Method Assurance Review Panel (MARP):

Overcrowding statistics derived from VOA number of (bed)rooms

Version Number: 2.0

Date: 10 March 2021

Author: Stephan Tietz, Emily Mason-Apps and Joe Herson

Revision History

Version	Date	Author	Summary of changes
1.0	27/10/2020	Stephan Tietz	Final draft
2.0	10/03/2021	Stephan Tietz	Pre-publication review

Note

In January 2021, the Office for National Statistics (ONS) published two articles based on the research covered in this paper and feedback received from the Method Assurance Review Panel. Readers are encouraged to read these two articles two understand ONS's latest research into the use of Valuation Office Agency (VOA) data to produce overcrowding statistics:

- Estimating the number of rooms in Census 2021: an update on deriving an occupancy rating from Valuation Office Agency number of rooms (1)
  The Office for National Statistics (ONS) plans to use administrative data from the Valuation Office Agency (VOA) to replace the number of rooms question on Census 2021. This article looks at the impact the use of VOA number of rooms in the 2021 Census has on room occupancy ratings, outlining the method we plan to use to account for definitional differences and asking users for feedback on the method.
- Admin-based levels of overcrowding (using the bedroom standard and Valuation Office Agency number of bedrooms), feasibility research: England and Wales (2)
  This article, considers the potential for using VOA number of bedrooms to support the production of sub-regional estimates of overcrowding in intercensal years (which the current surveys are unable to do due to sample size). Again, it asks users for feedback on the methods we're proposing here.

Most notably, these two articles expand on the research presented here by providing comparisons of level of overcrowding at national level and looking at differences between the census and admin data derived measures not just by tenure, but also accommodation type.

1 Purpose

This paper covers the strategy and analysis undertaken for assuring overcrowding statistics derived from Valuation Office Agency (VOA) number of rooms variable which will replace the number of rooms question on Census 2021.
It also introduces an almost identical strategy and the analysis undertaken for assuring overcrowding statistics derived from VOA number of bedrooms variable to support ONS's admin first transformation.

2 Recommendations

Members of the Panel are invited to provide feedback for:

I. ONS's strategy for assuring overcrowding statistics derived from VOA number of rooms in preparation for Census 2021
II. ONS's analysis underpinning the strategy for assuring overcrowding statistics derived from VOA number of rooms
III. ONS's strategy for assuring overcrowding statistics derived from VOA number of bedrooms as part of ONS's admin first transformation
IV. ONS's analysis underpinning the strategy for assuring overcrowding statistics derived from VOA number of bedrooms

3 Background

3.1 Assurance for Census 2021: VOA number of rooms

A Census 2021 topic consultation (3) found that number of rooms on the census primarily meets the same information need as number of bedrooms. Furthermore, the number of bedrooms data is considered to be more straightforward for respondents to answer compared to number of rooms, as evidenced in the 2011 Census Quality Survey (CQS) (67% accuracy of responses for number of rooms and 91% for number of bedrooms) (4). ONS's intention is to reduce respondent burden by using alternative data sources to provide information on number of rooms, which was announced in the Census 2021 White Paper (5).

Previous research (6) of the feasibility of using Valuation Office Agency (VOA) data to replace the number of rooms question, concluded that VOA data are of a sufficient statistical quality to replace the number of rooms question in 2021. Using VOA number of rooms for Census 2021 does imply a discontinuity with 2011 Census estimates (due to the definitional difference) that users need to be aware of.

The census included kitchens, utility rooms and conservatories in its number of rooms estimates, which the VOA data do not. Since most properties have a kitchen or cooking facilities, the number of rooms in the census data was generally higher than the corresponding number of rooms in the VOA data. Our research found that the VOA number of rooms variable was one less than the number of rooms in the census for 48% of addresses. The research also provided users with a matrix of 2011 Census number of rooms by VOA number of rooms to demonstrate the pattern of differences between the two measures.

It will not be appropriate to measure change in number of rooms from 2011 to 2021 and instead the census bedroom question can be used for comparisons over time. VOA number of rooms for Census 2021 will provide a high-quality relative measure of size, for example enabling the comparison of households across areas within the same time period.

VOA data are now a primary data source to be incorporated into the 2021 Census and will be used to produce the number of rooms output variable. The variable will be referred to as 'VOA number of rooms' (rather than just 'number of rooms') to distinguishing the VOA number of rooms variable from the 2011 Census number of rooms question to ensure users understand the definitional differences. This is the first time ONS will be using administrative data linked to the census to produce a census statistical output.

To incorporate VOA data into the 2021 Census, the two data sources will need to be linked together. This process will result in households that are missing a value for number of rooms, because of a failure of a household to link to the VOA data or missing information in the VOA data before linkage.

As census users have a need for complete coverage, a key aim of the England and Wales Census Programme is to produce a complete and consistent database through editing and imputation (E&I) of all census variables. This process will also reduce bias introduced by differences in the characteristics of linked and unliked records.

Therefore, similarly to all other census variables, E&I of VOA number of rooms will be required to make it fit-for-purpose. E&I will also be necessary where there are inconsistencies between the census number of bedrooms and the VOA number of rooms variables.

The edit and imputation (E&I) strategy for VOA number of rooms was presented (7) to the Census Methodological Assurance Review Panel (MARP) in June 2019. For the benefit of census users, a

methodological working paper (8) and summary (9) of the edit and imputation strategy for VOA number of rooms as well as quality assurance of admin data (QAAD) (10) for the VOA data has been published.

The VOA number of rooms variable will also be used to derive variables to measure overcrowding. In 2011, ONS derived a room occupancy rating which in turn was used for the derivation of the Housing Deprivation Index. The room occupancy rating is the difference between a hypothetical number of rooms required by a given household (e.g. one room per person) and the actual number of rooms. Sub-regional are ranked by occupancy rating to contribute in order to calculate the Housing Deprivation Index.

We are aware that an output for persons per rooms was published on Nomis (11) after 2011 Census, but we have no evidence that it is widely used.

3.2 Assurance for ONS's admin first transformation: VOA number of bedrooms

Previous research (6) of the feasibility of using Valuation Office Agency (VOA) data to replace the number of bedrooms question, concluded that the direct agreement rate between census number of bedrooms and VOA number of bedrooms is 76%, which is much higher than number of rooms (below 16%). Although there are also some definitional differences in the way bedrooms are counted on both sources, the proportion of addresses considered to be overcrowded (on a bedroom basis) was found to be similar (6% using VOA data compared to 5% in the 2011 Census).

While VOA number of bedrooms will not be incorporated into the design of Census 2021[1], we expect that the 2023 National Statistician's recommendation on the future of the census would suggest that number of bedrooms should not be collected on a survey given the availability of high quality information from the VOA.

Furthermore, there is a user need for annual, sub-regional statistics on overcrowding[2]. This has become more urgent as overcrowding has been identified as a possible driver behind Covid-19 infections and work is underway to explore if VOA property attribute data could be used for the analysis of the Covid Infection Survey (CIS) as well as incorporated into the linked Health Data Asset[3] (which plans to link a number of administrative data sources to Census 2011 to explore health outcomes and Covid-19 mortality). We like to note that some of this work requires ethical and legal review which is taken forward by the relevant teams within ONS.

PPA Housing has undertaken initial research[4] using the Family Resources Survey (FRS) (12) which demonstrated that overcrowding measures (e.g. Bedroom Standard) which are comparable to

[1] ONS will continue to ask the number of bedrooms question and use it to derive the Bedroom Standard.
[2] It is worth noting that the need to provide more context for occupancy ratings especially given the huge policy impact of occupancy ratings (e.g. determination of Housing Benefit reductions (16)). Therefore ONS is exploring the feasibility of providing a more comprehensive picture of overcrowding in England and Wales, beyond what is possible using occupancy ratings alone. To do this we are exploring how we can make use of floor area measure provided by VOA as well as Energy Performance Certificate data. This would allow us to produce statistics on "floor space per person" which can provide more context to understand "cramped" living condition as the size of rooms can vary dramatically.
[3] More detailed information is not yet publicly available.
[4] We are currently not planning on publishing this research, because of the different bedroom definition used in the FRS compared to the census. The Resolution Foundation has published occupancy rating tables (15) based on FRS.

existing census and English Housing Survey (EHS) (13) statistics[5] could be derived from large-scale surveys, if the underlying measure of the number of bedrooms (as well as household composition) are comparable. Additionally, we have demonstrated that it is feasible to link VOA data to survey responses. Together, these pieces of research have demonstrated the viability of linking ONS's large-scale household surveys to VOA number of (bed)rooms in order to derive overcrowding statistics (incl. Bedroom Standard) for sub-regional annually.

