Working Party on Financial Statistics

International comparisons of the measurement of non-market output during the COVID-19 pandemic

This joint ONS-OECD report is exploring differences in methodologies used by countries to measure non-market output and analysing their implications for international comparisons of GDP growth during the coronavirus (COVID-19) pandemic.

It will be presented at the 2021 joint meeting of the Working Party on Financial Statistics (WPFS) and the Working Party on National Accounts (WPNA) under Item 2.e.

Contacts: john.mitchell@oecd.org or Jorrit.zwijnenburg@oecd.org.

JT03483292
International comparisons of the measurement of non-market output during the COVID-19 pandemic

A report by the ONS and the OECD

1. Introduction and rationale

1. The coronavirus (COVID-19) pandemic and the measures taken to control its spread have had severe impacts on economies across the world. Restrictions on contact and travel reduced accessibility of many goods and services, and in some cases spurred transformation in how goods and services were delivered.

2. The effects of COVID-19 have also significantly affected the provision of non-market services, such as healthcare and education. The non-market sector is characterised by providing goods and services without economically significant prices, a feature that poses challenges to measuring its economic output. Various different approaches are available to meet these challenges, the use of which varies between countries and between components of non-market output. The COVID-19 pandemic has presented further measurement challenges, which may have directly affected the comparability of non-market output measures across countries, and indirectly the comparability of Gross Domestic Product (GDP) measures.

3. The purpose of this report is to inform on the range of approaches currently applied by National Statistical Institutes (NSIs) to measure non-market output, and investigate how different approaches may have contributed to differences in the measured growth of non-market output and GDP over the COVID-19 pandemic. To this end, the report provides a brief overview of the methodologies available, examines approaches as applied by NSIs, with a focus on a group of NSIs interviewed as part of this project. While this report discusses implications for the interpretability of non-market output, it does not aim to recommend specific compilation practices. This may be the topic of a follow-up study.

4. The paper is structured as follows. Section 2 provides the rationale for this report, presenting how different growth rates in non-market output impacted differences in GDP growth rates across countries. Section 3 details the recommended methodologies for measuring non-market services. Section 4 outlines which methodologies are currently in use by NSIs before explaining how these different methodologies have been affected by the COVID-19 pandemic in Section 5. Section 6 concludes the paper and provides some basic suggestions on how to move forward.

1 The authors of this report are John Mitchell, Jorrit Zwijnenburg and Rachida Dkhissi (OECD), James Lewis and Thomas Prendergast (ONS). The authors would like to thank Laura Caldwell and Sumit Dey-Chowdhury (ONS) for their important inputs and comments to this report.

2 For the project, the OECD and the ONS interviewed national account compilers from the respective NSIs in Australia, Canada, Italy, Ireland, France, Germany, Norway, the United Kingdom and the United States.
2. Background

5. The need to understand how COVID-19 may have affected the measurement of non-market output originates from the contribution of non-market output to differences in changes in Gross Domestic Product (GDP) across countries during the early stages of the pandemic. Figure 1 shows the differences in GDP growth rates in the second quarter of 2020 (when the first and most restrictive confinements were in place in many countries) across G7 countries. It shows marked differences, especially for the headline volume estimates. In some countries, substantial variation is also visible between the growth rates of volume and current price estimates of GDP.

Figure 1. Change in GDP, quarter-on-quarter, Quarter 2 2020, G7 economies, current price and volume estimates

![Graph showing GDP growth rates](image)

Source: OECD.

6. Even when considering different timings of restrictions across countries, i.e. by comparing the growth from the final quarter of 2019 to the second quarter of 2021 (see Figure 2), the difference in volume estimates of GDP still varies significantly, ranging from -3.9% in the Italy to +0.9% in the United States.
Figure 2. Cumulative change in GDP, Quarter 4 2019 to Quarter 2 2021, G7 economies, current price and volume estimates

Source: OECD.

7. Looking at the components of GDP, Dey-Chowdhury et al. (2021) identify that part of the difference in GDP growth arises from substantial variation in Government Final Consumption Expenditure (GFCE). GFCE predominantly consists of services that are provided by governments for free or at prices that are not economically significant, referred to as non-market services (2008 SNA §2.40). Across G7 economies, the UK observed the largest increase in current price GFCE in the second quarter of 2020, while recording the largest decrease in volume terms (see Figure 3). In the other G7 economies – where the GDP decreases were more modest – movements in current price and volume GFCE have been more closely matched.
8. The fall in volume GFCE in the UK, and to a lesser extent in France, in Quarter 2 2020 negatively affected GDP growth. Figure 4 shows the contribution of GFCE to GDP through 2020 and Quarter 2 2021. While the negative contribution for France and the United Kingdom in the second quarter is offset in the third and fourth quarter, the relatively larger variance of these contributions over time compared with other countries is clearly shown.\(^3\)

---

\(^3\) GFCE was not a significant contributor to the change in GDP for the UK or France over the whole period from the fourth quarter of 2019 to the fourth quarter of 2020. In fact, the contribution to the change in GDP from GFCE was similar across all G7 countries, ranging between 1.0 percentage-points (Japan) to -0.2 percentage-points (United States).
9. The pandemic had a similar impact on non-market services in most OECD\(^4\) countries, with many experiencing extended period of remote learning for schools and disruption to healthcare services. Consequently, variations as observed in Figures 3 and 4 may imply differences in the measurement of non-market output rather than in the impact of the pandemic. At a minimum, it raises the question of how countries have measured changes in the provision of non-market services in their economic measures in volume terms.

10. The composition of GFCE may vary across countries, dependent on how specific goods and services are provided. For instance, healthcare may account for a large proportion of GFCE in countries with public healthcare systems, but only a small proportion in those with mostly private systems. This report focuses on the output of the three industries\(^5\) most concentrated in GFCE, that is: Public administration, defence and social security services (ISIC section O), Education (ISIC section P) and Human health and social work activities (ISIC section Q).

11. It should be noted that most countries expect higher revisions than average for 2020. In some cases, this is because shifts in measurement practice produced

---

\(^4\) Organisation for Economic Cooperation and Development. The OECD consists of the following countries: Australia, Austria, Belgium, Canada, Chile, Colombia, Costa Rica, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom and United States.

\(^5\) According to the International Standard Industrial Classification (ISIC).
methodological splits between annual and quarterly national accounts. Even when methodological changes did not occur, differences between the indicators used for annual and quarterly compilation may have been exacerbated due to COVID. Due to this, a review of the non-market output ultimately recorded by countries once all annual data has been included may well prove useful to confirm that the trends observed above are still in place and to assist in separating differences due to methodology with differences due to availability of indicators.

3. Methods to measure non-market output

12. As non-market services are provided for free or at prices that are not economically significant, the measurement of their output poses unique challenges due to the lack of market prices. However, as noted by Schreyer (2010), regardless of whether services are provided by market or non-market units, “the measured volume of non-market services should be the same as the one for measurement of the volume of market services, and vice-versa, as long as the services are the same”. Therefore, different methods may be used to measure non-market output, ensuring that the results closely approximate those that would have been observed if the service were provided by the market.

13. That said, it is important to note that regardless of the method chosen, there also remains a level of conceptual ambiguity in measuring non-market services, which extends beyond the choice of methodology. In this regard, it is difficult to determine the exact level of output that is being produced for some specific types of non-market services, making it difficult to measure the change in quantity over time. Are firefighters producing more output if they attend more fires, or is their state of readiness production in itself? If it takes the same inputs to incarcerate 10 prisoners as 11, should production increase with this additional inmate? And what constitutes one unit of education output? Is it based on inputs, student numbers, or the quantity of learning conveyed?

