

## Donor-based imputation methods for admin data: How to replace the number of rooms question on the Census

Tietz, S<sup>1\*</sup>, Mealor, A<sup>1</sup>, Leather, F<sup>1</sup>, and Dent, A<sup>1</sup>

<sup>1</sup>Office for National Statistics

The Census White Paper recommends removing the number of rooms question and to utilise administrative data instead. Previous research demonstrated that Valuation Office Agency (VOA) data can be used for this purpose in principle. However, users of Census micro-level data expect a utility dataset without missingness.

This research project explored whether VOA number of rooms variable is suitable to undergo edit and imputation (E&I) within the standard census framework (i.e. donor imputation). We examined how linked admin and survey data challenges assumptions underlying the E&I process. We linked the 2011 Census with VOA data and attempted to impute VOA number of rooms using auxiliary variables from the Census to predict the missing and inconsistent values. This includes the question of whether we should allow questionnaire data to be changed where they are inconsistent with admin data.

We demonstrated that it is possible to predict VOA number of rooms from census variables, despite some assumptions being partially violated (definitional and time-frame issues, the possibility of subpopulations in the data). This was true for missingness prior to linkage and missingness due to a data linkage failure. Additionally, we successfully tested the design principal of favouring survey data over alternative data where the two are inconsistent.

We observed that some local authorities had large percentages of imputed data, and there was an increase in the percentage of properties where VOA number of rooms was equal to census number of bedrooms. This affects end-user interpretation of the data.

The result is not a general endorsement that all linked survey-admin data can be effectively treated by standard E&I procedures. However, our research can be used as a blue print for other proof-of-concept studies on imputing admin data.

\*Corresponding Author:

Email Address: [stephan.tietz@ons.gov.uk](mailto:stephan.tietz@ons.gov.uk) (S Tietz)

