

Bernard Jenkin MP
Chairman
Public Administration Select Committee
Committee Office
House of Commons
7 Millbank
London
SW1P 3JA

Monday, 19 November 2012

Dear Bernard,

Thank you for your letter of 16 October to Jill Leyland, asking a number of supplementary questions arising from the oral evidence she and I gave to the Committee in September. Your first question arose from some comments that I made, and my response is set out below.

- 1) Please could you specify the mechanisms by which you believe that data, such as the data on dwellings and their attributes, could be released or shared, and make an assessment of the likely benefits of such a release (Q4)?
 - a. What are the barriers to such release at present?

The issue to which I drew attention in answer to Q4 of the oral evidence session is the impact on users when disclosure control concerns appear excessive or poorly implemented. Although the Statistics User Forum (SUF) recognises that effective data collection depends on confidence in the protection of personal data, there needs to be a proper balance struck between risk of disclosure and the utility and value of datasets. At present, users feel that the approach of the Government Statistical Service (GSS) can be overly conservative, with some notable exceptions such as the release of Census data.

The example referred to in oral evidence was data on dwellings and their attributes produced by the Valuation Office Agency (VOA), and I respond below to the Committee's queries in this particular case. However, it should be noted that there are many other examples where valuable research has been impeded by the difficulty of getting access to data at the level of disaggregation required for useful analysis. SUF has been collecting such examples and discussing them with the National Statistician's Office. I am glad to say that the GSS is setting up a Disclosure Control Task Force to address these issues, and we understand that they will be considering whether the GSS Guidelines should be revised to encourage data producers to consider 'sufficient uncertainty' rather insisting on zero risk of disclosure. We look forward to seeing the outcome of this work. In addition, we are advocating the need for an 'ombudsman', to whom users can take their case if they reach an impasse in discussions with the data providers concerned.

It is not only users outside government whose work is impeded by difficulties in accessing data; it can also be difficult for one government department to have access to data produced

by another. It seems to us that these general issues will need to be resolved if projects such as Beyond 2011 are to be successful.

In the particular case of data on dwellings and their attributes produced by VOA, such data would provide a resource for research on the housing stock at national and local levels; for analysis of the housing market by the building industry, estate agencies and the financial sector; and for social researchers.

Historically, the VOA has been very reluctant to provide access to the data they hold at any level of aggregation. This has been puzzling, as dwelling attributes are not classed as personal data under the Data Protection Act. The VOA has also cited a clause in the Local Government Act which states that data collected for Council Tax purposes should not be used for any other purpose. However, we note the fact that the VOA has been publishing statistics on the number of dwellings by Council Tax band right down to Output Area level on the Neighbourhood Statistics website for some years. This seems to indicate that this apparent legal barrier is not insuperable.

Another barrier is that of charging for data. For example, although it is possible to look up information about the attributes of each rateable non-domestic property on line, the VOA charges £1,100 plus VAT for a simple file for the whole country. This approach is at odds with the government's expressed desire for Open Data.

There are a variety of mechanisms by which VOA data might be made available. First, we would like to see all aggregated statistics published on the Neighbourhood Statistics website (NeSS), down to the lowest possible level, ie Output Area. However, given the richness of the dataset the scope for tabulations is very large, and so we would advocate consultation on which would be the most widely useful tabulations to include.

The wide scope for tabulations and the necessary restrictions on the number which could be included on NeSS argue for making the individual dwelling level data more widely available, so that users can produce their own analyses. Ideally, this would be in the form of a database made available periodically for others to use, either by secure download by authorised users or via the ESRC's Economic and Social Data Service. We believe that although the database would be very large, there are no technical obstacles to doing so. Even if access were to be limited to other government departments including the ONS, there could be potential for linkage with other sources, such as Land Registry transaction data to produce local house price indices which would be very valuable to users more widely.

Kind regards,

Jenny Church
Royal Statistical Society's Honorary Officer for Statistics Users