Deputy Director for Regulation

Sandra Tudor
Head of Profession for Statistics
Ministry of Housing, Communities and Local Government
(by email)

6 June 2018

Dear Sandra,

As you know the Office for Statistics Regulation committed to reviewing a selection of statistics publications published by the Ministry of Housing, Communities and Local Government (MHCLG) – these were *Net Additional Dwellings*, *New Build Dwellings*, *Affordable Housing Supply* and *Dwelling Stock Estimates*. This commitment was under a programme of planned code compliance checks. We chose to review these due to their importance in helping to understand housing supply and the housing stock in England.

We reviewed the statistics against the three pillars – trustworthiness, quality and value – that are fundamental to supporting public confidence in statistics and form the basis of the *Code of Practice for Statistics*[1]. We recently had a constructive conversation with the responsible statisticians about our review's findings and the current plans for these statistics. This letter confirms our review's findings.

We found several positive examples around MHCLG's statistics production relating to data accessibility and recent or ongoing work to review methods and improve data coverage:

- MHCLG's OpenData platform represents a strong innovation to provide flexible access to data for some of MHCLG's statistics outputs, and MHCLG's Data Strategy seems very promising to build on this innovation. We think there is scope for MHCLG to consider making a fuller range of its statistics outputs available in this way to offer greater accessibility to users, including the statistics we considered in this review.

- MHCLG's ongoing work to improve the coverage of the data sources used to produce *New Build Dwellings* is welcomed, and the statisticians told us about their engagement with some users and data suppliers to progress this work. We strongly encourage this remaining a priority, and we welcome MHCLG's commitment to publish a transparent plan for this ongoing work in its next publication scheduled for 28 June.

- MHCLG statisticians also told us about their recent work, with support from the Office for National Statistics, to review the current data sources and methods used to produce *Dwelling Stock Estimates*, and to consider alternative data sources. MHCLG reflected this work in its recent publication of *Dwelling Stock Estimates* published on 24 May.

- MHCLG's recent work to consider the imputation methods it uses in its statistics production to account for missing data, with the aim of introducing greater consistency and transparency about the methods used.

We found a number of areas where MHCLG should introduce improvements to the statistics publications reviewed, to publicly demonstrate appropriate quality assurances of the data, the methods used and their quality; and to further increase the value of these statistics by presenting more insight within the publications and enhance their coherence:

---

[1] https://www.statisticsauthority.gov.uk/code-of-practice/

- MHCLG should review and refresh the information it presents about the methods used to produce the statistics, their quality and the strengths and limitations of each statistic, to better support their use. MHCLG should also seek opportunities to clarify and introduce consistency to definitions and to its imputation methods in particular, building on its recent work to review the methods.

- MHCLG should publicly demonstrate its quality assurance of the data sources used to produce each statistic, including its ongoing work to improve the coverage of *New Build Dwellings* and publicly reflect recent developments to this work. We encourage MHCLG to take a proportionate approach that reflects the nature of the quality issues and the importance of the statistics.

- MHCLG should increase the value of these statistics by developing more insightful commentary that provides context and helps users to understand trends in the data and consider whether the visual presentation of data within the publications is the most effective means of presenting the key messages.

- MHCLG should enhance the coherence of these statistics, by clearly explaining how the measures presented in each statistics publication relate to each other; the extent of their comparability to similar measures produced for other countries in the UK; and the role of each statistic in contributing towards the evidence base for housing matters in England. We are aware that wider planned improvements to housing statistics[2] are in train, which include plans to improve coherence across housing statistics in the UK. We think these plans are really positive, and we look forward to seeing developments take shape.

In developing and introducing these improvements, MHCLG could build on its recent engagement with a small number of users and think more strategically about how it can engage with other key stakeholders to understand their needs, to inform and prioritise plans.

We have presented these findings in more detail in **Annex 1** of this letter.

We consider it appropriate that these publications retain National Statistics status given the value they have to users, including through their use in other National Statistics products. However, we feel that the areas outlined above need to be addressed in order to meet the high standards required of National Statistics and outlined in the *Code*. I would be grateful for your careful consideration of the findings in this letter and would welcome details of how you plan to address them. We hope that through these plans and subsequent developments we can be reassured so the statistics can continue to be published as National Statistics.

