# An investigation of overcrowding among the UK households

## Sandeep RAO[*], Chia-Hao CHOU[**]

**Abstract**

*Despite being a major economic power, the UK has been facing several social issues especially with respect to housing its population. The UK law on housing defines the minimum standards of the housing structure and the number of people occupying a given housing space. In spite of these legal requirements, the census data of housing shows that there is a steady increase in issues of adequate housing for the population and that the overcrowding in UK dwellings is a very common problem. This paper identifies various factors that determine the 'space' dimension of overcrowding in the UK housing market. We find that the 'space' dimension of overcrowding in the UK housing market as indicated by 'the total number of bedrooms households' is significantly related to the joint income of the household personals and tenure type, whereas factors like 'age under sixteen' and 'rurality' are not significant. We use a linear probability (regression) model for discrete dependent variables with multiple responses to test the hypotheses and provide robust estimations.*

*Keywords*: overcrowding, housing, linear probability model

## Introduction

The UK housing Act 1985, sec 324 defines overcrowding for the purpose of dwelling. This definition includes the standards specified for room and space (floor area). This legal definition enables a person to apply to the council as homeless if his dwelling is overcrowded and also seek local housing allowances. Overcrowding can cause significant problems relating to physical and mental health, may hamper childhood growth, education and development and also creates concerns regarding general health and safety among the people dwelling in overcrowded houses (*The Impact of Overcrowding on Health & Education: A Review of Evidence and*

---

[*] Sandeep RAO is a Commonwealth doctoral research scholar at the Strathclyde Business School, University of Strathclyde, Glasgow, United Kingdom; e-mail: sandeepkrao@outlook.com.
[**] Chia-Hao CHOU is a doctoral researcher at the Strathclyde Business School, University of Strathclyde, Glasgow, United Kingdom; e-mail: gbyhoward@gmail.com.

*Literature*, 2016). It is important to note that most of the policies relating to the household in the United Kingdom are based on the data collected from the household survey. The housing allowance as well as the legal restrictions on the number of occupants in the house with specific dimensions are all dictated by the housing law. The legal definitions and policies are formed based on the research data and output available from time to time on the annual household survey. This paper uses the annual survey of household data to analyse the importance of various demographic factors which influence the number of bedrooms the UK household has. Specifically, we look into the factors which determine the number of bedrooms in the UK housing, and empirically show that there is overcrowding in the UK housing.

While income and wealth effect theories suggest that, higher income of the household should lead to demonstration effects, which therefore should result in bigger living space. This means a greater number of bedrooms in the house when compared to the number of people dwelling in it. Further considering the UK policies on age and gender restrictions on sharing rooms, the number of rooms in the UK housing must be significantly influenced by these legal requirements. In our empirical study on the UK housing market we find income, ownership, household size are significant factors influencing the number of bedrooms in the UK housing while rurality and age group below sixteen are not as important. However, we also find that overcrowding, which is defined as the total number of rooms per person in the household, is significantly influenced not only by income and ownership type but also by the rurality of the dwelling and whether there are any members below the age of sixteen in the house (using a restricted regression model).

Through this study we contribute to the literature relating to the housing market and housing policy. We specifically identify the factors that influence the number of rooms in the UK housing market, that is the living space aspect, and investigate whether the UK housing policies have any significant impact on the overcrowding in UK housing.

In our empirical analysis we find that two important policy factors - age and the number of people sharing rooms - have no significant influence on the number of rooms in the UK household. Interestingly, we also find that, contrary to the wealth demonstration effect, fully employed household member has a significantly negative relationship with the number of rooms. We also find that people living in rural areas do not have a significantly different preference in the number of rooms in the UK households as compared to their urban counterparts. Finally, we also provide empirical support that there is overcrowding in UK housing. These findings lead us to suggest that stricter enforcement of the UK housing policies are required, and the government of the day must relook into various factors which affect the UK housing as the data suggests that reality does not completely concur with the policies.