Therefore, ONS's strategy for assuring overcrowding statistics derived from VOA number of rooms in preparation for Census 2021 will have to align with and work alongside our strategy for assuring overcrowding statistics derived from VOA number of bedrooms as part of ONS's admin first transformation.

Housing policy often uses various measures of overcrowding to assess living standards and inform policy. In the UK, overcrowding is often measured using occupancy ratings of which the Bedroom Standard is the most common.

Overcrowding is often measured using an occupancy rating. An occupancy rating is obtained by subtracting a hypothetical number of bedrooms (or rooms) recommended for a household from the actual number of bedrooms (or rooms) it has available. A household is considered overcrowded if it has fewer bedrooms (or rooms) available than recommended (negative occupancy rating), or under-occupied if it has more (positive occupancy rating). This makes the occupancy rating a straightforward way of measuring overcrowding and under-occupancy. Number of bedrooms is the base for the most widely used overcrowding measure knows as the Bedroom Standard.

The 2011 Census for England and Wales estimated (14) that 1.1 million households had an occupancy rating of '-1 or less' which means that they are "lacking" one bedroom and would be considered overcrowded. 16.1 million households had an occupancy rating of '+1' or '+2 or more', which means they are technically under-occupied.

[5] Similar measures should be producible from the Welsh Housing Condition Survey (WHCS) (17), but are not generally published.

4 Discussion

4.1 Disclaimer

All results presented below are derived from linking pre-edit and imputation 2011 Census with VOA data using Unique Property Reference Numbers (UPRN) assigned to address strings. This matches the methodology for the feasibility research (8) referred to above. Additionally, we have linked to 2011 Census final household-level record data to attach the required (bed)room variables. Proportions of overcrowding have been suppressed if an LA had fewer than 10 overcrowded households or fewer than 100 households in total for a given tenure and accommodation type. We are confident that the results presented here are representative but need to warn the reader that full quality assurance is still underway. Therefore, this research does not provide official statistics on housing and should not be used for policymaking or decision-making.

4.2 Assurance for Census 2021: VOA number of rooms

4.2.1 Assurance strategy

Our strategy for assuring overcrowding statistics derived from VOA number of rooms in preparation for Census 2021 consists of three parts:

1. Demonstrating that census number of rooms and VOA number of rooms measure similar statistical concepts, and provide guidance to users where the two differ
2. Confirming that occupancy ratings derived from VOA number of rooms are similar to occupancy ratings derived from census number of rooms, if we take account of the definitional differences between VOA number of rooms and census number of rooms
3. Demonstrating that the geographical distribution of the proportion of overcrowded households derived from VOA number of rooms is similar to the geographical distribution of the proportion of overcrowded households derived from census number of rooms, and provide guidance to users where the two differ

4.2.2 Similarity between 2011 Census number of rooms and VOA number of rooms

Our feasibility research (8) for using donor-based imputation for census outputs on number of rooms using VOA data examined the relationship between both census and VOA number of rooms with census household variables using several simple regression analyses. Census household variables were used as predictors and number of rooms were outcome variables. Separate analyses were performed for both the census and VOA versions of number of rooms and the coefficients representing the relationship between rooms and other variables were compared.

The relationship between household variables and the 2011 Census number of rooms variable are similar to the relationship with the VOA number of rooms variable across all of the census household variables which allows us to infer that they are measuring similar statistical concepts. The number of rooms predicted using the VOA number of rooms variable was approximately one room less than the number of rooms in the census. This is consistent with the definitional differences between the two sources.

We have expanded on this initial research by exploring if the census household variables can predict the difference between census number of rooms and VOA number of rooms. If Census number or rooms and VOA number of rooms are truly statistically similar, then we would expect the difference to be independent of the predicter variables.

Table 1 lists the regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census accommodation type. We note that the intercept is

close to 1 (which we would expect due to the definitional difference between VOA number of rooms and census number of rooms) and that the unstandardised beta coefficient is close to 0 for most accommodation types. Detached houses, however, have a relatively large beta coefficient of 0.70. It is likely that these types of properties have conservatories and utility rooms, which are counted by VOA but not the census. To a lesser extent that is also true for semi-detached houses (B=0.23). For block of flats we can observe the opposite effect (B=-0.17), which might indicate that flats are less likely to have a kitchen as a separate room. For most of the other 2011 Census household variables, the intercept is close to 1 while unstandardised beta coefficient varies from -0.45 and 0.38.[6] We think that these results, in particular the quite clear narrative emerging from 2011 Census accommodation type, will help users to interpret their own results when starting to use VOA number of rooms, especially when attempting comparing it to results from the 2011 Census number of rooms question.

Table 1 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census accommodation type

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	1.04	
Detached	0.70	0.22
Semi-detached	0.23	0.08
Terraced	0.07	0.02
Block of flats	-0.16	-0.04
Converted/shared house	-0.16	-0.02
Commercial building	-0.18	-0.01
Reference category: Caravan/other mobile structure		

4.2.3 Room occupancy rating based on 2011 Census number of rooms and VOA number of rooms

Occupancy rating is the difference between the actual number of rooms available to a household and a hypothetically required number of rooms for that household. This means occupancy ratings are always an integer. A household with an occupancy rating of +1 or more[7] would be consider under-occupied and one with an occupancy rating of -1 or less would be considered overcrowded.

We would expect occupancy ratings for census and VOA to measure a similar statistical concept as they are a simple difference between census/VOA number of rooms and a measure independent of the data collection (i.e. the number of required rooms is the same). Additionally, we would expect the VOA room occupancy rating to also be typically one less than the census room occupancy rating. Our further analysis has demonstrated that this is indeed the case. We found the median difference between these two measures for all Local Authorities (LA) was -1, while the average varies between -0.6 and -1.7. Boxplots for a sample of the LAs can be found in Section 7.

[6] The full set of results can be found in Section 6.

[7] The threshold used by the English Housing Survey (EHS) to determine underoccupancy is +2 or more while ONS has published census statistics for +1 and +2 or more.

4.2.4 Geographical distributions for proportion of overcrowded households based on the room occupancy rating

Occupancy ratings are often analysed by applying a threshold, for example, to identify the proportion of overcrowded places for a certain geography. Changes in the proportion of households considered overcrowded are likely to draw public and political attention, and therefore we have explored it in more detail.

ONS and the English Housing Survey (EHS) consider households with a occupancy rating of -1 or less to be overcrowded. We would expect a census room occupancy rating threshold of -1 or less to be comparable to a VOA room occupancy rating threshold of -2 or less.

Figure 1 to Figure 3 reproduce choropleths for households in the private rental sector (PRS), the social rental sector (SRS) and owner-occupied households. Please, note that the scale for census and VOA choropleths are not identical which highlights that despite the similarity of the geographical distribution there are differences between the two measures even after adjusting the threshold for VOA to -2 or less. These differences are particularly stark for owner-occupied households.

Figure 1 Proportion of overcrowded households in the private rental sector using 2011 Census number of rooms with a threshold of -1 or less (left) and VOA number of rooms with a threshold of -2 or less (right) for LAs in England Wales. Please, note that the scales are not identical.

