Committee for the Coordination of Statistical Activities **SA/2017/7**

Thirtieth session **19 September 2017**

Muscat, 27 – 28 September 2017

Item 5 of the provisional agenda

==========

Nowcasting techniques to improve timeliness

Prepared by OECD

1. Introduction

1. Like statistical quality frameworks produced by many national statistics institutes and other international organisations, the OECD's Quality Framework and Guidelines for OECD Statistical Activities recognises a number of dimensions by which quality can be assessed. The OECD's framework includes seven dimensions: *relevance; accuracy; credibility; timeliness; accessibility*; *interpretability; and coherence.*

2. An overarching principle in the framework is the interrelated nature of the various dimensions, with improvements in one often directly improving another, for example improvements in any of relevance, accuracy, timeliness, accessibility, interpretability, coherence can work to improve credibility. But implicit in the framework is a recognition that the dimensions may also be competing, which reinforces the need to look at all dimensions of quality.

3. Perhaps chief in this regard, and a central issue, for statistical offices throughout the world is the potential tension between accuracy and timeliness. Recent years, and especially in the wake of the recent financial crisis, have heightened the need of policy makers for more timely data from statistical offices across nearly all domains of the statistical information system, and especially in the areas of macroeconomic and income distribution statistics.

4. Further pressure in this regard has arrived through the proliferation of private sector data providers who are beginning to occupy 'data spaces' through the provision of more timely measures that purport to measure the reference indicator or, more commonly, provide a useful proxy – for example sentiment indicators of businesses used as signals of economic growth. Increasing availability of Big Data sources, for example Google Trends, are also creating new pathways for occupying these spaces.

5. Whilst it is probable that in many cases these alternative measures would not satisfy the standards typically used by statistical offices to assess quality (as defined above), especially when weighing up accuracy or relevance versus timeliness, it is clear that they are gaining increasing currency. This is certainly true among users such as the media and financial markets but it is also increasingly the case among policy makers.

6. One question, or rather dilemma, this poses for producers of official statistics relates to the relative importance currently put in any trade-offs between accuracy and timeliness. In some statistical areas at least, it appears that users are currently settling for relatively higher importance for timeliness.

7. That is not necessarily to say that the response should be to recalibrate the relative importance attached to these two dimensions of quality by official statistical agencies to match those currently requested by users. Where unofficial sources are clearly not fit-for-purpose, official statistics agencies should be clear on this and adopting similar measures of quality used by unofficial statistics providers is likely to impact significantly on the credibility of statistical offices, including in areas where timeliness and accuracy are of the highest quality.

8. But equally, it is not to say that official statistical agencies should not reflect on whether more could be done to better meet the needs of users in this regard. Among many ways, this could be done through an examination of: *compilation methods*, where new technologies may provide scope for improved timeliness without any impact on quality – for example by improving internal software systems, automated data collection (e.g. in consumer price collection) etc.; *survey approaches*, where, technological advancements may have improved the ability of firms to respond in a more timely fashion, or indeed to

provide data for a more recent reference period; and also through a greater adoption of '*nowcasting*' techniques. This issues paper provides examples of nowcasting approaches used within the OECD Statistics Directorate as a way of motivating wider implementation at the national level, even if only as inputs into larger geographic zone totals.

2. What is nowcasting?

9. The draft UN *Handbook on Rapid Estimates*, circulated for global consultation in October 2016, provides a definition of nowcasting and a detailed description of the various statistical approaches that can be used to create nowcast estimates (summarised below). The Handbook makes a distinction between Flash and Nowcast estimates by defining the former as '*normally calculated on the basis of incomplete data, however produced using the same statistical or econometric model as for regular estimate'* and the latter as '*A very early estimate produced for an economic variable of interest over the most recent reference period calculated on the basis of incomplete data using a statistical or econometric model different from the one used for regular estimates'*.