14. Another conceptual challenge is how to assess changes in the quality of some non-market services. Many countries follow the guidance of the European System of Accounts (2010) which directs that estimates of non-market output not be adjusted for any change in quality. However, some non-European countries may try to capture this within their estimates. In this regard, the distinction between quality and quantity change is also not always very clear. For instance, are high graduation rates of school students the result of the quantity of teachers, creating lower teacher to student ratios or do they reflect the quality of the teaching occurring?

15. While this report focuses on the compilation methods rather than the conceptual questions, some of these questions have become identifiable points of difference in how countries assessed the impact of the pandemic on the delivery of non-market services.

16. The practical guidance for compiling non-market output is set out in key national accounts references, including; the 2008 System of National Accounts (SNA), the European System of Accounts 2010 (ESA) and the Eurostat Handbook on prices and volume measures in the national accounts (2016), as well as existing commentary such as the OECD handbook Towards Measuring the Volume Output of Education and Health Services (2010). This section briefly outlines current guidance for measuring current price and volume estimates of non-market output.

---

Current price estimates

17. The 2008 SNA defines market output in current price terms as “the value of goods and services sold at economically significant prices” (see 2008 SNA §6.99). For units only producing non-market output, the 2008 SNA suggests that output may be valued in current price terms as the sum of the total costs of production. These costs include intermediate consumption; compensation of employees; consumption of fixed capital; and other taxes (less subsidies) on production (see 2008 SNA §6.130). Following this guidance, there is little variation between countries in the methodology used for current price estimates of non-market output.

Volume Estimates

18. The measurement challenge is far greater when measuring output in volume terms. The typical approach for the volume measurement of market services involves the deflation of the output measure in current prices. In practice, deflators used are typically applied at industry section or sub-section level or at the level of specific types of output, such as components of Consumer Price Indices (CPIs) or Producer Price Indices (PPIs). This has the benefit of ensuring volume estimates are all in the same price base and changes in the quality of products detected through prices are captured in the volume of output.

19. The measurement of non-market output differs, as there is no explicit price from which to construct a deflator. Therefore, alternative methodologies must be sought. These can be grouped into one of four categories:

- deflation using output prices
- deflation using input prices
- direct output indicators
- direct input indicators

20. The first two categories are considered “indirect” methods as the volume estimate is created “indirectly” through first estimating the current price estimate and then deflating using a chosen proxy price index. The second two categories are considered “direct” methods as the indicator is used to directly move forward the volume estimate irrespective of the current price estimate.

Deflation using output prices

21. Deflation using output prices is the approach most similar to the conventional approach to measuring the volume of market services. This approach involves deflating the

---

7 This approach implies that for nominal non-market output, gross operating surplus is assumed to be equal to consumption of fixed capital (depreciation), thereby making net operating surplus equal to zero. In the special case of units that produce both non-market and market output, the non-market output component of such units is valued in current price terms as “the difference between the total costs of production minus the revenues from market output.”

8 However, quantity indicators, such as those based on employment data, may also be used for measuring market output, particularly for early quarterly estimates, additionally quantity indicators are often used for volume estimates of trade in goods.

9 For instance, see Roe (ONS, 2014) “Deflation of Short-term Output Indicators”.

Unclassified
output of non-market services using deflators constructed from price data associated with market output produced, such as components of CPI or PPI.

22. When applying this method, compilers should be mindful that the composition of services provided by the market and non-market sector might vary. For instance, hospital service providers in the market sector may provide a different range of elective surgery (such as cosmetic surgery and treatments with long waiting lists) than non-market providers. Therefore, the output deflator calculated for the market sector may not be representative of the services provided by the non-market sector. Similarly, growth in prices and costs for market services will be subject to market forces such as competition, whereas for non-market services cost growth will usually be influenced by budgets and government policies on efficiency. Therefore, the use of output prices is most suitable when the composition of market and non-market services are relatively similar, although even then, its use implies that the relationship between expenditure and volume is the same between the market and non-market sector.

**Deflation using input prices**

23. Deflation using input prices involves deflating the output of non-market services using deflators constructed from price data associated with the inputs used, predominantly labour inputs and intermediate consumption (goods and services consumed as inputs during production).

24. Using input prices may be more accurate than using output prices due to the level of disaggregation at which these prices can be applied. The limitation of this approach is that, by definition, it assumes that there is no change over time in productivity for non-market services. Consequently, the input price approach is generally considered inferior to the direct output indicators method discussed below. However, the use of input prices to deflate output may be appropriate where the variety of services provided is too great to enable groupings for relatively homogenous activity types that can be counted using direct output indicators.

**Direct output indicators**

25. At its simplest level, use of direct output indicators involves changes in the level of output volumes. It is driven by non-monetary indicators related to the service provision in question, independently of the service expenditure. The indicators used explicitly relate to the output, such as student numbers or medical appointments. The 2008 SNA (§15.122) states that the use of direct output indicators is the recommended approach for non-market services where the appropriate data is available.

26. Where a service provides a range of outputs of varying value, weights usually need to be applied. For instance, hospital services can vary greatly in value, with a major surgical operation requiring a greater weight in assessing the volume changes of healthcare services than a brief outpatient consultation. For non-market output, where there is an absence of economically significant prices to distinguish and weight different activities, costs may be used instead. In this regard, a common direct output indicator is the cost-weighted activity index, where growth in more costly activities has a greater effect on the quantity change in output although the overall impact is still dependent on the weight that is given to this activity. Due to this, compilers should be mindful of changes in the composition of services over time. This issue is particularly pertinent over the pandemic, where many services, such as elective surgery, were restricted, and new services, such as COVID-19 testing, were created.
Direct input indicators

27. Direct input indicators also use non-monetary indicators to assess the change in output, independently from its monetary level. However, unlike those focusing on specific outputs, these indicators focus on volumes of inputs, such as staff numbers or staff-hours.

28. As with direct output indicators, when input indicators refer to inputs of differing costs, such as employees at different pay bands, it may be appropriate to calculate a cost-weighted index by weighting staff numbers by the respective costs of different staff groups. It is important to note that the specific choice of input indicator could lead to significant differences. For example, as shown during the pandemic, the use of employee numbers and employee hours worked are both considered direct input indicators but may produce very different results in the situation of furloughed employees.¹⁰

29. As with other measurement approaches focused on inputs, this approach does not allow for changes in productivity over time. Therefore the use of output indicators is preferred.

Factors determining the choice of non-market output methodology

30. The choice of methodology will depend on the nature of the service provided. Chief among these differences is the distinction between those that are individually consumed, that is received by specific individuals (such as education and healthcare), and those that are collectively consumed, that is received simultaneously by a (section of) society as a whole (such as policing or national defence).

31. In general, the activities that define individually consumed non-market services, such as surgical procedures and school lessons, are easily identified. Where adequate data is readily available, direct output indicators are typically preferred in the key national accounts’ references mentioned earlier in this section. In contrast, identifying the discrete activities of collectively consumed non-market services and acquiring the data needed to produce a direct output indicator is often more difficult so input indicators may be needed for collectively consumed non-market services.

32. In practice, national statistics institutes (NSIs) are constrained in what methodologies they employ for non-market output by the timeliness and quality of available data. If more data becomes available over time, methods may change accordingly, with later estimates possibly relying on better methodologies than earlier estimates. This may also cause revisions, particularly when events affect the provision of non-market services (such as the COVID-19 pandemic).

Comparability of non-market output measures

33. The different methodological approaches may well bring about difficulties in comparability between non-market output measures across countries. For instance, input-based measures may not account for productivity improvements that output indicators

¹⁰ Furloughed employees includes those that are temporarily stood down from undertaking duties but are still receiving some or all of their standard pay. While payments are paid by the employer, during the pandemic employers with furloughed workers were usually the recipient of various similar government support schemes designed to maintain the employer-employee relationship, this includes the Kurzarbeit in Germany, the Chômage Partiel in France, Jobkeeper in Australia and the temporary wage subsidy scheme in Ireland.
would capture.\textsuperscript{11} Methodological choices may also generate different results depending on national institutional contexts – for instance, if a country’s education system allows for or incentivises part-time study, using number of students enrolled as an output indicator rather than number of students adjusted for full time equivalency may produce different results.