We are aware that the most recent publication of *Dwelling Stock Estimates* was on 24 May. We welcome MHCLG introducing some improvements in that publication including demonstrating publicly recent work to consider the suitability of data sources and methods used to produce the statistics; and starting to publicly reflect the ways in which MHCLG quality assures the data sources used in the context of their use to estimate dwelling stock in England. My team will continue to liaise with your team about future developments to *Dwelling Stock Estimates,* including to develop more insightful commentary in the publication.

We also understand that both the annual *Net Additional Dwellings* and *Affordable Housing Supply* publications will next be published in November this year; and that the next quarterly publication of *New Build Dwelling* is scheduled for June 2018. For these publications, we expect you to demonstrate a clear commitment to addressing the findings during the course of 2018, with updates on MHCLG's progress, so that these publications are improved in a way that represents the highest standards expected of National Statistics.

I appreciate that there are many pressures on your team which at times make it difficult for them to undertake or communicate developments in the way MHCLG may like. I would encourage you to

---

[2] https://gss.civilservice.gov.uk/blog/ripe-for-development-building-better-housing-statistics-for-the-uk/

make use of support available, such as from the Government Statistical Services' Good Practice Team who would be happy to work with you.

Please do not hesitate to get in touch if you would like to discuss any aspects of this letter further.

We look forward to hearing from you.

Your sincerely

Mary Gregory

Deputy Director for Regulation

Annex 1: Compliance check approach and findings

What approach did we take when reviewing the publications?

Last year we committed to reviewing a selection of publications produced by the Ministry of Housing, Communities and Local Government (MHCLG).

We used the *Code of Practice for Statistics* as a framework guiding this high-level review, and highlighted areas that MHCLG should consider to increase the Trustworthiness, Quality and Public Value of these statistics, to demonstrate the highest standards expected of National Statistics. As well as reviewing the statistics publications listed below, we also considered findings from our Systemic Review of the *Public value of Housing and Planning Statistics in the UK*[3], which we published in November 2017. These findings were informed by our engagement with over 60 people including representatives from 43 organisations to understand how housing and planning statistics overall - including the four statistics publications considered in this compliance check - meet the needs of users and how they can be improved. We have included references to relevant sections of the review where they apply. The publications we reviewed were:

A.   Net Additional Dwellings (National Statistics, published annually)
B.   New Build Dwellings (National Statistics, published quarterly)
C.   Affordable Housing Supply (National Statistics, published annually)
D.   Dwelling Stock Estimates (National Statistics, published annually)

Findings

National Statistics status brings with it a commitment to the *Code of Practice for Statistics.* Complying with this framework ensures statistics are of **public value,** are **high quality** and are produced in a way **worthy of trust.** While we can reflect on some positive aspects of these statistics, there are also broad areas that we think MHCLG should strengthen in line with the *Code.* These points should be considered by any organisation publishing National Statistics.

1.   Demonstrating trustworthiness

a. Orderly release of statistics

As part of this review we looked at the extent to which the statistics publications are published to a clear schedule that takes into consideration the uses of the statistics and the decisions they inform. Overall, we found that the forthcoming release schedule for the four statistics publications is clearly published.

2.   Methods, Quality and Quality assurance

a. Suitable data sources

Both the annual *Net Additional Dwellings* and quarterly *New Build Dwellings* statistics publications present a measure of new build 'completions'. The annual measures presented in *Net Additional Dwellings* are higher than those in the quarterly *New Build Dwellings* publications (when aggregated), due to using different data sources. MHCLG describes the *Net Additional Dwellings* measure as the "primary and most comprehensive" and the *New Build Dwellings* figures as "a leading indicator" of new build 'completions'. Many users we spoke to during our Systemic Review

---

[3] https://www.statisticsauthority.gov.uk/publication/public-value-of-statistics-on-housing-and-planning-in-the-uk/

(see paragraphs 3.31-3.34 of that report) expressed a need to have robust statistics with a more complete coverage that are more frequently available than once a year.