The rest of the paper is structured in the following way: review of literature and hypotheses development in section 2, data used in the study and its descriptive

analysis in section 3, empirical analysis in section 4, discussion of empirical findings and conclusion in section 5 and 6, respectively.

## 1. Literature review and hypotheses development

There is a lot of research which discusses the role of the state in formulating policies and interventions to housing, housing size and acceptable standards of living (Chen, 2014; Gao and Asami, 2011; Ha, 2002, 2005; Kim, 1993; Lee and Ngai-ming, 2006; Lim, 1987; Seong-Kyu, 2010; Yip and Chang, 2003).

Myers *et al*. (1996) have addressed overcrowded housing as an important factor influencing housing policy because it is detrimental to both physical and mental health of the people. For instance, overcrowded housing can lead to high infant mortality (Cage and Foster, 2002; Ormandy, 2014), have deleterious impact on child's health development (Leventhal and Newman, 2010) and even wellbeing (Solari and Mare, 2012). In an incident reported by Patrick Cassidy (2013) on emergency medical call, overcrowding in houses had apparently caused danger and dissatisfaction among the residents. However, despite these consequences, overcrowding still exists due to reasons such as 'affordability'. Studies show that the "unaffordability" of housing will yield a negative impact on children (Harkness and Newman, 2003; Newman, 2008). Higher than average prices of housing not only drag the children's mathematic achievement, but also extend to the wage rate in their young adult period (Blau and Haurin, 2017). Further overcrowded housing has been shown to have a direct relationship with the financial status and ability to pay for housing among the migrant workers in Shenzhen, China. In their study, Tao, Wong and Hui (2014) show that the migrant workers demonstrate a high degree of tolerance for 'overcrowded housing' due to their financial inability. Overcrowding is also an issue with respect to immigration (Ward, 1971).

Previous research shows that annual gross income of the householder directly impacts the housing affordability (Newman, 2008), social community isolation (Van Zandt, 2007) and younger household's financial development (Blau and Haurin, 2017). The status of the householder's employment implies the frequency of moving (Böheim and Taylor, 2002), which can have a great impact on the financial investment decisions especially relating to the type of house ownership and size of the house.

Myer *et al*. (1996) use various characteristics such as income, household size, ethnicity, place of dwelling (metropolitan) to find their impact on various housing policy. They show that various factors such as ethnicity, age, immigration, and poverty do not appear to significantly influence the housing market. The subjective characteristic of overcrowded houses makes the objective to measure standards more necessary.

Irit Sinai (2002) uses logistic regression to determine the number of rooms in the households based on the income of the occupants. The literature on housing also

provides evidence of the age and health issues, both mental and physical, directly influencing the number of rooms the household has (Ferraro *et al.*, 2014; Kimhy *et al.*, 2006; Lopoo and London, 2016; Rey-Ares *et al.*, 2016; Ruback and Pandey, 2002; Yu *et al.*, 2015). The age of the household will also cause the degree of overcrowding. The space needed for childhood development will be varied to accommodate the requirement for older household members (Newman, 2008). This could also be influenced by the type of the house. Olotuah (2010) shows that different types of houses (detached house, multiple family dwellings) can have a differential impact on children's mental health development.

One of the major factors that influences the size of the house and the number of rooms is the tenure (Ermisch and Jenkins, 1999; Tiwari and Hasegawa, 2004). While Ermisch and Jenkins (1999) consider housing tenure as one of the dimensions to examine the determinants of residential mobility using the first five waves of the British Household Panel Survey (BHPS), Tiwari and Hasegawa (2004) use tenure as one of the dimensions in the nested multinomial logit model to estimate the choice probabilities and demand elasticities of housing in Tokyo. (Huang *et al.* 2015) confirm the positive relationship between homeownership and residential satisfaction, that is 'sense of belonging, the engagement of social affair and better opportunities for children's education'.

In their paper, Butler *et al.* (2013) constructed a social deprivation index using the townsend index in which the "percent living in overcrowded conditions (more persons in a dwelling unit than number of rooms)" was also used as a major factor of analysis.