4. How National Statistical Institutes (NSIs) currently measure non-market output

34. In 2010, the OECD working paper \textit{Towards Measuring the Volume Output of Education and Health Services: A Handbook} provided an inventory of countries’ practices based on a survey conducted by the OECD and Eurostat. As the COVID-19 pandemic brought new challenges to the measurement of non-market output, Eurostat and the OECD conducted a new inventory based on a “Questionnaire on Price and Volume Measures for Collective Non-market Services, Health Services and Education Services during COVID-19”.

35. The results of these two surveys have been crucial in understanding the prevalence and nature of different non-market output methodologies. However, such surveys can only collect limited information, leaving further questions on how different methodologies have responded to the consequences of the COVID-19 pandemic. To answer these questions, the OECD and the UK Office for National Statistics (ONS) jointly conducted an information gathering exercise, engaging with eight National Statistics Institutes (NSIs) on the following topics:

- their basic methodology and data sources for measuring non-market output and government final consumption expenditure (GFCE)
- any differences in methodology and data sources between first and final estimates
- any adjustments applied to reflect the impact of the COVID-19 pandemic
- implications of the methodology used for interpreting changes in non-market output during the pandemic.

36. The surveys have shown that sum-of-costs methods are used across the board in both annual and quarterly national accounts to estimate current price output of non-market services. For volume estimates, all four methodologies described in Section 3 are used. Several countries report the use of multiple methods for components of a service category – for example, deflation of inputs may be used to measure the output of public education while deflation of outputs is used to measure the output of private education.\textsuperscript{12}

37. Data compiled by the OECD show that in the pre-pandemic period, differences in methodology were not a clear predictor of differences in volume GFCE growth rates. Broadly speaking, differences are small and it is difficult to distinguish the effect of differences in methodology compared with economic factors such as difference in the delivery mechanism of government services.

38. Similarly, differences in methodology were not a strong predictor of growth for combined ISIC sections O, P and Q (hereafter referred to as industries OPQ) in the pre-

\textsuperscript{11} An example in healthcare would be if the number of people employed was used as an input indicator, but technological improvements allowed for an increased number of treatments to be performed by a single doctor. If two countries had the same technology but used different output measurement techniques, their results may vary due to method rather than the economic reality.

\textsuperscript{12} For this reason, percentages as presented in this section may not sum to 100%.
pandemic period.\textsuperscript{13} Year-on-year growth rates in countries that made heavy use of direct output measurements showed slightly higher volatility than in countries predominantly relying on deflation using input prices, but once again, these differences are small and difficult to untangle from wider economic factors.

39. Trilateral discussions between the OECD, ONS and NSIs enabled a more detailed study of the practical application of the different methodologies for non-market services across countries. This section explores the themes uncovered from this work for each of the three main industries of the non-market sector. Tables 1-3 summarise the methods used to measure these industries across OECD countries.\textsuperscript{14}

**Public Administration and Defence (Section O)**

40. ISIC section O includes public administration, defence, law enforcement and other collective non-market services, which are overwhelmingly publicly funded and provided. As discussed in Section 3 the collective consumption aspect of most services in ISIC section O prevents the use of direct output volume measures for most services in this industry. Direct input indicators are commonly used by OECD members for volume measures of section O output, typically focusing on labour inputs. Use of data on hours worked or employee numbers are both common.

<table>
<thead>
<tr>
<th>Table 1. Public Administration and Defence (Section O)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input-based measures</strong></td>
</tr>
<tr>
<td><strong>Indirect methods</strong></td>
</tr>
<tr>
<td>Canada, Norway (ANA)</td>
</tr>
</tbody>
</table>

**Direct methods**

<table>
<thead>
<tr>
<th><strong>Direct input indicators</strong></th>
</tr>
</thead>
</table>

Note: Applies to both quarterly national accounts (QNA) and annual national accounts (ANA) unless specified.

*Assumed for output of collective non-market services, such as national defence or fire services. Although most of section O consists of these types of services, some aspects are not regarded as collectively consumed, such as subsidised cultural institutions.

Source: Eurostat and OECD.

41. In a minority of countries, indirect estimation is also used for ISIC section O, with volume output estimated by deflating output expenditure using input prices. Deflators used

\textsuperscript{13} Industries OPQ is the aggregate of ISIC sections O (Public administration and defence), P (Educational services) and Q (Human health and social care services).

\textsuperscript{14} Data is based on a Eurostat-OECD “Questionnaire on Price and Volume Measures for Collective Non-market Services, Health Services and Education Services during COVID-19”, which was self-reported by countries in 2020. For simplicity, countries are placed in the quadrant that reflects their predominant methodology used in quarterly and annual compilation. This does not take into consideration each country’s level or mix of private or public involvement in the delivery of these services.
include industry-specific deflators derived from national accounting systems and components of producer price indices (PPIs). For instance, the United Kingdom uses industry specific deflators extracted from national accounts to deflate output of policing and some other government services, while the United States deflates output of civilian government employees using wage data.

42. Direct output volume indicators are applied rarely. For instance, the United Kingdom measures elements of section O, including fire protection services, prisons, probation and legal aid services, using the cost-weighted activity approach with output indicators. However, a large majority of the services in ISIC section O in the United Kingdom have no suitable activity data available for the direct output methodology, and most of this industry is measured through indirect input approaches.

43. For both direct and indirect inputs approaches, data are often available for initial quarterly estimates and more mature estimates. However, for many countries, a short period of forecasting is required for producing preliminary outputs regardless of the methodology subsequently used.

**Education (Section P)**

44. As an individually consumed service, education output is more easily measured by direct output indicators. This is reflected in countries’ practices, the vast majority (81%) use direct output indicators. Furthermore, several countries use input-price deflation (16%) while a smaller number apply deflation of outputs (13%) and/or input indicators (10%). Many countries use indicators such as the number of pupils or students enrolled to measure volume change in education services in annual national accounts. Cost-weighted activity is commonly used, with different weights applied, to account for differences in costs among schooling types. Alternatively, direct output volume may be captured using pupil or student hours.

**Table 2. Education (Section P)**

<table>
<thead>
<tr>
<th></th>
<th>Input-based measures</th>
<th>Output-based measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Indirect methods</strong></td>
<td><strong>Input price deflation</strong></td>
<td><strong>Output price deflation</strong></td>
</tr>
<tr>
<td>Canada, Japan, South Korea, Colombia</td>
<td>United States</td>
<td></td>
</tr>
<tr>
<td><strong>Direct methods</strong></td>
<td><strong>Direct input indicators</strong></td>
<td><strong>Direct output indicators</strong></td>
</tr>
<tr>
<td>Canada, Ireland, Norway (QNA), Spain</td>
<td>Australia, Austria, Belgium, Chile, Czech Republic, Denmark, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Luxembourg, Mexico, Netherlands, New Zealand, Norway (ANA) Poland, Portugal, Slovak Republic, Slovenia, Sweden, South Africa, United Kingdom</td>
<td></td>
</tr>
</tbody>
</table>

Note: Applies to both quarterly national accounts (QNA) and annual national accounts (ANA) unless specified. Source: Eurostat and OECD.
45. Input cost deflation and direct input indicators, including the number of teachers employed or their hours worked, are also used for education. Often this is because the data is available more rapidly than the data for output-based measures. Combinations of approaches are also possible, for instance, Ireland uses direct output and direct input indicators whereby changes in student numbers have a greater weight but labour inputs are also a factor.