10. Because the distinction between the two concepts, as defined above, can be blurred in practice, for ease of exposition, the concept of nowcasting used in this note is broader and includes both nowcast and flash estimation approaches as currently defined in the UN Handbook. This broader concept also includes estimation approaches to disaggregate data where a more aggregated statistic may be available, including for *recent* periods that may not be the *most recent* reference period. For example if disaggregated statistics are not available say at year T-3, this paper defines statistical methods to estimate the components in T-3 as also being nowcast. In this sense the reference to '*recent*' precludes methods to estimate historic data from the definition, although in practice the statistical approaches are likely to be similar.

11. In practice and in theory there are numerous statistical methods used to nowcast, too many to mention in this short note but, for reference, the UN Handbook attempts to provide as exhaustive an inventory as possible. Methods vary depending on the nature of the estimation required and the availability nature of other complementary data, ranging from time-series models that resemble pure forecast approaches, to very simple methods that extrapolate fixed relationships in populations to larger more timely aggregates or through the use of proxy variables, particularly for measures of growth.

12. A useful additional distinction can be made here by differentiating between those approaches that are essentially 'unbounded' i.e. share characteristics with traditional forecasting approaches and those that are 'bounded' i.e. are estimated within a set of known constraints, for example estimating components of a known aggregate, as the latter category can be expected to have smaller margins of error.

13. To meet user demands, both types of approach are adopted within the OECD Statistics Directorate but, in general, unbounded approaches are restricted to estimates of missing countries needed in zone totals, such as the OECD and G20 zones, and only then for those countries whose weight in the zone total is relatively small. Importantly however, and as a reflection of their unbounded nature, the OECD does not publish individual estimates for the countries in question.

3. Examples of nowcasting in the OECD Statistics Directorate

14. Like other international organisations, as a general rule, in its role as provider of internationally comparable statistics, the OECD Statistics Directorate does not typically engage in the production of nowcast estimates. For those countries whose timeliness may be an issue for the OECD's dissemination of statistics, the OECD's role, through its Committee on Statistics and Statistical Policy and Working Parties, is to advise and provide a platform for countries to share their experiences in a way that motivates the development of official data in countries that may have longer time lags – be that through identifying improvements in compilation methods, advances in survey design, identifying new statistical sources, or indeed through the advocacy of nowcasting approaches.

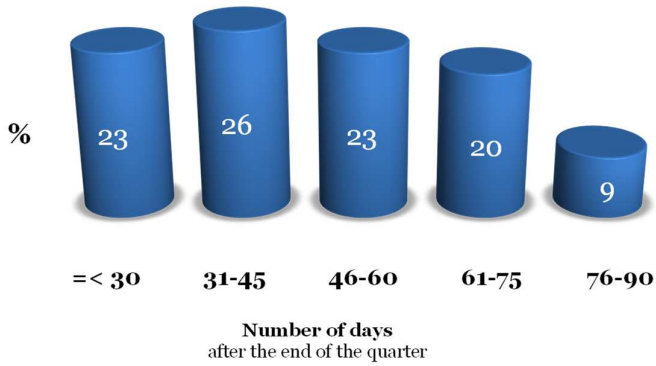
15. However, these mechanisms can often take time to deliver results and in some instances, in particular where there is strong user demand, the OECD may estimate data for missing countries. The following section provides an overview of the main areas where the OECD engages in, or is exploring the use of, nowcasting approaches, together with the methodologies and dissemination strategies applied.

3.1. Unbounded Nowcasting

16. The main area where unbounded nowcasting occurs is in the production of the OECD's quarterly GDP estimates of zone totals for the OECD and G20. The OECD's target here is to produce estimates for the OECD zone total 50 days after the reference period. The difficulty however is that not all countries produce official estimates in a sufficiently timely manner to meet this schedule (Figure 1).

17. For European economies unable to meet this schedule, nowcast estimates are produced by Eurostat in order to derive the European Union aggregate, and these estimates are used directly by the OECD as part of the Eurostat-OECD cooperation and commitment to reduce country response burdens. For non-EU OECD and G20 countries, the OECD creates its own nowcast estimates: Canada (61 days), India (67), Iceland (70), Turkey (73), South Africa (74), New Zealand (76), Argentina (82), Russian Federation (91), and Saudi Arabia (115).

Figure 1. OECD and Key Partner Country release lags for Quarterly GDP Estimates, %



18. The nowcasting approaches vary by country, partly reflecting the availability of complementary data sources but also partly reflecting the contribution (weight) each country makes to the overall zone total.