The objective of this research is to study 'space' dimension of overcrowding in UK households. This is done by using 'The number of bedrooms the households has' (NBedsX) as a close proxy for the 'space' factor of the overcrowding as per the UK Housing act 1985 legal definition. Specifically, the aim of this study is to identify the factors (variables) that affect the NBedX in the UK. This study uses the 'bedroom standard', that is the number of bedrooms required by the household to avoid undesirable sharing (*English Housing Survey: HOUSEHOLDS, Annual report on England's households, 2011-12*, 2013) for understanding space dimension of overcrowding. Thus, in view of the said objective and literature we formulate several hypotheses and test them by using the multiple regression model.

**H1**: 'Total number of bedrooms in household' has no significant relationship with the income of the household. This is tested by using annual gross income of the HRP[1] and partner (Joint_income). In line with the income and wealth demonstration effects, when the aggregate income of the HRP and partner increases, they should prefer to secure houses with a greater number of rooms. We expect Joint_income to have a positive coefficient.

---

[1]Household Reference Person: see English Housing Survey for complete definition, available at https://www.gov.uk/government/collections/english-housing-survey.

**H2**: 'Total number of bedrooms in household' has no significant relationship with the tenure of the household. – This hypothesis is tested with the help of own and ownL dummies. For tenure1 we create two dummies to capture the essence of the relation between NBedsX and ownership. The study considers two types of ownerships, the house is owned outright by the house hold (own) or is bought with mortgage (ownL), we test for significance of both these variables. We expect 'own' to have a high positive coefficient whereas 'ownL' to have a low positive coefficient. Under the utility theory, we expect that when household members purchase houses, they make long term investments and prefer to derive the maximum benefit out of such investments, therefore they should seek more number of rooms in the house.

**H3**: 'Total number of bedrooms in household' has no significant relationship with 'no people below age 16 living in the house'. – We test this hypothesis using dummy variable 'sixteen'. The dummy is equal to one for households with no persons below age 16 and equal to zero otherwise. We expect a positive and significant relationship between dummy variable of 'sixteen' with the total number of rooms because many housing laws require people above the age of 15-16 to have a separate room. This is also in line with the privacy preferences of young adults.

**H4**: 'Total number of bedrooms in household' has no significant relationship with 'the primary person having full-time employment'. – This hypothesis is tested with the help of the dummy 'fulltime' with value of one if the primary person has full-time employment or with value zero otherwise. In line with the income effect, a full time employed person is more financially secured, and hence we expect a positive coefficient for this variable.

**H5**: 'Total number of bedrooms in household' has no significant relationship with 'Rurality of location'. – We test this hypothesis using the 'urban' dummy with value of one if the person lives in an urban area or with value zero otherwise as per the morphology COA classifications. Urban housing being more expensive than rural housing, we expect a negative relationship between housing in urban areas and the total number of rooms in the house.

**H6**: 'Total number of bedrooms in household' has no significant relationship with 'number of people in the household'. – We test this hypothesis using the 'household size'. The UK regulations on housing clearly define the maximum number of people who can share bedrooms in a house, and this legally restricts more people sharing rooms. Assuming that the households abide by this law we expect a positive coefficient greater than one. Any coefficient lower than one can indicate a greater number of people sharing the rooms and thus, overcrowding in the housing.

## 2. Data and descriptive analysis

After the financial crisis in 2008, the UK Housing market experienced a steep decline. This intensified the housing issues in the UK. 'English Housing Survey: households, annual report on England's households, 2011-12' (2013) data on

overcrowding revealed that the crisis continued to spiral upwards up until 2012-13. The same survey shows that in 2013, however, as per the statistics available with the council of mortgage lenders on the total mortgage disbursements per quarter, the housing crisis showed signs of improvement. This upward trend, however, does not seem to relieve the overcrowding issue. That is why the 2012-2013 census data becomes particularly important for this study. The data applied in this study is English Housing Survey, 2012-2013: Household data. The survey consists of primary household data set collected from interviews with household members in England. This primary data collection is a regular annual exercise conducted by the Government of UK. This paper uses the entire household data collected through this survey for the 2012-2013 year. The total sample size of observations as available through this survey is 13652 with 44 variables, which include both nominal and scale variables. The data analysis is conducted using R-3.4.2 and STATA analysis packages and the linear probability (regression) model for discrete dependent variable with multiple responses.