MHCLG recently introduced an ongoing programme of work to improve the coverage of data sources used to produce *New Build Dwellings* and is working with some users and data providers to improve consistency of recording – which is a positive development. MHCLG told us it intends to publish a plan for improvements in its next publication scheduled for 28 June, to increase the transparency of its ongoing work, which we welcome. Because of the high profile of house building and clear user need for accurate and timely data, we expect this plan to include a clear assessment of whether the statistics are of a suitable quality for users' needs and an indication of when users can expect that the coverage will be more complete. This will help to ensure that users' trust in the statistics is not eroded further.

Dwelling Stock Estimates statistics are based on a number of data sources, including the Census, English Housing Survey and Labour Force Survey. In the latest publication, MHCLG describes recent work by ONS to consider the existing methodology as part of research to explore whether reliable estimates of private rented dwelling stock for smaller geographies could be produced; and describes its own consideration of alternative data sources, such as data from HM Land Registry and Valuation Office Agency, though there are limitations due to data granularity and data accessibility.

b. Definitions and description of methods

Each of the four statistics publications considered include a section on definitions used and/or a glossary of key terms, which provides helpful guidance to users. There is scope to review and clarify the definitions used in *Affordable Housing Supply*, since the use of different terms such as affordable housing 'delivered', 'starts', 'starts on site' and 'completions' is potentially confusing, without a clear sense about the extent to which these elements are comparable across different types of 'affordable housing' across England. MHCLG statisticians told us that there are historical reasons why definitions may have been different across different funding programmes, and agreed there would be value in introducing greater consistency and are considering work to do this.

Dwelling Stock Estimates, *Net Additional Dwellings* and *New Build Dwellings* each include a description of the methods used to produce the statistics. Overall, the descriptions are generally clear, with the exception of the use of imputation for missing data. In each case, there is limited description about the use of imputation, the specific technique used, and a justification of why this approach was chosen and its likely impact on the quality of the statistics. In response to our feedback, MHCLG has subsequently included a description of their use of imputation methods to produce *Dwelling Stock Estimates* in its most recent publication, which we welcome. MHCLG statisticians also told us that they recently held a workshop to review their use of imputation in statistics production to ensure greater consistency. This is positive, and we encourage MHCLG to build on this work to ensure that it uses imputation methods in a consistent and transparent manner across its publications, to support their use.

Affordable Housing Supply does not include any information about the methods used to produce the statistics, although it does present detailed descriptions of the data sources used. MHCLG told us they do not use imputation methods, since they assume the data is complete as it is linked to grant funding, and there is some auditing of the data by data suppliers. MHCLG should do more to explain the methods it uses to produce these statistics, and be clearer how it, and its data suppliers, quality assure the data used to ensure it is of suitable quality for its use (see Section 2c. below).

c. Assured data quality – Quality, strengths and limitations, and quality assurance

The statistics publication each include a section on 'Data Quality' which describes aspects of quality and the strengths and limitations of the statistics, to varying degrees. These descriptions would benefit from a review and refresh since they appear outdated and do not always give a clear sense of the quality of the outputs, and the reliance users can place on the statistics. For example:

- *Dwelling Stock Estimates*: In the latest 24 May publication, MHCLG statisticians have taken steps to respond to our feedback by including an enhanced description of the quality of these statistics, to reflect its recent work with ONS to review the data sources and methodology used, as described in section 2a.. MHCLG has also included a high level assessment of the strengths and limitations of the data sources used to produce the estimates, and its judgement about why the data sources are fit for purpose. Given the limited timeframe for MHCLG to introduce improvements in this most recent publication, going forward MHCLG should reflect whether further assurances are needed, that takes into account the risk of quality concerns and known assurances of each data source.

- *Net Additional Dwellings*: Descriptions about data quality is limited and relies mainly on response rates of data collections used to produce the statistics. It is unclear what consideration MHCLG have given to the risks to accurate data recording, the checks by data suppliers or data checks by MHCLG themselves.

- *New Build Dwellings:* This publication presents greater detail about MHCLG's recent work to improve response rates and more consistent data recording (see Section 2a.), with the aim of improving data coverage. There is scope for MHCLG to be clearer about what reliance users can place on the statistics given the known data undercount; and an indication of how long it will take for MHCLG to resolve this issue.