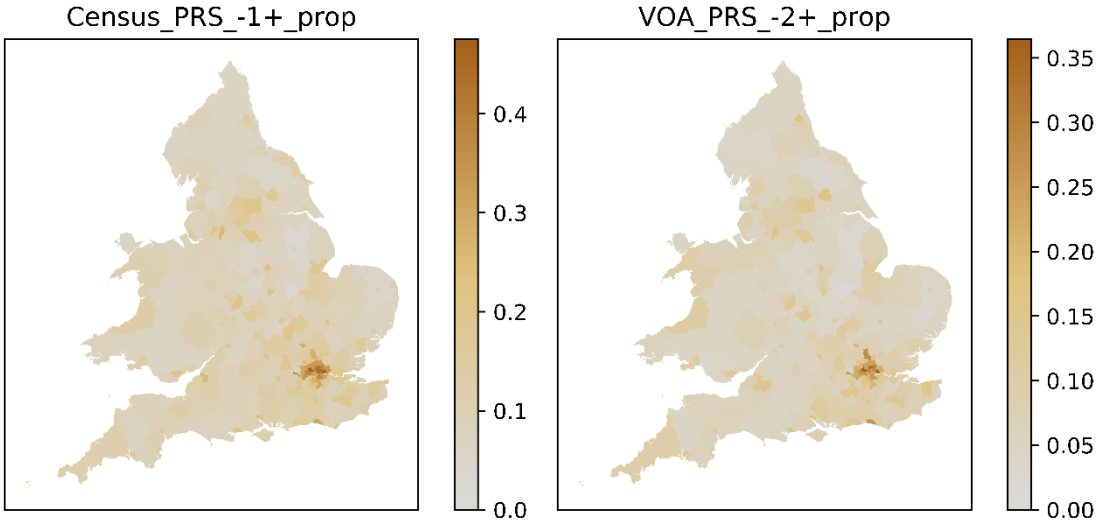


Figure 2 Proportion of overcrowded households in the social rental sector using 2011 Census number of rooms with a threshold of -1 or less (left) and VOA number of rooms with a threshold of -2 or less (right) for LAs in England Wales. Please, note that the scales are not identical.

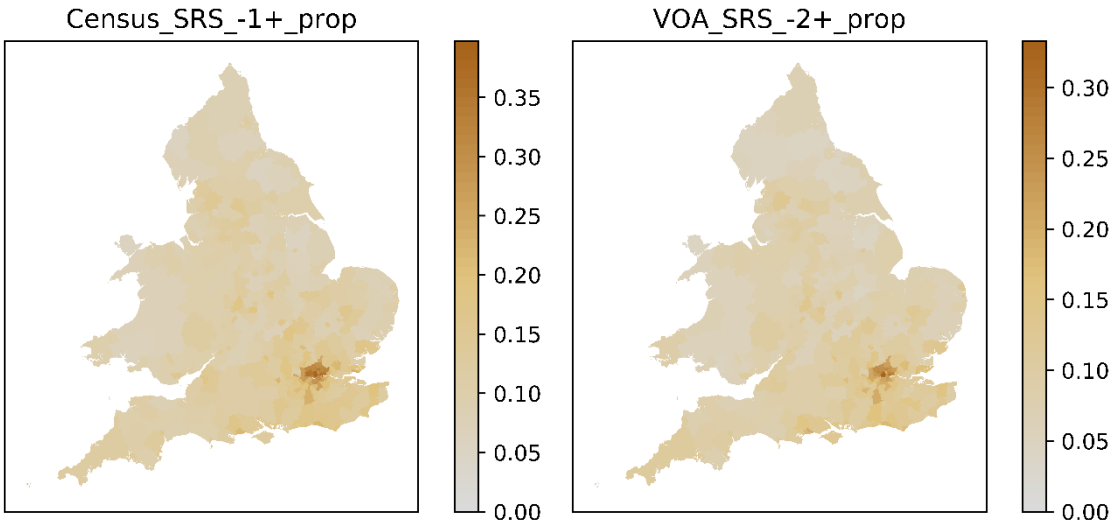
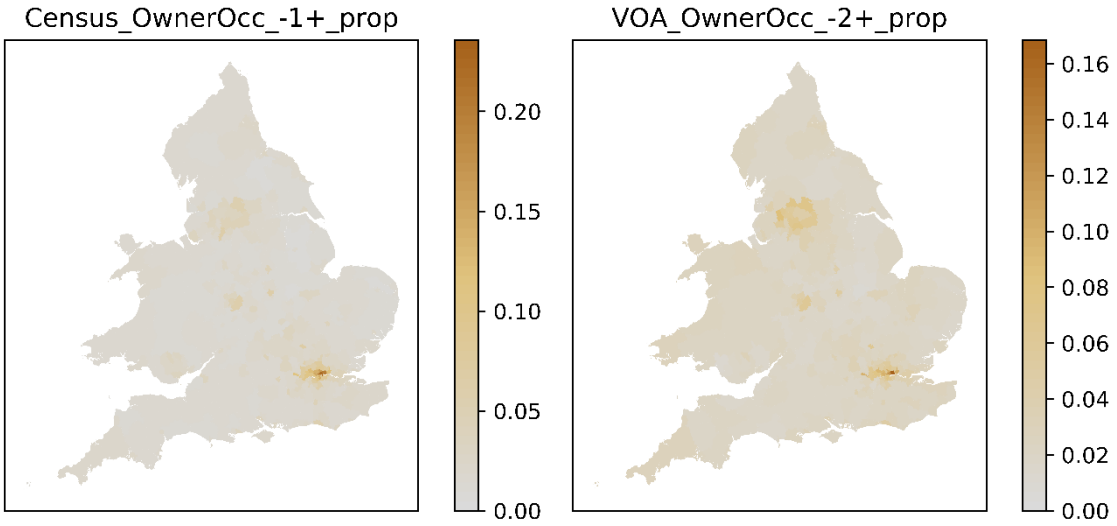


Figure 3 Proportion of overcrowded owner-occupied households using 2011 Census number of rooms with a threshold of -1 or less (left) and VOA number of rooms with a threshold of -2 or less (right) for LAs in England Wales. Please, note that the scales are not identical.