46. Countries using output indicators for their annual results will typically make quarterly estimates following an annual path based on the most recent benchmark data. In some cases, the quarterly path may vary across the year dependent on quarterly patterns. In many cases, enrolment data is only available annually and output is spread smoothly across the four quarters using a linear trend or other estimation technique.

Healthcare (Section Q)

47. As with education, healthcare is an individually consumed service. However, while in most countries education is mostly provided as a public service, substantial variation can be seen across OECD countries regarding the relative size of healthcare services provided by private and by public entities. In the United States, for instance, hospitals are private entities charging economically significant prices, and as such, their output is not considered part of non-market output.

48. Responses to the various surveys, coupled with information available publicly from NSIs and Eurostat, show that over a third of the countries (35%) use input prices to deflate output in order to generate the volume of health services. Almost as many countries use direct output indicators (32%), while a smaller group (23%) deflate using output prices. The latter group includes countries (10%) where healthcare is predominately delivered by the private sector, so market prices are more readily available. A further 19% of countries use direct input indicators.

Table 3. Healthcare (Section Q)

<table>
<thead>
<tr>
<th></th>
<th>Input-based measures</th>
<th>Output-based measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect methods</td>
<td><strong>Input price deflation</strong></td>
<td><strong>Output price deflation</strong></td>
</tr>
<tr>
<td></td>
<td>Austria, Canada, Chile, Colombia, Czech Republic,</td>
<td>Germany, Japan, Luxembourg, South</td>
</tr>
<tr>
<td></td>
<td>Denmark (QNA), South Korea, Latvia, Poland</td>
<td>Africa, United States</td>
</tr>
<tr>
<td>Direct methods</td>
<td><strong>Direct input indicators</strong></td>
<td><strong>Direct output indicators</strong></td>
</tr>
<tr>
<td></td>
<td>Canada, Ireland, Denmark (QNA), Mexico, New Zealand (QNA)</td>
<td>Australia, Belgium, Denmark (ANA), Finland,</td>
</tr>
<tr>
<td></td>
<td>, Norway (QNA), Slovak Republic, Spain</td>
<td>France, Hungary, Italy, Netherlands, New</td>
</tr>
<tr>
<td></td>
<td><strong>Note</strong>: Applies to both quarterly national accounts</td>
<td>Zealand (ÂNA), Norway (ANA), Portugal,</td>
</tr>
<tr>
<td></td>
<td>(QNA) and annual national accounts (ANA) unless specified</td>
<td>Slovenia, Sweden, United Kingdom</td>
</tr>
<tr>
<td></td>
<td><strong>Source</strong>: Eurostat and OECD</td>
<td></td>
</tr>
</tbody>
</table>

49. In annual national accounts, many OECD countries use direct output indicators to measure most aspects of non-market healthcare output. In most cases when output indicators are used, measurement of hospital output is derived from Diagnostic Related
Groups (DRGs) or something similar such as the UK’s Healthcare Resource Groups (HRGs).  

50. As DRG data is only available annually in most cases, there is substantial variation in the indicators used to measure hospital output in quarterly national accounts. Initial quarterly estimates are usually composed of more timely but less detailed direct output indicators, based on projections of the annual data. Alternatively, countries use input prices to deflate current price estimates. For instance, Australia has access to early DRG data at a less granular level of detail for use in initial estimates, while the United Kingdom employs a mixture of highly aggregated indicators spanning the main service components of hospitals, primary care and prescriptions for initial estimates. Input indicators such as employment levels are also used. Such data often has the advantage of being available for quarterly national accounts, allowing it to be used as a broad indicator for the level of production without providing the additional level of information such as the compositional mix or resources used.

**GFCE and non-market output**

51. It should be noted that while output of industries O, P and Q is often an important component of GFCE, these categories do not directly correspond to each other. This is because industry-level output may not only concern non-market output but also market output, such as driving lessons provided by private entities. Additionally, not all non-market output is captured in ISIC industries O, P and Q; for instance, some heavily subsidised cultural services fall in ISIC section R, also feeding into GFCE.

52. The market structure of all these industries varies substantially between countries – for example, the bulk of healthcare spending in countries with highly centralised public healthcare services such as Norway or the United Kingdom would be reflected in GFCE, while Germany, which has a largely private healthcare infrastructure but an extensive state-funded insurance mechanism, the majority of health spending may be recorded as social transfers in kind. In contrast, health spending in the heavily market-based United States is more likely to be part of private consumption expenditure (PCE). These differences imply that divergences between GFCE and section OPQ trends may be wider in some countries than in others.

53. Furthermore, as was the case for the estimates of non-market output, methods of calculating volume GFCE figures differ between countries. Many NSIs construct the portions of volume GFCE coming from industries O, P & Q with the same measures and data used to calculate the volume output of those industries. This is done, amongst others, in Canada, France and the United Kingdom. Other countries, such as Australia, deflate current price GFCE estimates to create their GFCE volume figures. As such, divergences may also arise between the trends associated with these measures due to methodological differences.

---

15 DRG (Diagnosis Related Groups) systems are used to classify treatments into medically meaningful and relatively homogenous groups. This provides a basis for applying differential cost weights to different treatments to produce a volume output measure.

16 While divergences may occur at the aggregate level on a quarterly bases, at a detail level, products are balanced during the standard compilation of Supply-Use Tables, which also benchmarks annual estimates.
5. How has COVID-19 affected measures of non-market output?

54. The events of 2020 brought unique challenges to the measurement of non-market output. Standard methods that were considered adequate for measuring non-market output prior to the pandemic required investigation to reflect shifts in the delivery of non-market services, such as the introduction of new treatments in healthcare or the switch to virtual education. As a result, some National Statistical Institutes (NSIs) implemented methodological changes, temporary adjustments, or incorporated additional information, such as data reflecting lower school attendance, to augment existing compilation methods. The extent to which NSIs adapted their methods and implemented adjustments varies substantially by country and industry. This section discusses methodological changes implemented by NSIs for Sections O, P and Q, before comparing published outputs and assessing the pandemic’s impact. It then contrasts the output of the aggregated non-market sector (industries OPQ combined) across a wider set of countries before discussing some general considerations related to compilation practices during 2020.

Public Administration and Defence (Section O) output measurement during COVID-19

55. As mentioned in Section 4, there is greater similarity between countries in the methods used to measure the output of section O than those used for sections P and Q, as almost all countries use direct input indicators to compile their volume estimates for this section. The COVID-19 pandemic did nonetheless raise measurement challenges for the output of section O. This included accounting for the different policy responses countries applied, which ranged from substantial layoffs of local government employees in the United States to furloughing public employees in France, while Australia reported that government employees were able to continue their work despite the restrictions.

56. The degree to which any reduction in output would automatically be captured by section O volume measures is largely dependent on the method used. Direct input indicators, such as the number of employees, or volume estimates derived using input prices, would not automatically show any reduction in the amount or cost of labour where public sector workers remained employed, even if the amount of work they were able to do was severely limited. Some countries, such as Canada and France, felt that this did not properly reflect economic reality, so they collected additional data on the reduction in actual hours worked to make adjustments to public administration output. For these countries, this led to a noticeable reduction in section O output. In Canada, for instance, volume output for public administration declined by 5.1% in the second quarter of 2020.17

57. However, it appears that most countries made no specific additional adjustment to section O output. This was either because any reduction in production was not deemed material or the indicator in use already reflected any declines. For instance, hours worked is used as direct input indicator in several countries, which would already automatically pick up the reduction caused by employees shifting to a reduced workload. In most countries, the disruption observed was only minimal, with some countries even showing small increases (e.g., Australia (0.7%), Norway (0.2%) and the United Kingdom (0.7%)).