- *Affordable Housing Supply:* Brief information about data quality is presented, but there are some aspects of quality mentioned - such as risk of double-counting – that are not explained further, in terms of how MHCLG try to account for this; or 'other' type of affordable housing that cannot be verified by other data sources. There also appears to be a reliance on the data systems being linked to grant funding as an indicator of the data quality, without a clear sense of how this assumption is validated.

All the statistics publications are based primarily or partially on administrative data sources. However, overall, MHCLG presents limited information about the quality assurance of the data used to produce the statistics. There is no clear public demonstration of MHCLG applying the Authority's Regulatory Standard *Quality Assurance of Administrative Data*[4] to demonstrate how each data source is assured from their initial collection to their use in producing the statistics. As set out in the *Code*, quality assurance arrangements should be propionate to the nature of quality issues and the importance of the statistics in serving the public good. MHCLG statisticians told us that they will work to provide narrative assessments of how the data sources used to produce these statistics meet the assurance requirements set out in our Regulatory Standard, and will engage with the GSS Good Practice Team to support this work.

---

[4] https://www.statisticsauthority.gov.uk/osr/monitoring/administrative-data-and-official-statistics/

3.   Increasing the Value of the statistics

a. Data Accessibility and innovation

MHCLG have invested significant resources in developing its OpenData[5] platform for a selection of its National and official statistics and data outputs, which allows users to explore aggregated data and create bespoke data outputs according to their specific needs. The platform currently hosts 208 linked datasets, a portion of which relate to housing and planning statistics. This is a welcome development to improving data accessibility. At present, data from *Dwelling Stock Estimates* and some data tables from *Affordable Housing Supply* are available on the platform, but data presented in both *Net Additional Dwellings* and *New Build Dwellings* are not. It is not made clear if there are plans to make data from these statistics available on OpenData in future, and it is unclear how MHCLG prioritises what statistics are made available in this way. MHCLG statisticians told us that they have a longer-term vision of adding more to their data portal but there are no definite dates for making data from *Net Additional Dwellings* and *New Build Dwellings* available in this way.

Some users we spoke to as part of our Systemic Review (paragraph 3.15) told us that the format of the 'live data tables' that MHCLG publishes alongside its statistics publications do not present time series' or aggregates by regions. As a result, users find it time consuming to create their own aggregated tables. There would be significant value in MHCLG exploring with users whether there is demand for accessing all the data from these statistics outputs using OpenData, to provide users with greater flexibility, to inform MHCLG developing a prioritised schedule of migrating data to the platform. MHCLG statisticians told us that they are also in discussions with ONS about the potential to migrate housing data into the GSS Data Discovery project, though these discussions are in the early stages.

b. Increasing insight and enhancing coherence

Around 50-75% of the contents of these statistics publications are technical notes and much of the commentary is focused on trends. Given the importance and profile of the topics presented in each publication, we would welcome a stronger focus on drawing out valuable insights. There is scope to increase the insight and enhance the coherence of these statistics publications in several ways:

- *Considering the wider context:* The number of homes, and the types of homes being built in England, and how these contribute to the dwelling stock in England represent just some of the measures that could be used to monitor and evaluate housing policy. However, these measures will need to be considered alongside other measures to understand a broader picture about different aspects of housing, such as the supply and demand of homes in England; the changing trends in housing need and housing need for different groups of society; understanding the difference between housing demand and need; whether there is a housing shortage both nationally and locally; whether the right type of houses are being built in the right places etc.

- *Considering how other statistics that measure the same concept fit together*: As described in Section 2a., *Net Additional Dwellings* and *New Build Dwellings* both present a measure of new build 'completions', but these measures differ due to an undercount in the measure presented in *New Build Dwellings.* As a result, users cannot get an accurate picture on the number of new homes that have been built more frequently than annually, which limits opportunities to monitor and influence building activity.