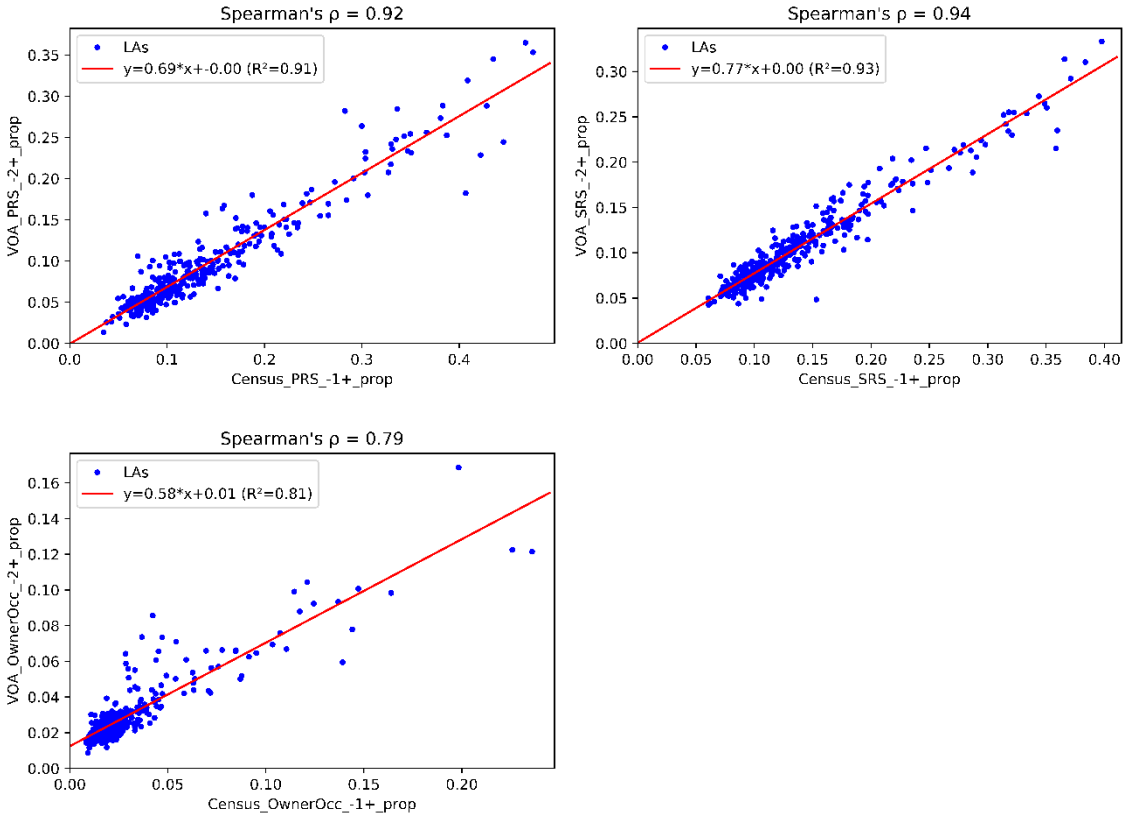


We have used linear regressions to further explore the relationship between the proportions of overcrowded households within LAs as measured using the census number of rooms variable and the VOA number of rooms variable. If we treat the proportion of overcrowded households using the census room occupancy rating as the predictor for the VOA measure, we found that a large proportion of variation could be explained for private and social rental sector households but a smaller amount of the variation for owner occupied households (R^2 = 0.95, 0.96 and 0.73 respectively). We also looked at the intercepts which were close to 0 and the gradients which were 0.69, 0.76 and 0.57, respectively. An intercept of 0 and gradient of 1 would indicate that we can

expect the two measures to be identical.[8] Scatter plots to illustrate these results can be found in Figure 4.

The fact that users are interested in the proportion of overcrowded households, also indicates that it is not just important to compare if the census and VOA measure are similar for each LA, but also if LAs would rank in a similar order for the two different measure. We have used Spearman's rank correlation to assess this. For private rental sector and social rental sector, Spearman's rank correlation coefficient was 0.95 and 0.96, respectively. This indicates that census measure and VOA measure are almost monotonical. However, owner-occupied households had a lower Spearman's rank correlation coefficient of 0.73.

Figure 4 Scatter plots of the proportion of overcrowded households within LAs using VOA number of rooms with a threshold of -2 or less against proportions using 2011 Census number of rooms with a threshold of -1 or less for private rental sector (top left), social rented sector (top right) and ownner occupied households (bottom left). The red lines are linear regressions and Spearman's rank correlation coefficient is provided for context.



Our hypothesis is that owner-occupied households are living in accommodation types (e.g. detached houses as shown in Section 4.2.2) that are more likely to have rooms (other than a kitchen) which are counted by the census and not the VOA, e.g. conservatories and utility rooms.[9] We will draw this

[8] Results for simple regression using other combinations of occupancy rating thresholds can be found in Section 8.

[9] In January 2021, the Office for National Statistics (ONS) published two articles Estimating the number of rooms in Census 2021: an update on deriving an occupancy rating from Valuation Office Agency number of rooms (1) which explores this hypothesis. Also, see the note at the start of the article.

out more clearly for our planned publication by rerunning this type of analysis for accommodation type.

It will be crucial to clearly communicate that the room occupancy rating for Census 2021 differs to the 2011 variable. In 2011 this variable was made available via NOMIS and the UK Data Service. Tables are generally referred to as "occupancy rating (number of rooms)" and we would strongly advise that tables need to be labelled "occupancy rating (VOA number of rooms)" for Census 2021. ONS might also want to consider if this derived variable should be included into the flexible table builder and how we could directly link to additional resource (e.g. our planned publication) to help data users interpret this variable accurately.

The room occupancy rating also feeds into Housing Deprivation Index which itself contributes towards the widely used Index of Multiple Deprivation (IMD).[10] The Housing Deprivation Index is calculated by flagging households as deprived and then ranking the proportion of deprived households within geographical areas. As demonstrated above, it will be essential to apply a threshold of -2 to identify overcrowded households as opposed to a threshold of -1 as used for 2011 Census. This partially adjusts for the definitional difference between census and VOA number of rooms.

There will be some households which would not be flagged as deprived using the VOA number of rooms as it does not count some types of rooms (other than kitchen, e.g. conservatories and utility rooms) which the census does. These types of rooms should probably not be counted when assessing if a household is overcrowded anyway. Furthermore, the room occupancy rating is one of multiple measures feeding into the Housing Deprivation Index which minimises the risk that these changes affect the overall measure. We think that changing the threshold is a better option and provides more consistency with 2011 Census then replacing the (VOA) room occupancy rating with the (census) bedroom occupancy rating (ie. Bedroom Standard) in 2021.

We have no plan to provide specific analysis for persons per rooms as these outputs don't seem to be in wide use and the above findings for number of rooms can be generalised.

We invite the Panel to provide feedback on the suitability and robustness of our analysis, and suggest further analysis if necessary to provide a suitable level of assurance to census users.

4.3 Assurance for ONS's admin first transformation: VOA number of bedrooms

4.3.1 Assurance strategy

Our strategy for assuring overcrowding statistics derived from VOA number of bedrooms as part of ONS's admin first transformation will largely be similar to our strategy for assuring overcrowding statistics derived from VOA number of rooms in preparation for Census 2021. The strategy consists of three parts:

1. Demonstrating that census number of bedrooms and VOA number of bedrooms measure similar statistical concepts, and provide guidance to users where the two differ
2. Confirming that occupancy ratings (i.e. the Bedroom Standard) derived from VOA number of bedrooms are similar to occupancy ratings derived from census number of bedrooms
3. Demonstrating that the geographical distribution of the proportion of overcrowded households derived from VOA number of bedrooms is similar to the geographical

[10] The English Index of Deprivation makes use of the room occupancy rating derived from number of rooms. The Welsh Index of Multiple Deprivation, however, uses the bedroom standard derived from number of bedrooms instead.

distribution of the proportion of overcrowded households derived from census number of bedrooms, and provide guidance to users where the two differ

4.3.2 Similarity between 2011 Census number of bedrooms and VOA number of bedrooms

We have started examining the relationship between both census and VOA number of bedrooms and census household variables using several simple regression analyses. This mirrors our previous research for VOA number of rooms. Initial findings confirm that the relationship between household variables and the 2011 Census bedrooms variable are similar to the relationship with the VOA bedrooms variable across all of the census household variables[11]. This isn't a surprise given the high agreement rate between census and VOA number of bedrooms, but it is an additional confirmation that both variables measure almost the same statistical concepts.

As we have done for number of rooms, we explored if the census household variables can predict the difference between census number of bedrooms and VOA number of bedrooms. If the two measure are statistically similar than we would expect the difference to be independent of the predictor variables.

Table 2 shows the results for 2011 Census accommodation type. We note that the intercept is close to 0 and that the unstandardised beta coefficient is close to 0 for all accommodation types. For all other 2011 Census household variables, the intercept is close to 0 while unstandardised beta coefficient varies between from -0.09 and 0.11, with the notable exceptions for census number of rooms and number of usual residents which have an intercept of -0.33 and -0.15, respectively.[12] These results are really encouraging for users as they suggest that while there are some small differences, the two measures are very similar.

Table 2 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of bedrooms from 2011 Census accommodation type

	Unstandardised beta coefficient *(B)*	**Standardised beta coefficient (β*)***
Intercept	0.04	
Detached	0.07	0.04
Semi-detached	0.03	0.02
Terraced	0.01	0.00
Block of flats	-0.03	-0.02
Converted/shared house	0.01	0.00
Commercial building	0.02	0.00
Reference category: Caravan/other mobile structure		

4.3.3 Bedroom Standard based on 2011 Census number of bedrooms and VOA number of bedrooms

We have calculated the bedroom occupancy rating (known as Bedroom Standard+ using census number of bedrooms and VOA number of bedrooms. Our analysis has shown that the median difference between these two measures for all LAs is 0, while the average varies between -0.2 and

[11] We have not included these results in favour for the simple regressions for the difference between 2011 Census number of bedrooms and VOA number of bedrooms.