The main dependent variable - the 'total number of bedrooms the household actually has' is used as a proxy to 'space dimension of overcrowding'. In constructing a regression model, this study looks at finance and household member characteristics & structure as important determinants of overcrowding. Therefore, two variable categories relating to income and age from Myers, Baer and Choi (1996) have been adopted in the model.

In table 3.1 of the housing survey (*English Housing Survey: HOUSEHOLDS, Annual report on England's households, 2011-12*, 2013), we notice that the overcrowding rate varied by tenure presented as: owner occupiers: 1.3%, social renters: 6.6%, and private renters: 5.7%. The same report shows the average annual gross income (HRP & Partner) as: owner occupiers: USD 40,504, social renters: USD 17,550 and private renters: USD 30,146. From the information provided in the descriptive graphs and statistics in the report, this paper tries to check whether an interaction between tenure type and the average annual gross income (HRP & Partner) might affect the household bedroom numbers, the main dependent variable. Households which have persons below age group 16 usually have a lower number of bedrooms because they tend to share it with other younger members of the family. It is also a requirement to have a separate bedroom for persons, whether single or couple, above the age of 16 in the UK for the purpose of local housing allowance.

Various factors from the survey (*English Housing Survey: HOUSEHOLDS, Annual report on England's households, 2011-12*, 2013) which are considered for the model are shown in table 1.

## Table 1. The variables considered for regression model

| Name | Variable Label | Type | Categories |
|------|----------------|------|------------|
| **NBedsX (Dependent)** | Total no of bedrooms household actually has | nominal | - |
| **Urban** | Rurality - morphology (COA) | nominal | urban > 10k, town and fringe, village, hamlets and isolated, dwellings |
| **hhtype6** | Household type - 6 categories | nominal | couple with no dependent children, couple with dependent children, lone parent with dependent children, other multi person households, one person under 60, one person aged 60 or above. |
| **sixteen** | Number of persons under 16 in household | Scale | - |
| **accomhh** | Type of accommodation for household | nominal | detached house or bungalow, semi-detached, terrace/end of terrace, purpose built flat/masionette, flat conversions/rooms. caravan or boat, other |
| **tenure1** | Tenure group 1 | nominal | own outright, buying with mortgage, shared ownership, council tenant, employer, organisation (inc property company), relative/friend , individual |
| **Joint_income** | Annual gross income of the HRP and partner | Scale | - |
| **fulltime** | Employment status (primary) of HRP | nominal | full time work, part-time work, retired, unemployed, full time education, other inactive |
| **hhsizex** | Number of persons in the household | nominal | - |

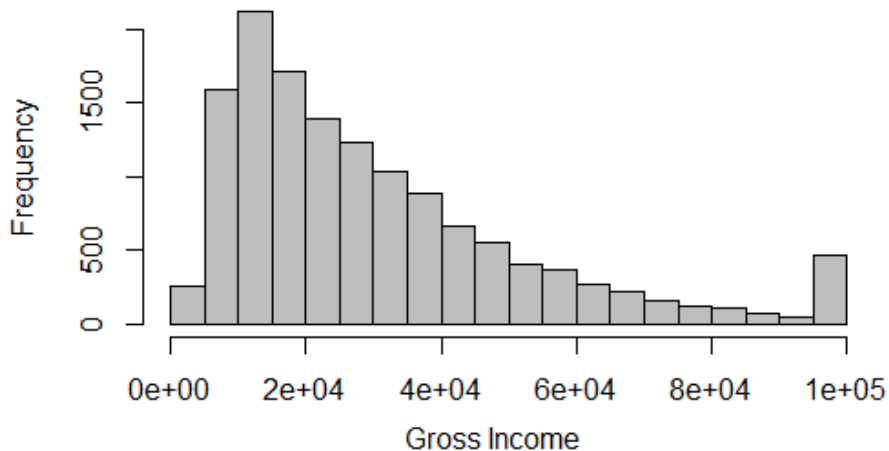*Source*: own representation.