19. For Canada where monthly estimates of GDP are produced, two months of data for the relevant quarter are usually available and so the OECD generates estimates only for the missing month. The method projects the cumulative annual growth rate (month on same month in the previous year) for the missing month. In quarters where forecasts are produced by the OECD Economics Department for its Economic Outlook projections, these are used instead.

20. For New Zealand and Iceland, which, combined, have a very small weight (less than 0.5%) in the OECD zone total, the nowcasting approach assumes that they grow in line with growth in other OECD countries.

21. For Turkey, a more sophisticated modelling approach based on GDP trend-restored estimates from the OECD's Composite Leading Indicator, together with measures from the index of industrial production and consumer confidence, is used. For the latest quarter, the model estimated year-on-year growth of 1.8% growth, which compared favourably with the eventual outturn estimate of 2.1%.

22. For India, South Africa, Argentina and Russia, with a combined weight of less than 15% of the G20 total, the approach currently applied imposes growth observed in other G20 countries. For Saudi Arabia, with the longest time-lag in the G20, no estimates are currently generated and, so, the OECD's G20 press release is de facto a G19 release but the OECD is exploring nowcasting approaches.

23. As noted earlier, in all cases, estimations for the countries are not disseminated by the OECD but are only used in the construction of zone totals.

3.2.Bounded Nowcasting

3.2.1 Trade in Value-Added

24. The OECD-WTO TiVA database has helped to provide a new prism through which international trade, and, in particular, international fragmentation of production can be viewed. However, the underlying national supply-use and input-output tables required to produce TiVA estimates are typically not available until at best, two to three years after the reference periods, and, in addition, involve a lengthy process of integration within a coherent global accounting framework.

25. This necessarily means that current benchmark TiVA estimates are only available with a lag (currently 6 years) – currently covering the period 1995-2011. While TiVA itself has helped to motivate the development of more timely national Supply-Use and Input-Output tables, which will in turn help to reduce the lag, it will be some time before these developments can be fully integrated within the TiVA statistical information system. To tackle this issue, the OECD has developed a procedure and method to nowcast national input-output tables using more timely official national account data, coupled with timely trade statistics, to develop, nowcast inter-country input-output tables and in turn, nowcast, TiVA estimates (currently covering the period 2012-2014).

26. This section summarily describes the approach used by the OECD to develop nowcast TiVA estimates. A full description of the approach, together with the detailed adjustments made for each country is available at http://www.oecd.org/std/its/tiva-nowcast-methodology.pdf.

27. National supply-use and input-output tables provide detailed information on the inter-relations and inter-dependencies between industries, and indeed between industries and consumers. However, as noted, these tables are generally only available sometime after the

reference period. Nevertheless many countries produce sufficient statistics for nowcast TiVA estimates to be derived. All countries for example produce more timely estimates of GDP (usually around one to two quarters after the reference period), and its core components, such as value-added and output by industry and totals for categories of final demand, as well as estimates of imports and exports. Countries also typically produce very detailed and very timely data relating to merchandise trade, including by partner country. In this regard, the key piece of the information jigsaw required to produce nowcast national input-output tables, is intermediate consumption by industry broken down by product, including by whether those products were produced domestically or abroad.