58. These examples show that (with a few notable exceptions) in most countries output of section O was not significantly impacted by the pandemic; and that, by contrast with

---

17 Canada uses the North American Industrial Classification System, whereby Public administration is most similar to section O of the ISIC.
education and healthcare services, it was not a significant source for international differences in non-market output.

**Education (Section P) output measurement during COVID-19**

59. Almost all OECD countries – and all countries interviewed for this report – saw substantial disruption to normal education practices, with remote learning implemented for pupils for part of the pandemic (where possible).

60. Some countries reacted to this by making specific adjustments to education output reflecting these disruptions; this was done by collecting information on the impact of remote learning or changing the fundamental method in use. For instance, Norway continued using a direct input indicator (number of people employed in education), but they introduced adjustments for a short period of time to reflect the transition period to remote learning, lower capacity in primary schools and the cancellation of exams. Germany applied specific adjustments to output to reflect the complete shutdown of pre-primary education in Quarter 2 2020, the magnitude of which varied between sub-national units depending on local circumstances.

61. Some countries applied more aggregated adjustments to cover the reduction in output provided to pupils who had switched to remote learning. Italy and the United Kingdom both used survey data to investigate the reduction in the amount of learning materials delivered to students learning from home. A further adjustment was made by France and the United Kingdom to discount the quantity of education output due to a larger proportion of the education service being provided through parental input, which is outside the production boundary of the national accounts. France also made an adjustment reflecting a higher than normal rate of attrition observed in students attending school via remote learning relative to the level observed in a normal year.

62. These adjustments appear to have had a substantial effect on observed education output. The changes instigated by Norway resulted in a 7.9% decline in education output in the second quarter of 2020, while the United Kingdom sustained a 40.2% decline. Similarly, although France does not produce separate growth rates for section P, the overall fall in output of industries OPQ for France in the second quarter of 2020 was second only to the United Kingdom among G7 nations, with the education adjustments likely being an important driver for this.

63. The countries listed above acknowledged that the adjustments were put in place as it was perceived that the move to a remote learning environment led to a lower level of education services. However, this view was not shared across all NSIs. In this regard, the majority of NSIs in Europe did not apply such a correction, as Eurostat guidance suggested “services delivered remotely count as non-market output just as those delivered ‘physically’. There is no implicit change in the volume of the service delivered.” This was followed up with an additional clarification, specific for education, “where the services...”

---

18 The SNA2008 production boundary (that is, the economic activity that contributes to a country’s gross domestic product) excludes unpaid household services. This includes activities such as cleaning, cooking and the supervision of children which would be counted if undertaken in exchange for payment from a third party but are excluded if done for an individual’s own benefit. Therefore if the additional activity by parents is considered an extension of this “unpaid household services”, it should not be included as production. However, since the activity was above what is normally expected from parents, it could be considered that the parents were “volunteering” their services to assist in the school in the production of education services, in this case the activity would fall back within the SNA production boundary (see SNA2008 ¶ 29.157).
have shifted towards remote teaching and more homework, with all pupils engaged somehow, it seems reasonable to assume that output is more or less unchanged compared with a normal situation.”19 Several countries (Australia, Canada, Germany and Ireland), when interviewed, confirmed that this guidance was consistent with their own views.

64. Some of the countries that did not apply specific adjustments still recorded decreases, due to actual decreases observed in the direct input or output indicator used. For example, Canada, which uses the direct input indicator of hours worked, supplemented this indicator with information on hours worked from their Disaster and Catastrophic Events (DCE) survey. This was not a remote learning adjustment – in fact, Statistics Canada advised that “output in [the education] service industry and in public administration will be unaffected by the closures except insofar as the employees report reduced actual hours worked”.20 Rather, the survey was used to better reflect the widespread layoffs of school support staff when school buildings closed and students switched to remote learning. Overall, this yielded a 10.7% decline in the output of education services in the second quarter of 2020.

65. Australia, which was also subject to school closures and remote learning in this period (although not on such a large scale), also chose not to apply a remote learning adjustment. However, due to their use of the relatively stable pupil numbers’ series as a direct output indicator of education services, Australia observed a 0.2% growth in education output volume in the second quarter of 2020.

66. The output of ISIC section P is broader than government final consumption expenditure (GFCE). It includes private schooling and wider education services usually funded by household consumption. For instance, driving schools faced extended closures and despite accounting for only a small proportion of education, as noted by Norway, the complete reduction in activity had a significant effect on total education output. Dynamics such as these may contribute to greater divergences between GFCE measures and OPQ measures, emphasising the varying effects of COVID-19 on these two measures.

67. In conclusion, differences between countries in output growth for ISIC section P over 2020 was heavily influenced by whether adjustments have been applied to account for changes in education provision. Importantly, adjustments of this sort have not been applied by all NSIs, and where they have, they have not been applied consistently, leading to a wide range of results. Based on the available data and discussions with NSIs it appears that the largest falls in education output were in those countries that applied adjustments to reflect a perceived reduction in production due to remote learning. In countries where no adjustments were applied or no additional data incorporated, direct output methods (which are often based around student enrolment) typically show little change in response to the pandemic. In contrast, direct input indicators, based on teacher numbers, hours worked or wages, may have shown some reduction, but not to the extent observed for those countries where additional adjustments were undertaken.


20 For more information on Statistics Canada recording of COVID-19 in the national account see www150.statcan.gc.ca/n1/pub/13-605-x/2020001/article/00001-eng.htm.
Healthcare (Section Q) output measurement during COVID-19

68. The COVID-19 pandemic had a profound impact on healthcare services, with widespread cancellations of non-urgent appointments and procedures, resources redeployed to care for COVID-19 patients, large-scale additional expenditure on personal protective equipment (PPE), and the creation of new services such as COVID-19 testing and contact-tracing.

69. As outlined in Section 4, whereas several countries rely on direct input or output indicators for their annual estimates, most countries use projections, input price deflation methods or more aggregated indicators to generate quarterly estimates for healthcare. Due to the rapid changes in healthcare activity seen in early 2020, many of these existing quarterly methods were suddenly inadequate to represent the changes in the allocation of healthcare resources and the provision of healthcare services.

70. Many NSIs responded by using more timely activity data replacing their conventional data sources. This allowed the increased use of new COVID-19 related treatments at the expense of non-urgent healthcare services to be properly reflected in output. Norway, for instance, gained access to monthly DRG figures from hospitals, enabling timely accounting for changes in the case mix of hospital care. The United States, where the primary method for compiling healthcare services involves deflating by input prices, incorporated timely new sources such as credit card and lab data to reflect the impact of COVID-19.

71. Additionally, while some countries are able to incorporate case mix effects on a quarterly basis through the DRG system, this was further complicated by the likely higher expenses for COVID-19 treatments and the disruption of DRG recording mechanisms where staff were redeployed from their regular specialties to help with emergency care. Many countries, including Australia, France and Norway, assigned COVID-19 treatments to existing DRG categories reflecting severe respiratory disease. This allowed the existing categories to reflect the higher level of COVID-19 cases in hospital, but at the same time may result in those categories receiving a higher cost weight in the future, once the higher costs have fed through into base period weights. In some countries, such as Italy, adjustments were applied to existing DRG categories to reflect the higher cost of COVID-19 treatments.

72. However, some of the resources deployed to support these services came from outside the healthcare industry (e.g. the military) and thus without changes to the accounting systems to assign these resources to the changed activity, the costs of these services may not always be fully captured in ISIC section Q.

73. As with education, it is important to understand the difference between what is covered in healthcare output and what will feed into GFCE because in many countries, some healthcare services are provided by the market. Therefore, household consumption expenditure on healthcare services will also be affected by changes in the output of ISIC section Q. Indeed, early in the COVID-19 pandemic, privately funded non-elective treatment, including cosmetic surgery, was widely cancelled and many regular primary care services, such dentistry and ophthalmology, experienced significant disruption. This may be an important explanation why ISIC industries OPQ sometimes showed larger drops than reflected in GFCE.