As we can observe, most of the variables are nominal, whereas annual gross income and number of persons under age 16 are scalar. Dummy variables (defined later) are created to capture the effects of the categories in some of the nominal data. The descriptive statistics of NBedsX (the dependent variable) and Joint_income (a key independent variable) are presented in table 2.

**Table 2. Descriptive Statistics values**

| Name | Mean | Std Dev | Skewness | Kurtosis |
|------|------|---------|----------|----------|
| **NBedsX** | 2.751 | 1.008797 | 0.38304 | 3.931754 |
| **Joint_income** | 30487 | 22554.74 | 1.37392 | 4.550249 |

*Source*: own calculations.

Table 2 shows that the mean annual gross income of the household is GBP 30,487 and the skewness and kurtosis (from Table 2 and Figure 1) indicate that the distribution is not normal and highly skewed towards the left.

**Figure 1. Histogram annual gross income**



*Source*: own representation.

Figure 2. shows the bar graphs of various independent variables considered for the study.

**Figure 2. Bar charts of various variables**
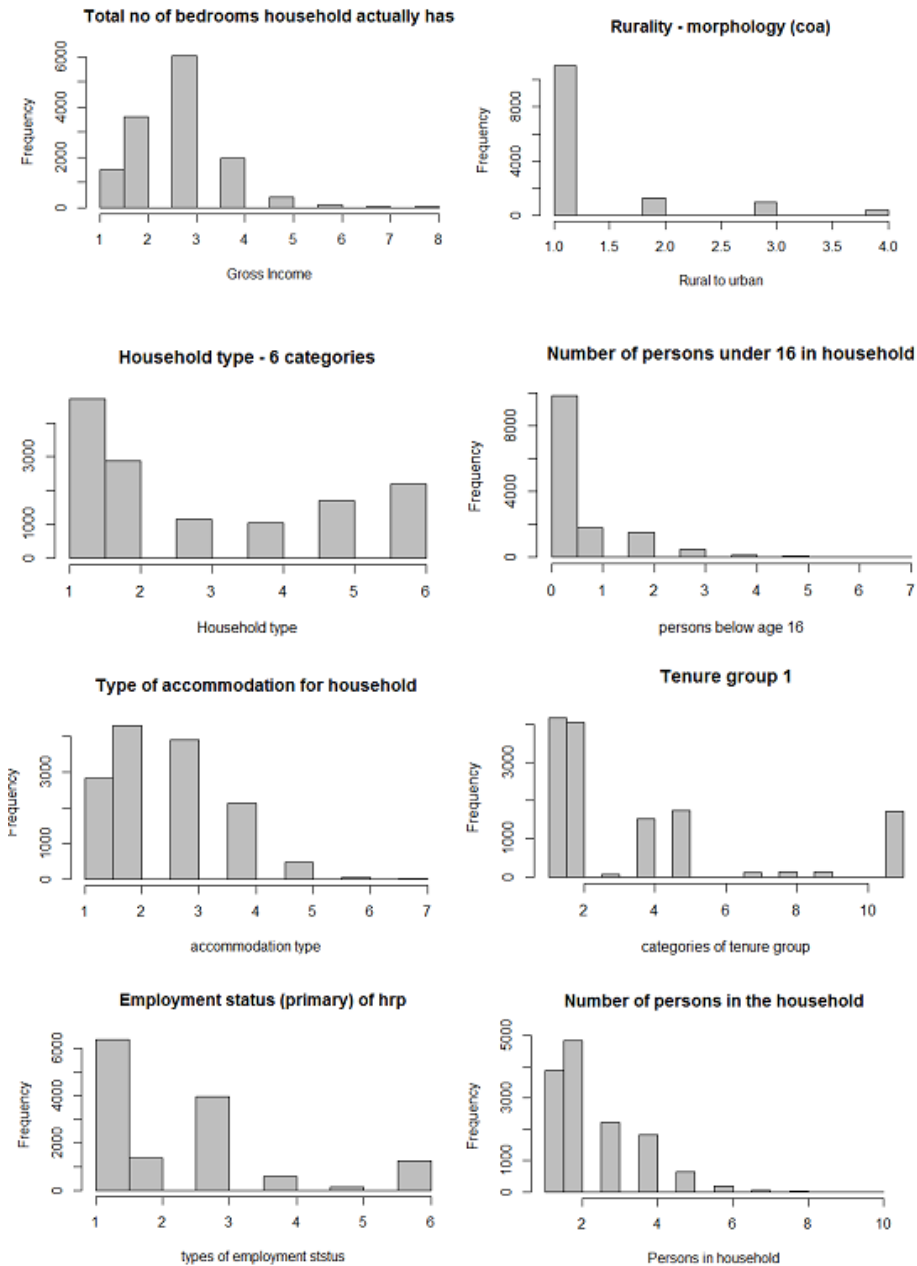


*Source*: own representation.

Rurality classification is based on Census Output Area (COA) morphology which are urban>10K, town and fringe, village, hamlets and isolated dwellings. It is observed that the largest category for this variable belongs to the Urban>10000 people with the frequency of 11052 households. The household by categories are as follows; couple with no dependent children (hhtype6 -1), couple with dependent children (hhtype6 -2), lone parent with dependent children (hhtype6 -3), other multi person households (hhtype6 -4), one person under 60 (hhtype6 -5), one person aged 60 or above (hhtype6 -6). It can be observed that the maximum representation is of couples with no dependent children with a frequency of 4708 households.

In the persons under 16 in household category the maximum representation is with zero people under 16. This shows one key demographic feature of the population: that most of the households surveyed had people above the age group of 16. This data, which is highly skewed, can have and important influence on the results of the regression analysis.

The type of accommodation for households are categorised as detached house or bungalow (accommhh2), semi-detached houses with terrace (accommhh3), purpose built flat mass unit flat (accommhh4), conversions rooms (accommhh5), caravans and boats (accommhh6), and other category (accommhh7). The highest frequency belongs to semidetached houses with a frequency of 4302 representing 31.5% of the total sample collected.