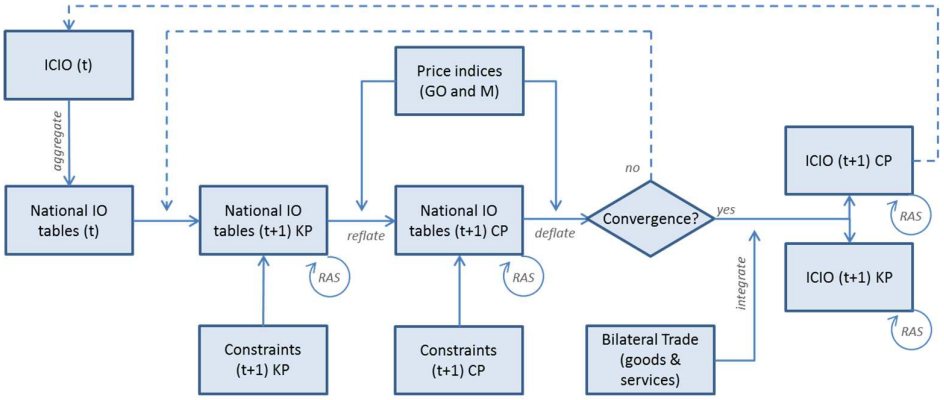
28. Although the OECD's nowcast TiVA estimates only show estimates in current prices, in order to maximise the use of official data in their construction, and also to strengthen the robustness of the underlying assumptions used to generate estimates, the approach uses available official statistics in both volume (constant) and current price terms. It does this through an iterative procedure that begins by assuming that intermediate consumption in year T+1, for given unit of output (where official timely statistics are typically available) is equivalent to the ratios observed in year T, *in volume terms.*

29. As such the initial assumption is that the production technology between the two periods has not changed: i.e. the same quantity of inputs is used to produce the same quantity of output, which avoids potential distortions introduced by differential price movements that may arise when intermediate consumption relationships are estimated only in current prices. Similar approaches are used to estimate components of final domestic demand – i.e. assuming, in the first instance, that consumption patterns are unchanged in volume terms. For imports and exports, merchandise trade data are available at a very detailed product level and, so, estimating consumption patterns requires few assumptions. At this stage all information required for a national input-output table are available and these estimates are made coherent (balanced) using the RAS balancing technique.

30. However, recognising that differential price movements may also induce changes in technology and in sourcing patterns, the nowcast approach also embodies a current price view of production, achieved through a reflation of volume estimates of intermediate consumption (with deflators derived from official available national data). The derivation of the current price tables from the initial volume estimates, through reflation, generates estimates that may not align with official current price estimates of value-added or final demand published by national statistical offices, and so an iterative procedure begins whereby the current price estimates are balanced and constrained to official estimates of value-added and key national accounts components before being deflated to derive new volume estimates, which are then constrained and balanced, in turn, to official volume estimates of output, value-added and components of final demand before being reflated again. The process continues until both current price and volume estimates converge to official published estimates of value-added and output by industry, trade and components of final demand. In this respect an important strength of this 'dual', 'simultaneous' approach is that it maximises, as input, all available national data – in volume terms and current prices.

31. With this approach, estimates of input-output tables can also be generated for periods T+2 and T+3, depending on available data, using the previous year's nowcast (or official) supply-use or input-output table as the starting point for production functions or consumption patterns. This also helps to avoid introducing statistical volatility in time-series. A simple schematic of the approach is illustrated in Figure 2 below.

Figure 2. The iterative nowcasting procedure for TiVA



32. The approach has been validated by applying the methodology to earlier years (1995-2011) where national supply-use and input-output tables are available and comparing the nowcast results generated with the equivalent benchmark TiVA estimates. Nevertheless, to reinforce the point that nowcast estimates necessarily require a greater degree of assumptions than the benchmark estimates, the nowcast TiVA database provides only a subset of the variables available in the benchmark dataset and the underlying input-output tables are only made available to users on request and only then to create derivative indicators.

33. The lower level of detail also reflects the fact that in some economies not all of the necessary constraining data – value-added, output, and components of GDP – are available in the nowcast years at a sufficient level of detail, especially with regards to volume measures of output or price indices to derive volume measures; where nowcast estimates are also necessarily derived.

3.2.1 Distribution of income, consumption and wealth

34. The OECD and Eurostat launched a joint Expert Group on Disparities in the National Accounts in 2011 to develop methodology to compile measures of the distribution of income, consumption and wealth across household groups, using a combination of detailed micro data sources and national accounts concepts and aggregates.