74. Finally, healthcare may be an important component in the explanation of the differences between current price and volume estimates observed in non-market output. The large-scale purchases of PPE seen in many OECD countries, which made substantial contributions to current price increases in healthcare output and GFCE, did not directly

Unclassified
influence output indicators based on activity data. Therefore, several countries may have recorded large drops in the output of healthcare services in volume terms, while at the same time recording an increase in current price estimates, leading to high implied price deflators.

75. In summary, although a wide range of methodologies is used by NSIs to measure healthcare on a regular basis, the methodologies, along with additional data sources that NSIs have developed, appear to have picked up the fall in standard healthcare output resulting from the pandemic. However, a variety of adjustments and additional data sources have also been incorporated by NSIs. Some adjust for new services such as COVID-19 testing and some source new data that enable the high cost of treating COVID-19 patients to be accounted for immediately, rather than waiting until the annual benchmarks are compiled. While these adjustments have no doubt contributed to a more accurate representation of the production actually taking place during the pandemic, disentangling the often opposing effects of changes in regular indicators and the application of additional adjustments presents substantial challenges.

Comparing declines in Healthcare and Education with the impact of COVID-19

76. Despite the lack of consistency in methodologies and adjustments, there seems to be a relationship across countries between the change in healthcare and education output as reported by countries and the degree to which they have been impacted by the pandemic (such as in terms of death rates). Among the five countries for which individual industry section output data was available for healthcare (albeit with some differences in classification systems), Australia, Norway and the United Kingdom use direct output indicators, while Canada and the United States deflate output using input prices. Despite the different methodologies, the primary driver of the decline in output of healthcare services in volume terms appears to be the severity of the pandemic and the scale of the resulting disruption to healthcare services. Generally, where COVID-19 death rates are higher, indicating a more severe disruption to the healthcare system, healthcare output shows a sharper decline (see Figure 5). All these countries show a substantial fall in healthcare, with the United Kingdom recording the highest fall at 19.5%, followed by the United States (14.9%) and then Canada (13.4%). Even Australia, that experienced only 4 deaths per million by the end of June 2020, displayed an 8.1% fall in healthcare output in volume terms reflecting reductions in some healthcare services taken as a precautionary response.

---

21 Many different indicators could have been used as a metric to represent the severity of the impact of COVID-19 on a country. While many indexes exist, the components and weights applied to them could be considered quite subjective. Alternatively the prevalence of infection is heavily dependent on testing rates, which varied widely across countries in the initial stages of the pandemic. While deaths per million is also impacted by other influences (i.e. quality of health care or the general health and demographics of the population) it was deemed a more neutral indicator that cases per million. That said, COVID-19 deaths may be partly influenced by the overall level of health service provision and so the relationship between these two factors may not flow entirely in one direction. Finally the data source notes the difficulty in separating deaths from COVID with those deaths that occurred with COVID.
Figure 5. COVID-19 deaths per million (to June-2020) versus healthcare and education output, Quarter 2 2020

Sources: ABS, Statistics Canada, Statistics Norway, BEA, ONS, Our World in Data. Selection of countries based on those for which individual industries data was available.

77. Figure 5 shows a similar pattern for education, with those countries that experienced a more severe impact of the COVID-19 pandemic recording a larger fall in education services. In most cases, the falls in healthcare output are more severe, roughly 5-6 percentage-points larger than for the education industry. The exception is the United Kingdom, where in the second quarter of 2020, the education industry recorded a fall over 20 percentage-points larger than the healthcare industry, but also around 30 percentage points more than the other countries publishing education output. This difference is likely due to the adjustments implemented to reflect both the lower relative amount of education services provided to remote learners as well as the further adjustment related to the larger contribution from parents in providing these services (which is considered to be outside the production boundary). As mentioned, Norway and the United States also made adjustments to reflect a reduction in the quantity of education output due to remote learning. However, Figure 5 suggests that these adjustments were of a more subdued nature.

78. Overall, there is a positive correlation between the drop in healthcare and education output and the severity of the pandemic. However, this cannot explain all of the differences between countries, and it is clear that the output declines are also affected by differences in methodology and additional adjustments put in place.

Aggregate industries OPQ output during COVID-19

79. Since disaggregated industry data is not available for all countries on a quarterly basis, additional comparisons can only be done at the aggregated level. When the output of
industries OPQ is compared with the severity of the pandemic for a larger range of countries (see Figure 6), the relationship established in Figure 5 is not as obvious as when focusing on Section P and Q separately. To a certain extent this may be expected because the more aggregated numbers and additional countries reflect a variety of industry composition, differing methodological approaches, varying strictness of lockdown policies and other variables that affect the output of industry OPQ and the number of COVID deaths. Therefore, while there appears some connection between declining output of industry OPQ and higher COVID death rates, the position of some countries on the graph would appear to refute this, even when methodological approaches and adjustments are taken into consideration.  

Figure 6. COVID-19 deaths per million (to June-2020) versus aggregate industries OPQ output, Quarter 2 2020: selected countries

Source: Our World in Data, OECD.

22 For example, based on submissions to the 2020 Eurostat/OECD survey, both Poland and Slovakia reported generally identical quarterly methodology for non-market output (see Annex for detail). However, despite Poland recording the relatively higher impact of COVID-19 at 38.7 deaths per million, it recorded positive growth in industries OPQ, while Slovakia, which recorded 5.1 deaths per million, reported a decline in industries OPQ of 5.9% for the second quarter of 2020. As mentioned, there are many possible reasons for this, including the composition of industries OPQ in each country or the level of COVID-19-related restrictions. Alternatively, it may be that adjustments or additional data have been applied by Poland or Slovakia that may not have been reflected in the 2020 survey.
80. As no two countries are identical in the composition of their economies and of industries OPQ, in their experience of COVID-19 or in their methodology, it is not possible to correct for all possible variables that might impact on a country’s position on the graph. However, when reviewing the countries that show larger deviations from the trend, there are some common traits between countries with similar positions in Figure 6. Since almost all countries are using a similar method for section O, the differences are predominately in approaches to measuring education and healthcare.

81. Spain and Ireland show relatively small decreases in output of industry OPQ relative to other countries with similar COVID-19 deaths per million persons. Of those listed on the graph, these two are the only countries that use the direct input indicator ‘number of employees’ to estimate the quarterly output of both healthcare and education. This choice is important because in most countries, the level of full-time employment barely changed as employees either switched to remote working or were maintained as employees with the help of government support. Their reliance on this type of indicator, along with the fact neither Spain nor Ireland made any additional adjustments to the health or education industries to reflect any changes in these industries, may explain why they observed such a negligible decrease.