The tenure of household is one of the important factors under consideration when a family decides to move into a house. The type of tenure or the ownership can determine the number of bedrooms in a house. The survey uses the following categories under the tenure factor: own outright, buying with Mortgage, shared ownership, council tenant, HA tenant, employer organisation property, relate or friend, and individual. Among the tenure factors, the data shows 'buying with mortgage' and 'to own the house outright' tenures have the highest frequency of approximately 4100 each.

We also consider the employment status to be one of the determining factors for the size of the house one would prefer. This is because the income of the household is directly correlated to the type of employment status of the occupants. Accordingly, the categories are full-time work, part-time work, retired, unemployed, full-time education and other inactive status. We have a maximum representation from full-time work with the frequency of 6382, part-time work with a frequency of 1347 and retired personal with the frequency of 3956 households.

The bar graphs of various other factors are shown in Figure 2 which shows that the frequency is not normally distributed for any of the variables. The distributions are highly skewed to the left. We have not corrected the data for non-normality for the study. This is one of the limitations of the results of the regression models discussed below.

## 3. Empirical analysis

For the empirical analysis to determine the factors that impact the number of bedrooms the UK housing has, we undertake the multiple regression analysis. The general Linear regression model of multiple discrete dependent variable can be written as follows

$$E(y \mid x_1, x_2, ... x_n) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ..... + \beta_k x_k$$

The regression results can be interpreted as the expected value of the dependent variable 'y' for the given change in the variable $x_j$ under the multiple regression assumptions. The coefficient $\beta$ is the effect of increase in the independent variable x on the expected value of the dependent variable. This expected value takes the value between the range (maximum and minimum) of possible discrete values on a continuous scale. Thus, the coefficients show the average changes in the dependent variable when the independent variable changes by one unit ($\Delta x = 1$), under the condition of other conditions remaining unchanged (ceteris paribus).