- *Considering how each of these statistics publications relate to each other:* The four outputs cover inter-related aspects of the housing market, all of which are likely to be of interest to users wanting to understand housing matters; affordable housing supply will contribute to new

---

[5] http://opendatacommunities.org/

build activity; which in turn will contribute to the total housing stock in England. There is little attempt within each statistics publication, or elsewhere, to provide a coherent overview of this bigger picture, although an attempt is made in *Net Additional Dwellings* to clarify the difference between that publication and *New Build Dwellings*. Given the wide interest by many stakeholders to get a better understanding of the picture of housing from land acquisition through to build completion, a coherent overview of these statistics would help to aid interpretation.

- *Considering the extent of comparability of these statistics with related statistics across the UK:* Housing is a devolved matter across the UK. Despite this, some users will want to understand trends in housing across the UK. Each of the statistics publications we reviewed include general descriptions of related statistics published in other countries of the UK, but there is little attempt to explain the extent to which these measures are comparable, nor any guidance to support users' interpretation of differences in definitions. The one exception is *Affordable Housing Supply* which does include more detailed descriptions of related measures in Wales, Scotland and Northern Ireland; but it does include a clear sense of the extent of comparability of the measures. MHCLG told us there are plans to improve coherence across housing statistics, as part of wider improvement plans led by the recently established Cross Government Housing Statistics Group, that should help to address this aspect.

- *Considering whether the visual presentation of data is the most effective means of telling a story:* The presentation of data within the publications tend to focus on large tables of data 'counts' at a point in time, and some charts. There is scope to review whether the publications would benefit from making more use of charts and infographics that help to clearly present key messages. For example, the use of data maps to provide snapshots of the most recent variations by local area is helpful, but this approach could be more insightful if it were able to illustrate the longer-term picture. MHCLG could explore whether making data available on its OpenData platform could allow users to create their own visualisations to show geographic trends.

We are encouraging statistics producers to consider areas where bringing together different parts of a data story would benefit users, using different means to present the key messages using insightful commentary to explain patterns and trends; how each of the statistics contribute to the 'bigger picture' on housing; and using appropriate data visualisations to support the commentary. This goes beyond signposting users to related publications, and seeking opportunities to provide a coherent narrative that answers the key questions users have about a topic. Overall, we can see that more can be done to link related topics together to help users understand the different statistics published by MHCLG in context.

MHCLG told us of the inherent difficulties in striking a balance between presenting impartial commentary and providing more context and possible explanations for trends over time. We welcome their desire to engage with the GSS Good Practice Team about how they could develop commentary that is more insightful, whilst remaining impartial. The statisticians also told us that MHCLG was running a pilot for its Data Strategy to develop interactive maps, that could provide more flexibility for users seeking to create their own infographics, but there are no firm timescales yet for these plans. We think these data developments sound promising, and strongly encourage MHCLG to engage with the Good Practice Team to support their developing improved commentary within these publications.

c. Reflecting the range of users and uses

MHCLG publishes some information within the statistics publications about the uses of the data but these descriptions tend to focus on uses in housing policy. Given the wide range of users of these statistics, there is scope for MHCLG to explore and reflect publicly a wider range of examples of how the data inform policy and practice decisions in government and beyond.

MHCLG publishes a user engagement strategy which each statistics publication links to. Based on this policy, it is not clear how MHCLG engage with a wider range of users outside the department to understand their experience of the statistics and ways the statistics could be developed to support their use and maximise their public value. During our Systemic Review, some users told us they perceived there were limited opportunities for them to engage with MHCLG statisticians to share their feedback about the statistics (paragraph 4.6). MHCLG statisticians acknowledged they don't have much direct engagement with users but told us of recent examples where they'd had dialogue with some users and data suppliers, as part of ongoing work to improve the coverage of *New Build Dwellings.* We think this is positive and MHCLG could be more transparent about this engagement activity and any tangible outcomes as a result. More broadly, we think there is scope for MHCLG to think more strategically about key stakeholder groups, and how it can create more sustainable dialogue with these groups over time to understand their needs, and how it can consider these needs in its improvement plans. MHCLG told us that it will be considering how to improve user engagement as part of wider planned improvements by the recently established Cross Government Housing Statistics Group, we look forward to seeing how these plans develop.