82. Belgium and Sweden are countries with a relatively high number of COVID-19 deaths per million persons, but both recorded relatively small decreases in education and healthcare output. Based on the Eurostat-OECD questionnaire (OECD, 2020) and publicly available methodological information, both use projections based on annual data to estimate their initial quarterly estimate of non-market healthcare and education. As forecasting would not be able to pick up an event such as the one experienced in the second quarter of 2020, and as no other adjustments were mentioned by either country, this may explain the smaller decreases in these two countries. In contrast, France and the United Kingdom recorded greater falls in the output of non-market services while experiencing similar COVID-19 death rates. These countries use direct output indicators for healthcare, which may have been able to better capture the impact on non-COVID-19 healthcare. At the same time, both countries use some quarterly projections based on annual estimates for education, similar to Belgium and Sweden. However, unlike Belgium and Sweden, France and the United Kingdom made specific adjustments to reflect the switch to remote learning. Italy also uses direct output methods and made some additional adjustments to reflect the lower number of pupils reached by remote learning and to account for the higher costs of COVID-19 patients. However, Italy showed a relatively small decline in output of industry OPQ when compared with the substantial COVID-19 impact experienced.\footnote{The decline in Italy is reflected in both the first Quarter and second Quarter of 2020 due to its earlier imposition of COVID-related restrictions compared with other European countries.}

83. Figures 7 and 8 show time series of OPQ output for OECD countries for which data were available, divided into those using direct output methods to measure both healthcare and education services, and those using alternative methods for at least one of these industries. Countries using direct output methods to measure both sections P and Q exhibited changes in OPQ ranging from +1.4% to -23.2% between the fourth quarter of 2019 and the second quarter of 2020, followed by rebounds beginning in the third quarter. For the countries using alternative methods, the change over this period was less marked, ranging from an increase of 1.5% to a decrease of 14% with all but three countries in the +1.5% to -3.4% range. This likely reflects the response of output indicators to the amount of services provided, especially in healthcare, while methodologies driven by input-factors may not have changed much over the pandemic. However, the wide variation among countries that do not use direct output for healthcare and education irrespective of any
adjustment is clear. For instance, Canada follows a pattern of substantial decline followed by a recovery, whereas Ireland shows only very small movements across quarters.

Figure 7. Volume change in output for aggregated sections OPQ, countries using direct output methods for both sections P and Q, Quarter 4 2019 – Quarter 2 2021, indexed to Quarter 4 2019 (=100)

Source: OECD.
Correlation between GFCE and Industries OPQ during COVID-19

84. In some of the countries studied, the effect of the COVID-19 pandemic on GFCE has been similar to that on Industries OPQ but in other countries there are notable differences. Figure 9 shows the change in these indicators for the countries that were interviewed. In Canada, France, Norway and the United Kingdom, volume GFCE and Industries OPQ output have tracked each other quite closely, all demonstrating correlation coefficients of more than 0.85 over 2020. Notably, these countries all have particularly high degrees of non-market healthcare and education output and use the same data sources to calculate both GFCE and Industries OPQ output. A weaker, but still present relationship is visible in Italy, which exhibits similar characteristics.

85. Far greater disparities are seen in Australia, Germany, Ireland and the United States, all of which exhibit equivalent correlation coefficients of less than 0.2 over 2020. Furthermore, in Australia and the United States the relationship between Industries OPQ output and GFCE is negatively correlated. This can be partially explained by differences in market structures – hospitals in the United States are privately run, whereas healthcare in Australia is structured around co-payments systems. However, methodological differences are also important - Australia calculates GFCE by deflating current price spending whereas its volume estimates for OPQ output primarily come from direct input and output indicators. Ireland calculates GFCE volume estimates based on a slightly different methodology.

---

24 Differences in Australia were also impacted by activity within the childcare industry switching from household to government expenditure due to policy shifts implemented during the pandemic.
different mixture of input volume indicators and deflated output than for its OPQ output measures.

Figure 9. Volume change in OPQ output and GFCE for selected countries Quarter 4 2019 – Quarter 2 2021, indexed to Quarter 4 2019 (=100)
As a result, as illustrated in Figure 9, while France and the United Kingdom showed large falls in output and GFCE, a number of countries, including Australia, Germany, Ireland and the United States, showed an increase in GFCE in the second quarter of 2020 in contrast to the movement seen in these countries for industries OPQ. This reinforces the need to look at both industries OPQ output as well as GFCE when considering the contribution of non-market output to Gross Domestic Product (GDP) estimates.

An interesting measurement consideration regarding the correlation between GFCE and Section O raised by some countries, relates to the impact of COVID-19 on cultural institutions or infrastructure run by governments and municipalities, such as libraries, museums and toll roads. Norway provided additional monetary resources to these institutions to cover for their reduced income over this period in relation to reduced use of their services. In this case, the input indicators used by Norway to measure these institutions’ output on an industry basis – number of people employed – would not have captured this reduction, while GFCE may have even increased due to the additional government expenditure on subsidies. Similarly, if a direct output indicator had been used to compile output of section O in volume terms, this would have led to a decline, in line with the reduced amount of books borrowed, tickets sold, or cars on the toll road. This is not a conceptual flaw, but rather a practical consideration when using different sources and methodologies to derive what is conceptually the same output.
6. Conclusions

88. Due to the need to value output without the availability of explicit prices, non-market output poses a unique challenge in national accounts measurement. Whereas all countries derive current price estimates via a sum-of-costs approach, a variety of methods is applied to derive volume estimates.

89. Analysis of the growth trends over the decade leading up to the pandemic indicates that the difference in results by different methods is limited in normal times. However, the COVID-19 pandemic resulted in large-scale disruption to the delivery of non-market services, particularly in the second quarter of 2020, which led to significant changes and increasing differences in measures of non-market output and GFCE across OECD countries. Some of these differences may be reduced once benchmarked data is available. However based on current estimates, it appears that a mixture of reasons explains the differences in measures of non-market output and GFCE. Some differences may relate to how much countries have been affected by the COVID-19 pandemic and to the fact that the range of services included in non-market output and GFCE differs between countries; but methodological differences, both existing and in response to the pandemic, have also contributed. As this may have important consequences on the cross-country comparability of important macroeconomic indicators (including GDP), this report analysed to what extent methodological differences may have driven differences in countries’ measures of non-market output and GFCE. This was done by looking at the three main industries that predominantly make up non-market output and GFCE: public administration and defence, education and healthcare.

90. The effect of methodological choices on comparisons of output of these services differs somewhat for each of these industries:

- For public administration and defence, input-based methods predominate and the effect of COVID-19 on the regular delivery of these services is deemed to have been minor across all countries. Changes to methods and new adjustments were also relatively limited and did not lead to large differences across countries.

- For education, although methodological differences exist between countries, most countries use direct output indicators. While lower output in education appears linked to the severity of the pandemic, the larger decreases observed appear to be driven mainly by the varied adjustments that NSIs applied to account for any perceived COVID-19-related reduction, due to the educational services being delivered remotely. A decline in output of market providers also played a small but not insignificant part.

- For healthcare, a wider range of regular methodologies is applied across countries. However, despite the variety of methods combined with additional data and adjustments to account for pandemic effects, the differences in output generally show alignment with the relative scale of the impact of the pandemic. Furthermore, it is important to note that while in some countries the industry-level output data is used within GFCE calculations, ensuring closer correlation, in others GFCE and section Q output are compiled using different data.

91. Methodological differences, including those applied temporarily, are clearly important in understanding differences between countries in non-market output, especially between those that reported similar impacts from COVID-19. Almost every country that primarily used direct output indicators for healthcare and education consistently showed a fall in output of these industries during the first wave of the pandemic, although the scale of this fall varied widely depending on factors such as the application of additional
methodological adjustments. On the other hand, countries that used deflated output or direct input indicators as their basic method for non-market output showed greater variations in the reported change in non-market output over the pandemic.

92. Despite these differences, there is a clear relationship between the scale of the impact of the pandemic measured through mortality and the size of the fall in non-market output. This indicates some degree of comparability remains across despite the differences in methodology, and importantly provides some reassurance that broader comparisons between countries of GDP are not obscured by these methodological differences. Rather, comparisons of GDP growth, particularly the quarter-on-quarter changes observed over the peak of the pandemic, or where conclusions are being drawn from relatively small international differences in GDP growth, should consider the potential consequences of different non-market output methodologies.

93. These methodological differences are likely to continue to affect comparisons of non-market output for some time to come, with the pandemic having lasting effects on non-market services. For instance, the productivity of healthcare is likely to be negatively affected due to the additional resources required for infection control measures such as personal protective equipment and social distancing in healthcare settings. Direct output-based measures may be negatively affected by such changes, but input-based measures may increase. Likewise, input-based measures will respond to increases in inputs providing greater contingent capacity for potential future waves of infections, whereas direct output-based measures may not.