[12] The full set of results can be found in Section 7.

0.1. From this we can conclude that these two measures are very similar. Boxplots for a sample of the LAs can be found in Section 8.

4.3.4 Geographical distribution for proportion of overcrowded households based on the Bedroom Standard

To explore this further we looked at the proportion of overcrowded places within LAs using a threshold of -1. Figure 5 to Figure 7 reproduce choropleths for households in the private rental sector (PRS), the social rental sector (SRS) and owner-occupied households. Please, note that the scale for census and VOA choropleths are not identical which highlights that the geographical distributions are similar but not identical. The differences are most notable for owner-occupied households.

Figure 5 Proportion of overcrowded (Bedroom Standard = -1 or less) households in the private rental sector using 2011 Census number of bedrooms (left) and VOA number of bedrooms (right) for LAs in England Wales. Please, note that the scales are not identical.

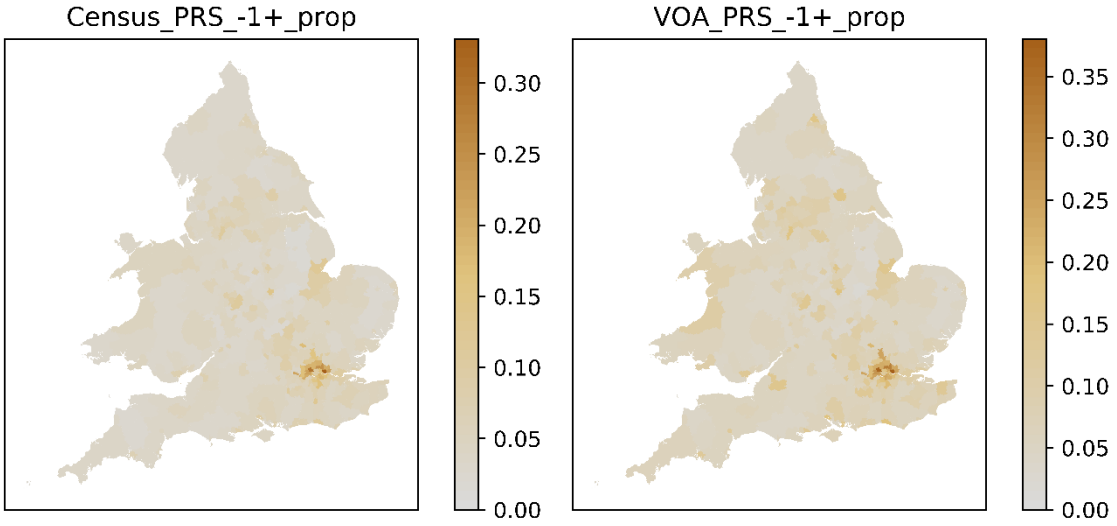


Figure 6 Proportion of overcrowded (Bedroom Standard = -1 or less) households in the social rental sector using 2011 Census number of bedrooms (left) and VOA number of bedrooms (right) for LAs in England Wales. Please, note that the scales are not identical.

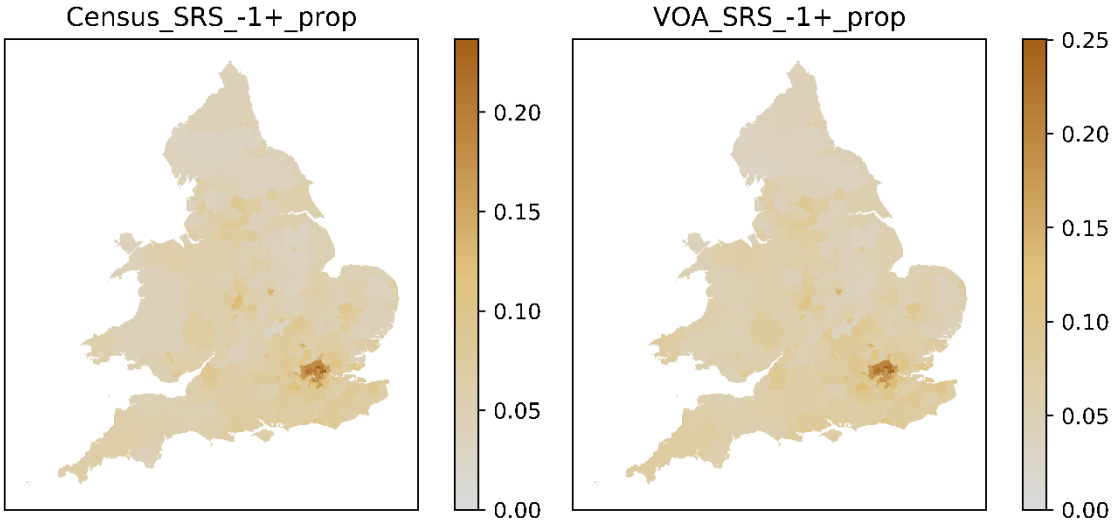
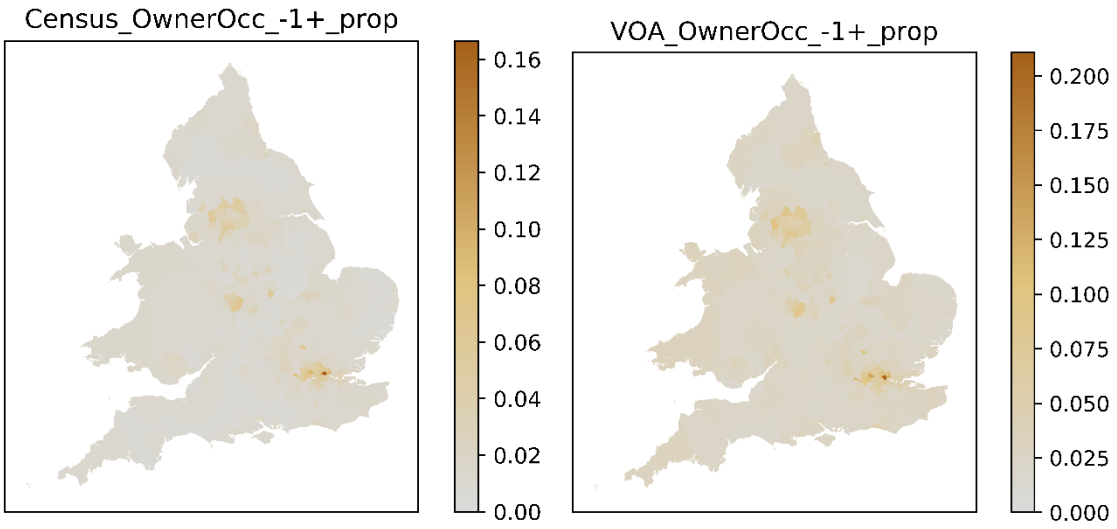