35. However the underlying micro data sources are typically only available significantly later than annual data on the household sector, which are usually available around nine months after the reference period. As such the OECD has begun to explore a number of options to derive nowcast distributional estimates that could be adopted by countries. Three broad options have been scoped:

- *A top down approach*: based on disaggregated and available national accounts components. This approach captures changes in distributions through aggregation effects. So for example if the share of net property income in overall income increases in T+1, and the share of the top quintile of households is higher in this category in T, the rising share of net property income will also result, all other things equal, in higher shares of overall income for the top quintile in T+1. The challenge here is to estimate the appropriate shares of each quintile within each category of income, consumption and wealth that should be used to extrapolate into period T+1. The simplest approach is to take shares in the previous period but for countries with longer time series more elaborate approaches could be used, for example using time-series approaches.

- *A bottom-up approach*: This approach extrapolates micro-data, using for example time-series approaches, which are subsequently aligned to the adjusted national accounts equivalents.
- *A meso approach*: This approach capitalises on the possible availability of data on more aggregated groups of households that can provide the basis for estimating quintile distributions.

36. Of the three options the OECD has been able to derive illustrative estimates for Australia and the Netherlands, where longer time series were available, for the top-down approach. These examined a number of options to estimate the shares of each quintile in each category of income, consumption and wealth, and while all options generally produced comparable results, they did generate differences.

4. Concluding remarks

37. Demand for more timely official statistics data has never been greater. Increasingly some of this demand is being met by alternative providers of statistics, creating risks that official statistics providers may become marginalised, at least with regards to those policy makers requiring very timely data. A number of approaches could be adopted by national statistics offices to respond to these demands and risks.

38. Improvements in compilation practices are one such area, especially given technological advancements such as digitised data collection. Improvements in survey design and in particular survey questions may provide another avenue. Firms today have access to digitised information in a way that should have significantly reduced response burdens since the surveys were first designed, this may mean that response times could be shortened with respect to days after the reference period.

39. A third avenue is through nowcasting. As shown with the examples above, the OECD's response to increased user demand is to increasingly scope and use nowcast approaches. Importantly, where estimates are nowcast, the OECD's dissemination strategy stresses this, and indeed, in most cases, only presents nowcast results indirectly, within a higher level aggregate, such as a zone-total, or as a derivative statistic, for example the import content of a country's exports within the nowcast TiVA database.

40. These innovations have been well received by users. However, while the estimates generated by the OECD are of acceptable quality, it is more likely than not, that comparable estimates produced at the national level will be of better quality. The OECD's work on distributions for example revealed differences depending on the estimation method used and national institutes are certainly better placed to determine the most appropriate method. Similar situations are likely to exist in other international statistics agencies.

41. That is not to say that national statistics agencies should necessarily produce nowcast estimates to fill these gaps and present these estimates as official statistics. As noted above the quality of statistics produced by national agencies depends on a number of factors. Improved timeliness may come at a cost of reduced accuracy with a corresponding negative impact on credibility. However, there does appear to be scope to consider the development of 'nowcast' estimates that could be used by international statistics agencies subject to certain conditions of use, such as those currently adopted by the OECD.

42. Such an exercise may also have direct benefits for national statistics agencies. One outcome of the exercise could be that the nowcast estimates are subsequently released as official

statistics, as a sufficiently long time series of nowcast estimates may reveal a high level of accuracy when compared with eventual outturn data.

5. Issues for discussion

43. Many international organisations are engaged in the 'nowcasting' of national statistics to meet the demands of their users. A wider dissemination of the methods used across subject areas can raise awareness within national statistics agencies and act as a catalyst for improvements at the national level.

    - *What mechanisms should/could be used by international organisations to raise awareness within national statistics offices?*
    - *Should national statistics agencies be encouraged to develop nowcast estimates of data they currently provide to international organisations, and, if so, what disclosure rules are needed?*

44. Dissemination strategies may vary across and indeed within international organisations. In some statistical areas nowcast estimates of national statistics are unpublished and only used for constructing zone totals or derivative statistics.

    - *What standards and dissemination strategies are needed to determine if nowcast estimates can be published? Are common guidelines needed?*

45. The extent to which nowcasting is used by national statistics offices is likely to vary considerably.

    - *Should a conference be organised/or surveys undertaken to determine practices across countries with a view to sharing international best-practices.*

46. Nowcasting is of course not the only avenue that can be explored to improve timeliness.

    - *What more can international organisations do to disseminate best practice in national statistics compilation methods and survey design/questions?*