94. To assist international comparisons of GDP over the pandemic period and subsequent economic recovery, or any economically volatile period, it is important that NSIs publish detailed metadata explaining both their regular methods for non-market output and any adjustments or additional data sources incorporated to account for crisis-induced changes. The following are few examples of material released by NSIs over the course of 2020.

- Canada: Recording COVID-19 measures in the national accounts.
- France: Detailed methodological notes on the second quarter of 2020.
- The United Kingdom: Coronavirus and the impact on measures of UK government education output: March 2020 to February 2021.

95. With countries often using different methodologies, interpretation, transparency and analytical potential would also be aided by NSIs publishing industry-level data on a quarterly basis, in order that the effects of methodologies, which are often industry specific, can be better understood.

96. Ideally, countries should continue to move to closer alignment for their standard compilation methodologies. Convergence on the use of direct output indicators for individually consumed services, as recommended by the SNA, would at least provide a similar starting point across countries for healthcare and education services – although whether this is possible will depend on mechanisms for delivering education and healthcare in each country as well as on data availability. Additionally, the international statistical community should continue to discuss and refine some of the concepts around the production of non-market services in order to ensure greater consistency between countries.

Unclassified
The impact on the quantity of education being produced when students are learning remotely compared with learning in a physical classroom is a clear example where international agreement would likely benefit cross-country comparability.

97. While there are no definitive answers, as shown by the debate this year, there is certainly room for more discussion on some of these concepts. There must be a continued desire to achieve general agreement on some of the issues, both conceptual and methodological, even if it requires a level of compromise. While macro-economic indicators like GFCE and GDP need to reflect the actual economy in order to remain correlated to other macro-economic outcomes like unemployment and prices, the ability to compare across countries should always be aimed for.
References


Roe, S. (2014) Deflation of Short-term Output Indicators, ONS.


## Annex A. Methods normally used to measure non-market output, by country

<table>
<thead>
<tr>
<th>Country</th>
<th>Public Administration</th>
<th>Education</th>
<th>Healthcare</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quarterly</td>
<td>Annual</td>
<td>Quarterly</td>
</tr>
<tr>
<td>Australia</td>
<td>Direct input indicators (hours worked)</td>
<td>Direct input indicators (hours worked)</td>
<td>Projection based on annual direct output indicators</td>
</tr>
<tr>
<td>Austria</td>
<td>Input-based methods(^{25})</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
</tr>
<tr>
<td>Belgium</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
</tr>
<tr>
<td>Canada</td>
<td>Direct input indicators (CoE, measured by hours worked) and input price deflation</td>
<td>Direct input indicators (CoE, measured by hours worked) and input price deflation</td>
<td>Direct input indicators (CoE, measured by hours worked) and input price deflation</td>
</tr>
<tr>
<td>Chile</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
</tr>
<tr>
<td>Colombia</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Input price deflation (public), direct output indicators (private; number of pupils enrolled)</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
</tr>
<tr>
<td>Denmark</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
</tr>
</tbody>
</table>

\(^{25}\)“Input-based measures” can consist of direct input indicators, input price deflation, or combinations of the two. Collective services are very often measured through direct input indicators, such as hours worked or number employed.
<table>
<thead>
<tr>
<th>Country</th>
<th>Input-based methods</th>
<th>Input-based methods</th>
<th>Projection based on annual direct output indicators</th>
<th>Direct output indicators (number of students)</th>
<th>Projection based on annual direct output indicators</th>
<th>Direct output indicators (DRG index)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>Direct input indicators</td>
<td>Direct input indicators</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (number of students)</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (DRG index)</td>
</tr>
<tr>
<td>Germany</td>
<td>Direct input indicators</td>
<td>Direct input indicators</td>
<td>Direct output indicator</td>
<td>Direct output indicator</td>
<td>Output price indicator</td>
<td>Output price indicator</td>
</tr>
<tr>
<td>Hungary</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (number of students)</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (DRG index)</td>
</tr>
<tr>
<td>Ireland</td>
<td>Direct input indicators (number employed)</td>
<td>Direct input indicators (number employed)</td>
<td>Direct output and direct input indicators (number of students and teachers)</td>
<td>Direct output and direct input indicators (number of students and teachers)</td>
<td>Direct input indicators</td>
<td>Direct input indicators (number employed in healthcare)</td>
</tr>
<tr>
<td>Italy</td>
<td>Direct input indicators</td>
<td>Direct input indicators</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (number of students)</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (DRG index)</td>
</tr>
<tr>
<td>Japan</td>
<td>Direct input indicators (employee compensation)</td>
<td>Direct input indicators (employee compensation)</td>
<td>Input price deflation</td>
<td>Input price deflation</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
</tr>
<tr>
<td>Latvia</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (number of students)</td>
<td>Input price deflation</td>
<td>Input price deflation</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (student hours)</td>
<td>[Market]</td>
<td>[Market]</td>
</tr>
<tr>
<td>Mexico</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Direct output indicators (public; number of students), output price deflation (private)</td>
<td>Direct output indicators (public; number of students), output price deflation (private)</td>
<td>Direct input indicators (public), output price deflation (private)</td>
<td>Direct input indicators (public), output price deflation (private)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Input-based methods</td>
<td>Input-based methods</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (number of students)</td>
<td>[Market]</td>
<td>Direct output indicators (volume index based on ICDs by age and discharge numbers)</td>
</tr>
<tr>
<td>Country</td>
<td>Methodology</td>
<td>Input Price Deflation</td>
<td>Output Price Deflation</td>
<td>Direct Output Indicators</td>
<td>Direct Input Indicators</td>
<td>Direct Output Indicators – Main Details</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------------</td>
<td>-----------------------</td>
<td>------------------------</td>
<td>--------------------------</td>
<td>-------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (student hours)</td>
<td>Direct input indicators (public), output price deflation (private)</td>
<td>Direct output indicators – Composite index of DRG, patient discharge and bed night data (public), deflation of outputs (private)</td>
</tr>
<tr>
<td>Norway</td>
<td>Direct input indicators (number employed)</td>
<td>Input price deflation</td>
<td>Direct output indicators (student hours)</td>
<td>Direct input indicators</td>
<td>Direct output indicators (DRG index)</td>
<td></td>
</tr>
<tr>
<td>Poland</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (student hours)</td>
<td>Input price deflation</td>
<td>Input price deflation</td>
</tr>
<tr>
<td>Portugal</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (number of students)</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (DRG index)</td>
</tr>
<tr>
<td>Slovakia</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (number of students)</td>
<td>Input price deflation</td>
<td>Input price deflation</td>
</tr>
<tr>
<td>Slovenia</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (number of students)</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (DRG index)</td>
</tr>
<tr>
<td>Spain</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (number of students)</td>
<td>Direct input indicators (number employed in education)</td>
<td>Direct input indicators (number employed in healthcare)</td>
</tr>
<tr>
<td>Sweden</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (student hours)</td>
<td>Projection based on annual direct output indicators</td>
<td>Direct output indicators (DRG index)</td>
</tr>
<tr>
<td>South Africa</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (number of students)</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
</tr>
<tr>
<td>South Korea</td>
<td>Input-based methods</td>
<td></td>
<td></td>
<td>Direct output indicators (student hours)</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Input price deflation and projection of annual output indicators</td>
<td>Input price deflation (public), output price deflation (private)</td>
<td>Input price deflation (public), output price deflation (private)</td>
<td>Direct output indicators (number of students)</td>
<td>Input price deflation (public), output price deflation (private)</td>
<td>Direct output indicators (HRG index)</td>
</tr>
<tr>
<td>United States</td>
<td>Direct input indicators (employment)</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
<td>Output price deflation</td>
</tr>
</tbody>
</table>