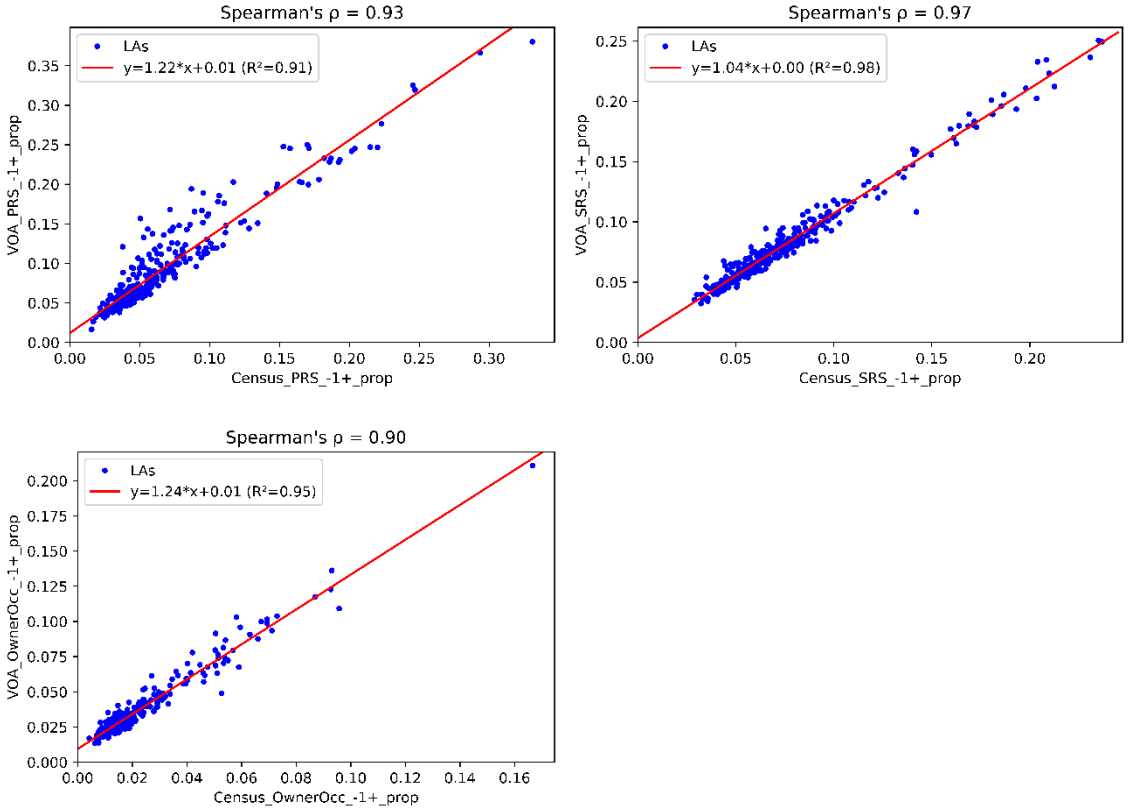
Figure 7 Proportion of overcrowded (Bedroom Standard = -1 or less) owner-occupied households using 2011 Census number of bedrooms (left) and VOA number of bedrooms (right) for LAs in England Wales. Please, note that the scales are not identical.



We have used linear regressions to explore if the proportion of overcrowded households in within LAs as measured using the census number of bedrooms variable can predict the proportion as measured using the VOA number of rooms variable. We found that a large proportion of variation could be explained for households of the different tenure types. R^2 was 0.95, 0.99 and 0.91 for private rental sector, social rental sector and owner-occupied households, respectively. The intercepts were close to 0 and the gradients which were 1.28, 1.04 and 1.25, respectively. This indicates that there is high similarity between the census measure and VOA measure. An intercept of 0 and gradient of 1 would indicate that we can expect the two measures to be identical. Scatter plots to illustrate these results can be found in Figure 8.

As for number of rooms, we have used Spearman's rank correlation to assess if LAs would rank in a similar order for the two different measure. The Spearman's rank correlation coefficient was 0.96, 0.98 and 0.86, respectively. This indicates that census measure and VOA measure are almost monotonical.

Figure 8 Scatter plots of the proportion of overcrowded (Bedroom Standard = -1 or less) households within LAs using VOA number of rooms against proportions using 2011 Census number of rooms for private rental sector (top left), social rented sector (top right) and owner occupied households (bottom left). The red lines are linear regressions and Spearman's rank correlation coefficient is provided for context.



As we would expect, VOA number of bedrooms provide overcrowding statistics very similar to census number of bedrooms, which should be reassuring to data users. While the VOA and census measures are much better aligned for number of bedrooms than number of rooms, we did observe that owner-occupied households are the least similar for number of bedrooms as well. This will require further analysis. VOA has a stricter bedroom definition compared to the census by requiring a minimum room size. Additionally, we can expect the VOA data to be more consistent then self-reported census data. We hope to paint a clearer picture in our planned publication by rerunning this type of analysis for accommodation type.

Our initial findings indicate that moving from number of bedrooms census question to VOA number of bedrooms will not provide perfect continuation in time series data. However, the benefits of being able to measure overcrowding annually and for sub-regional outweigh this downside. An integrated statistical system to replace the census would likely introduce a number of additional methodological changes which will prevent users from exact comparisons and will need to be considered in context.

5 Conclusion

Our initial findings of comparing census room occupancy rating and VOA room occupancy rating highlights that there are notable differences, which warrants ONS to publish an analytical output in advance of Census 2021.[13] We believe this can provide users with assurance and guidance when using these census outputs. We also hope to provide a recommendation on how the VOA room occupancy rating can be used for deriving the Housing Deprivation Index to ensure that we maximise continuation from the 2011 Census. We propose that the VOA room occupancy rating is calculated by subtracting a hypothetical number of rooms recommended for a household (or "rooms required") from the actual number of rooms it has available and adding 1, so that households are considered as overcrowded if they have a VOA room occupancy rating of "negative 1 or less". This new variable should be referred to as "occupancy rating (VOA number of rooms)".

Comparatively, the census Bedroom Standard and VOA Bedroom Standard are much more similar, an adjustment isn't needed. This provides an opportunity for producing annual overcrowding statistics for sub-regional of geography by linking ONS's large-scale household surveys with VOA data. Our analysis, once completed and published, will provide users with the required assurance that these new measures are comparable to existing statistics.

We are inviting the Panel to provide feedback on the analysis underpinning our strategy for assuring overcrowding statistics derived from VOA number of (bed)rooms it is suitability and robustness.

[13] In January 2021, the Office for National Statistics (ONS) published two articles:

- Estimating the number of rooms in Census 2021: an update on deriving an occupancy rating from Valuation Office Agency number of rooms (1)
- Admin-based levels of overcrowding (using the bedroom standard and Valuation Office Agency number of bedrooms), feasibility research: England and Wales (2)

Also, see the note at the start of the article.

6 Annex A: Simple regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census household variables

The table for 2011 Census accommodation type can be found in Section 4.2.2.

Table 3 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census landlord, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	1.41	
Housing association	-0.45	-0.09
Council	-0.43	-0.09
Private landlord	-0.46	-0.11
Employer	-0.24	-0.01
Relative/friend	-0.30	-0.03
Reference category: Other landlord		

Table 4 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census tenure, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	1.06	
Owns outright	0.33	0.11
Owns w/ mortgage	0.39	0.14
Shared ownership	-0.13	-0.01
Rents	-0.08	-0.03
Reference category: Lives here rent free		

Table 5 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census self contained status of accommodation, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.95	
Self-contained	0.33	0.04
Reference category: Is not self-contained		

Table 6 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census central heating status, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	1.28	
No central heating	-0.24	-0.03
Reference category: Has central heating		

Table 7 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census number of bedrooms, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.21	
No of bedrooms	0.38	0.30

Table 8 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census number of cars or vans owned, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.94	
No of cars or vans	0.28	0.20

Table 9 Regression results predicting the difference between 2011 Census number of rooms and VOA number of rooms from 2011 Census number of usual residents, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.94	
No of usual residents	0.14	0.13

7 Annex B: Simple regression results predicting the difference between 2011 Census number of bedrooms and VOA number of bedrooms from 2011 Census household variables

The table for 2011 Census accommodation type can be found in Section 4.3.2.

