group on Price Statistics (IWGPS). 2020. *Consumer Price Index: Business Continuity Guidance*. <https://statswiki.unece.org/x/roKSE>
- Lamboray, C., Evangelista, R., Konijn, P. 2020. *Measuring inflation in the EU in times of COVID-19*. EURONA Issue 2020. [https://ec.europa.eu/eurostat/cros/content/measuring-inflation-eu-times-covid-19-claude-lamboray-rui-evangelista-and-paul-konijn\\_en](https://ec.europa.eu/eurostat/cros/content/measuring-inflation-eu-times-covid-19-claude-lamboray-rui-evangelista-and-paul-konijn_en)
- Merrington, Leigh, Australian Bureau of Statistics. 2020. *Importance of transparency and communication and the value of scanner data*. <https://unece.org/info/Statistics/events/348358>
- Office for National Statistics (ONS). 2020. *Coronavirus and the effects on UK prices*. <https://www.ons.gov.uk/economy/inflationandpriceindices/articles/coronavirusandtheeffectsonukprices/2020-05-06>
- Office for National Statistics (ONS). 2021. *Coronavirus (COVID-19) and Consumer Price Inflation weights and prices: 2021*. <https://www.ons.gov.uk/economy/inflationandpriceindices/articles/coronaviruscovid19andconsumerpriceinflationweightsandprices/2021#collection-of-reference-prices-for-cpih-cpi-and-rpi>
- Reinsdorf, Marshall. 2020. *Is Inflation Underestimated?* IMF working paper No. 2020/224. <https://www.imf.org/-/media/Files/Publications/WP/2020/English/wpia2020224-print-pdf.ashx>

Sands, Helen. 2020. *Producing supplementary analysis to assess the impact of lockdown on CPI*. United Kingdom, Office for National Statistics (ONS). <https://unece.org/info/Statistics/events/348358>

Statistics Canada. 2011: *Glossary of Terms*. <https://www150.statcan.gc.ca/n1/edu/power-pouvoir/glossary-glossaire/5214842-eng.htm#i>

Statistics Sweden. 2020. *Measurement issues related to the coronavirus pandemic*. [https://www.scb.se/contentassets/9cd323f08a894b1fb50c82706fcb435a/measurement-issues-related-to-the-coronavirus-pandemic\\_202012.pdf](https://www.scb.se/contentassets/9cd323f08a894b1fb50c82706fcb435a/measurement-issues-related-to-the-coronavirus-pandemic_202012.pdf)

UNECE. 2000. *Glossary of terms on statistical data editing*. <https://unece.org/DAM/stats/publications/editing/SDEGlossary.pdf>

UNECLAC. 2020. *Producing the consumer price index (CPI) and the COVID-19 pandemic in Latin America and the Caribbean*. [https://statswiki.unece.org/download/attachments/278037166/S2000283\\_en.pdf?version=1&modificationDate=1589299675070&api=v2](https://statswiki.unece.org/download/attachments/278037166/S2000283_en.pdf?version=1&modificationDate=1589299675070&api=v2)

# Guide on producing CPI under lockdown

The lockdown that followed the outbreak of the Covid-19 pandemic in 2020 posed unprecedented challenges to produce the consumer price index (CPI) in many countries.

The closing of retail outlets and difficulties with collecting prices from outlets that remained open caused significant drops in the number of collected prices.

Statistical offices therefore had to explore new data sources and implement methods to make up for missing price observations and changes in data sources and collection methods to ensure the compilation of a reliable CPI.

Based on countries' experiences with producing the CPI during the pandemic, this Guide provides recommendations and emerging best practices for CPI data collection, calculation methods and communication under lockdown conditions.

Information Service  
United Nations Economic Commission for Europe

Palais des Nations  
CH - 1211 Geneva 10, Switzerland  
Telephone: +41(0)22 917 12 34  
E-mail: [unece\\_info@un.org](mailto:unece_info@un.org)  
Website: <http://www.unece.org>