Mark Pont, Assessment Programme Lead

Sean Whellams
Head of Profession for Statistics
HM Revenue and Customs (by email)

23 October 2020

Dear Sean

Statistics from the Survey of Personal Incomes

I am writing to you following our review of the HMRC's [Statistics from the Survey of Personal Incomes](#) (SPI). We reviewed the statistics against the [Code of Practice for Statistics](#) pillars of trustworthiness, quality and value.

I would like to take this opportunity to thank you and the statistics team for engaging so effectively with us throughout the review process. I am pleased to confirm that these statistics should continue to be designated as National Statistics. We found several positive examples of the way that HMRC produces and presents these statistics, including:

- the extent to which HMRC has established these statistics as a rich resource answering important questions about people's incomes and earnings. The statistics cover income and tax liability analyses for individuals in the UK, and helpfully include breakdowns for its constituent countries, other sub-UK geographic areas such as counties, and on individual characteristics of taxpayers, such as age. Understanding trends in income and earnings in our society helps us see how living standards may be changing and allows policy to be developed based on the data

- the extent to which HMRC wants users of these statistics and data to be at the centre of statistical production. People we contacted who are members of a Devolved Income Tax Working Group found their engagement with HMRC very useful. We also found excellent examples from the [Scottish Fiscal Commission](#) (SFC) in which HMRC improved the ways in which it met SFC's needs from the personal incomes data. In [the 2019 edition](#) of its annual Statement of Data Needs, SFC said "our three top priorities for income tax data from HMRC, ….have been delivered. These include ….a significant improvement to the main underlying survey data source we use in our forecast. …We appreciate the ongoing engagement from HMRC, and the work that has gone in to fulfilling our data requests"

- the quality assurance of the statistics. The quality of statistics can be seriously compromised when analysts do not fully understand the nature and quality of administrative source data they are working with. HMRC statisticians working on statistics from the SPI have a good appreciation of the wider organisational context in which they work. We found well-designed and managed systems and processes, and established relationships between the statisticians and data suppliers and with the Personal Tax modelling team within HMRC

- clear explanations of the strengths and limitations of the statistics. One of the ways that statistics producers build users' trust is through explaining clearly the strengths and limitations of the statistics and data in relation to different uses. HMRC provides comprehensive and detailed [supporting information](#) about personal incomes statistics identifying limitations of the methods and their application, explaining these to users. HMRC gives information to users about the underlying data sources, which they can take into consideration when weighing how best to use the data. As estimates of personal incomes are derived from a sample, HMRC also sets out the uncertainty in the results using confidence intervals

We have identified some areas where we consider that the statistics should be enhanced in order to ensure that the highest standards of the Code of Practice continue to be met:

Public Value

- While we found proactive and productive engagement around these statistics with users from other government departments and allied bodies such as official forecasters, HMRC should consider how to better take account of the needs of analytical users outside of Government. Examples are users in think tanks and academia and those who use the statistics from the SPI in a data archive called the [Public Use Tape (PUT),](#) which is maintained by the [UK Data Service.](#) PUT contains more-disaggregated data than the statistics in the tables accompanying the personal incomes statistics bulletin and users of PUT data tend to feedback queries about the data to the UK Data Service. Consequently, HMRC personal incomes statisticians might be unaware of some of the needs of such users. To ensure that the statistics and data are maximising public value, HMRC should engage proactively with a wider group of users

- Information requests from users can often provide helpful feedback about how users' experiences of using the personal incomes statistics and data might be enhanced. For example, the personal incomes statisticians improved the data in Table 3.1 of these statistics after a user request. We commend HMRC for its responsiveness to users' needs. HMRC should publish a summary of user feedback showing how feedback has influenced the development of these statistics to help prompt other users to respond to statisticians' invitations to feed back their requests

- Because there is a wide range of official statistics on income and earnings, the Office for National Statistics (ONS) has recently published an [updated guide](#) on sources of data on income and earnings. HMRC could helpfully provide a link to this guide from its personal incomes statistics web landing page

- We recognise that the statistics are not very timely, largely because taxpayers who are required to submit a self-assessment tax return have 10 months after the tax year to do so. To help mitigate this, HMRC should provide alongside the personal incomes statistics a link to the monthly [PAYE Real Time Information](#) (RTI) statistics that it publishes jointly with ONS

Quality

- HMRC's estimates of personal incomes are subject to sampling error. In the detailed [supporting information](#) about personal incomes statistics, HMRC provides some guidance about the precision of its estimates for sub-UK areas. In the guidance HMRC presents examples of the precision of data for parliamentary constituencies and for large and small counties. HMRC told us that many of the data requests it receives are for local authority breakdowns, some of which are subject to quite wide sampling variability. To support users' understanding of the levels of uncertainty in these small area estimates HMRC should provide some indication of their sampling

variability and keep under review the sampling arrangements for these statistics in response to user needs

Trustworthiness

- HMRC [states](#) that personal incomes statistics are published annually, usually in January or February. Over recent years, the first set of tables has been released as late as March or April. The uncertainty over timing of publication does not make for an orderly release of the statistics. HMRC should commit more decisively to a regular publication date for the first release and any revisions and update its metadata to reflect dates of release that it is confident it can commit to

- HMRC releases Tables 3.16 and 3.17, covering data on various data on income tax liabilities on non-savings and non-dividend income due for Scottish taxpayers and non-Scottish taxpayers as official statistics and the statistics in all the other tables as National Statistics. We are satisfied that the statistics in Tables 3.16 and 3.17 are produced to the same standards as the others in this suite and should therefore be labelled as National Statistics for consistency

Our economy team will continue to engage with you and your team in the coming months to follow up on the actions highlighted in this letter and, if you wish to discuss any aspects of this letter, we are happy to do so. I am copying this letter to the responsible team: Mel Whent, Anthony Rourke and Gordon McGregor.

Yours sincerely

Mark Pont
Assessment Programme Lead