Table 10 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census landlord, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.08	
Housing association	-0.09	-0.03
Council	-0.09	-0.04
Private landlord	-0.01	0.00
Employer	-0.06	0.00
Relative/friend	-0.06	-0.01
Reference category: Other landlord		

Table 11 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census tenure, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.01	
Owns outright	0.03	0.02
Owns w/ mortgage	0.11	0.07
Shared ownership	0.00	0.00
Rents	0.02	0.01
Reference category: Lives here rent free		

Table 12 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census self-contained status of accommodation, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.03	
Self-contained	0.03	0.01
Reference category: Is not self-contained		

Table 13 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census central heating status, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	0.07	
No central heating	-0.08	-0.02
Reference category: Has central heating		

Table 14 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census number of rooms, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	-0.34	
No of bedrooms	0.07	0.20

Table 15 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census number of cars or vans owned, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	-0.02	
No of cars or vans	0.08	0.10

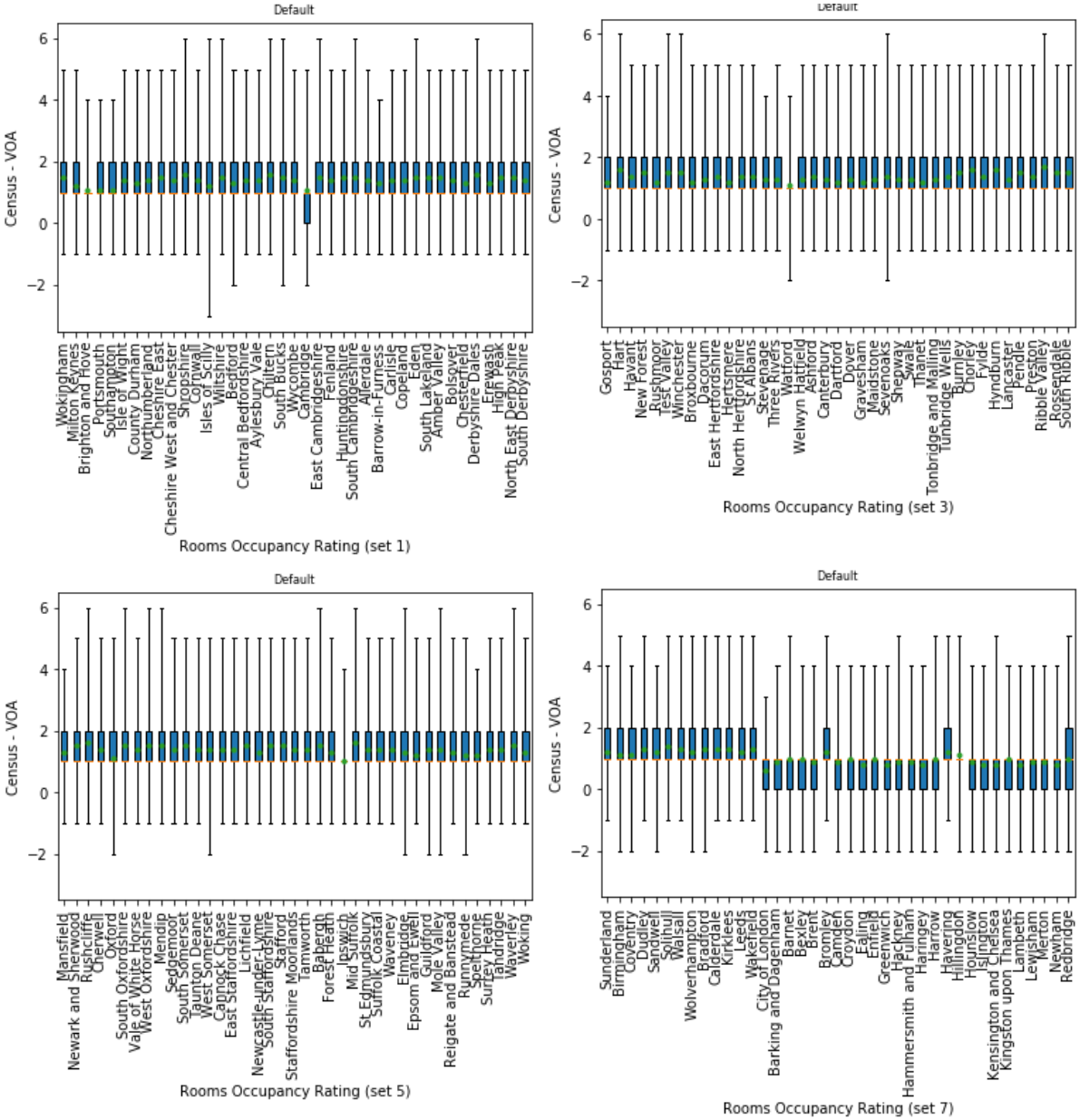
Table 16 Regression results predicting the difference between 2011 Census number of bedrooms and VOA number of rooms from 2011 Census number of usual residents, England and Wales

	Unstandardised beta coefficient *(B)*	Standardised beta coefficient (β*)*
Intercept	-0.14	
No of usual residents	0.09	0.16

8 Annex C: Comparison of census and VOA room occupancy rating by LA

The boxplots below show the difference between the census room occupancy rating and the VOA room occupancy rating for 160 LAs to demonstrate how similar distributions are across England and Wales. The median is a yellow dash, the mean a green dot and the interquartile range is marked with a blue box. The whiskers are the 1st and 99th percentile in order to exclude extreme outliers. Occupancy ratings are by definition an integer and therefore the median, the interquartile range and min/max will be integers.

Figure 9 Box plot presenting differences between 2011 Census room occupancy rating and VOA room occupancy rating for 160 Local Authorities.



9 Annex D: Simple regression and Spearman's rank correlation for proportion of overcrowded households using census/VOA number of rooms

The table below lists the results of simple regressions which use the proportions of overcrowded households within LAs as measured using the census number of rooms variable as predicter and proportions of overcrowded households within LAs as measured using the VOA number of rooms variable as outcome. Different combination of occupancy rating thresholds for the census and VOA have been explored. The Spearman's rank correlation coefficient is also recorded.

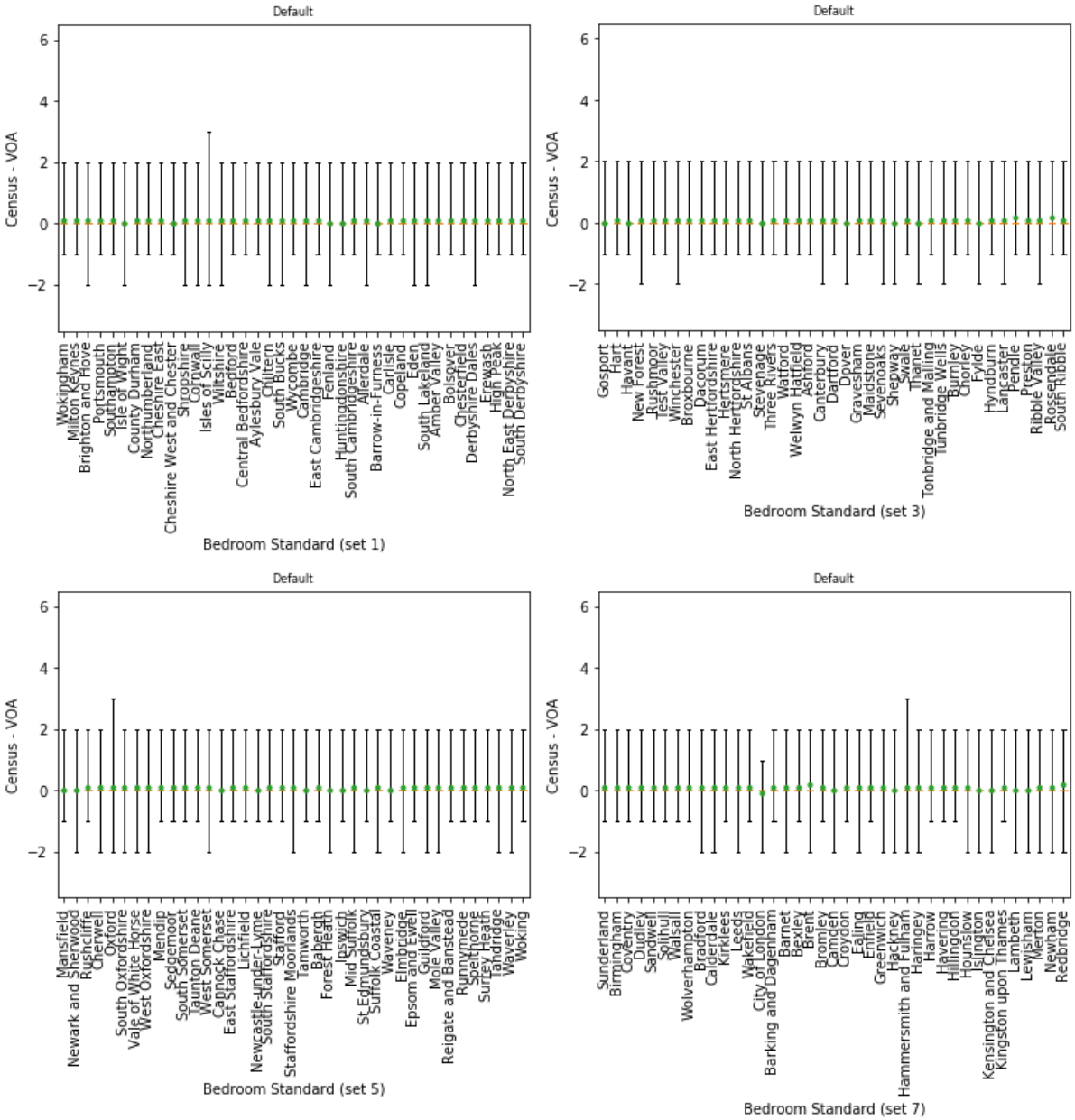
Table 17 Results of simple regressions which use the proportions of overcrowded households within LAs as measured using the 2011 Census number of rooms variable as predicter and proportions of overcrowded households within LAs as measured using the VOA number of rooms variable as outcome.

Tenure group	Census occupancy rating threshold	VOA occupancy rating threshold	Gradient	Intercept	Linear regression R^2	Spearman's Rank Correlation ρ
Private rental	-1 or less	-1 or less	1.26	0.21	0.85	0.91
Private rental	-1 or less	-2 or less	0.69	0.00	0.91	0.92
Private rental	-2 or less	-2 or less	1.70	0.04	0.88	0.90
Social rental	-1 or less	-1 or less	1.10	0.40	0.60	0.86
Social rental	-1 or less	-2 or less	0.77	0.00	0.93	0.94
Social rental	-2 or less	-2 or less	1.88	0.05	0.81	0.86
Owner-occupied	-1 or less	-1 or less	1.73	0.09	0.87	0.83
Owner-occupied	-1 or less	-2 or less	0.58	0.01	0.81	0.79
Owner-occupied	-2 or less	-2 or less	2.02	0.02	0.84	0.77

10 Annex E: Comparison of census and VOA Bedroom Standard by LA

The boxplots below show the difference between the census Bedroom Standard and the VOA Bedroom Standard for 160 LAs to demonstrate how similar distributions are across England and Wales. The median is a yellow dash, the mean a green dot and the interquartile range is marked with a blue box. The whiskers are the 1st and 99th percentile in order to exclude extreme outliers. Occupancy ratings are by definition an integer and therefore the median, the interquartile range and min/max will be integers.

Figure 10 Box plot presenting differences between 2011 Census Bedroom Standard and VOA room Bedroom Standard for 160 Local Authorities.



11 References

1. **Office for National Statistics.** Estimating the number of rooms in Census 2021: an update on deriving an occupancy rating from Valuation Office Agency number of rooms. *Office for National Statistics.* [Online] 2021. [Cited: 10 03 2021.] https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/housingcommunalestablishmentsandvisitors/estimatingthenumberofroomsincensus2021anupdateonderivinganoccupancyratingfromvaluationofficeagencynumberofrooms#levels-of-overcrowding.

2. —. Admin-based levels of overcrowding (using the bedroom standard and Valuation Office Agency number of bedrooms), feasibility research: England and Wales: January 2021. *Office for National Statistics.* [Online] 2021. [Cited: 10 03 2021.] https://www.ons.gov.uk/peoplepopulationandcommunity/housing/articles/adminbasedlevelsofovercrowdingusingthebedroomstandardandvaluationofficeagencynumberofbedroomsfeasibilityresearchenglandandwales/january2021.

3. —. The 2021 Census: Assessment of initial user requirements on content for England and Wales - Housing topic report. [Online] [Cited: 12 11 2020.] https://www.ons.gov.uk/file?uri=/census/censustransformationprogramme/consultations/the2021censusinitialviewoncontentforenglandandwales/topicreport12housing.pdf.

4. —. 2011 Census Quality Survey. [Online] 2014. [Cited: 12 11 2020.] https://www.ons.gov.uk/file?uri=/census/censustransformationprogramme/consultations/the2021censusinitialviewoncontentforenglandandwales/2011censusqualitysurveyreport.pdf.

5. —. Help shape our future: The 2021 Census of population and housing in England and Wales. [Online] 2018. [Cited: 12 11 2020.] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/765089/Census2021WhitePaper.pdf.

6. —. Estimating the number of rooms and bedrooms in the 2021 Census: An alternative approach using Valuation Office Agency data. [Online] 2020. [Cited: 12 11 2020.] https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/housingcommunalestablishmentsandvisitors/estimatingthenumberofroomsandbedroomsinthe2021censusanalternativeapproachusingvaluationofficeagencydata.

7. —. Imputing linked 2011 Census and administrative Valuation Office Agency (VOA) data: A feasibility study. [Online] 2019. [Cited: 12 11 2020.] https://uksa.statisticsauthority.gov.uk/wp-content/uploads/2020/07/EAP120-Imputing-linked-2011-Census-and-administrative-VOA-data-A-feasibility-study.docx.

8. —. ONS working paper series no 20 – Feasibility of using donor-based imputation for census outputs on number of rooms using Valuation Office Agency data. [Online] 2020. [Cited: 12 11 2020.] https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/housingcommunalestablishmentsandvisitors/onsworkingpaperseriesno20feasibilityofusingdonorbasedimputationforcensusoutputsonnumberofroomsusingvaluationofficeagencydata.

9. —. Estimating the number of rooms in Census 2021: an update on imputation methods for Valuation Office Agency data. [Online] 2020. [Cited: 12 11 2020.] https://www.ons.gov.uk/census/censustransformationprogramme/questiondevelopment/housingcommunalestablishmentsandvisitors/estimatingthenumberofroomsincensus2021anupdateonimputationmethodsforvaluationofficeagencydata.

10. —. Valuation Office Agency property attribute data: quality assurance of administrative data used in Census 2021. [Online] 2020. [Cited: 12 11 2020.] https://www.ons.gov.uk/peoplepopulationandcommunity/housing/methodologies/valuationofficeagencypropertyattributedataqualityassuranceofadministrativedatausedincensus2021.

11. **nomis.** Persons per room - People. [Online] 2013. [Cited: 12 11 2020.] https://www.nomisweb.co.uk/census/2011/qs410ew.

12. **Department for Work and Pensions.** Family Resources Survey. [Online] 2020. [Cited: 12 11 2020.] https://www.gov.uk/government/collections/family-resources-survey--2.

13. **Ministry of Housing, Communities & Local Government.** English Housing Survey live tables. [Online] 2019. [Cited: 17 11 2020.] https://www.gov.uk/government/collections/english-housing-survey-live-tables.

14. **Office for National Statistics.** Census suggests 1.1 million households in England and Wales were overcrowded. [Online] 2014. [Cited: 12 11 2020.] https://webarchive.nationalarchives.gov.uk/20160105214004/http:/www.ons.gov.uk/ons/rel/census/2011-census-analysis/overcrowding-and-under-occupation-in-england-and-wales/sty-household-occupancy-and-overcrowding.html.

15. **Resolution Foundation.** The Resolution Foundation Housing Outlook. [Online] 2020. [Cited: 12 11 2020.] https://www.resolutionfoundation.org/app/uploads/2020/04/Housing-Outlook-April-2020.pdf.

16. **House of Commons Library.** Under-occupying social housing: Housing Benefit entitlement. [Online] 2019. [Cited: 12 11 2020.] https://commonslibrary.parliament.uk/research-briefings/sn06272/.

17. **Welsh Government.** Welsh Housing Conditions Survey. [Online] 2020. [Cited: 12 11 2020.] https://gov.wales/welsh-housing-conditions-survey.