We contract both unrestricted (equation 1) and restricted models (equation 2) and test for the robustness of the results. The unrestricted regression model in equation 1 (UN model from now on) and the restricted model (RS model from now on) in the equation 1 are as follows:

$$NBedsX_{UN} = \alpha_{UN} + \beta_1(Joint\_income) + \beta_2(own) + \beta_3(ownL) + \beta_7(sixteen) + \beta_8(fulltime) + \beta_9(urban) \; \beta_{11} + (hhsizex) + \beta_4(Joint\_income*own) + \beta_5(Joint\_income*ownL) + \beta_6(hhtype6) + \beta_{10}(accomhh) + \varepsilon_{UN} \quad (1)$$

$$NBedsX_{RS} = \alpha_{RS} + \gamma_1 (Joint\_income) + \gamma_2(own) + \gamma_3(ownL) + \gamma_4(Joint\_income*own) + \gamma_5(Joint\_income*ownL) + \gamma_6(hhtype6) + \gamma_7(sixteen) + \gamma_8(fulltime) + \gamma_9(urban) + \varepsilon_{RS} \quad (2)$$

See table 1 for definitions of these variables. We create dummy variables for tenure1, sixteen, fulltime, urban. The dummy variables take the value of 1 for the variable tenure1 if the house is owned outright by the household (own) or is bought with mortgage (ownL) and is zero otherwise. The variable agen16 dummy takes the value of one if there are no household members under the age of 16 years and zero otherwise, fulltime dummy is one if the main household member has a fulltime employment, urban dummy takes the value of one if the person lives in an urban area or takes the value zero otherwise. We also interact for income and ownership in both the UN and RS models. We control for the six types of households and accommodation in these models (see table 1).

Table 3 shows the regression output of the unrestricted model (UN). The results shows that Joint_income, hhtype6, fulltime, accomhh, hhsizex and also the joint interaction between the Joint_income and the tenure type (own and ownL) are significant at 0.0001. They also show that variables 'sixteen' and 'urban' have no

significant relationship with the number of bedrooms in the household (NBedsX). Thus, we cannot reject the hypothesis H3 and H5. This means that the rurality and persons under age sixteen living in the house have no significant relationship with the total number of bedrooms the household has (NBedsx). The model has an adjusted R-squared of 0.5129.

We cannot reject H6 as the coefficient of the number of people is positively significant at 1% and with a value of 0.2631, this shows that for approximately every four people there is one room. This rate of people per room is bigger than the requirement under the UK housing law or any other international standards of housing. This clearly shows that there are more occupants compared to the number of rooms indicating overcrowding in UK households.

**Table 3. Regression coefficients of UN model**

| Coefficients: | Estimate | Std.Error | t value | Pr(>\|t\|) | |
|---|---|---|---|---|---|
| (Intercept) | 2.253 | 0.03519 | 64.028 | < 2e-16 | *** |
| Joint_income | 0.000006659 | 7.012E-07 | 9.496 | < 2e-16 | *** |
| own | 0.3323 | 0.02699 | 12.311 | < 2e-16 | *** |
| ownL | 0.2387 | 0.03044 | 7.841 | 4.8E-15 | *** |
| sixteen | 0.002972 | 0.023 | 0.129 | 0.89719 | |
| fulltime | -0.05072 | 0.01635 | -3.102 | 0.00192 | ** |
| urban | -0.006239 | 0.01609 | -0.388 | 0.69816 | |
| hhsizex | 0.2631 | 0.00892 | 29.502 | < 2e-16 | *** |
| Joint_income*own | 0.000003548 | 8.238E-07 | 4.307 | 1.67E-05 | *** |
| Joint_income*ownL | 0.000001994 | 7.943E-07 | 2.51 | 0.0121 | * |
| (hhtype6)2 | -0.1379 | 0.02678 | -5.15 | 2.64E-07 | *** |
| (hhtype6)3 | 0.173 | 0.02846 | 6.079 | 1.24E-09 | *** |
| (hhtype6)4 | 0.3466 | 0.02567 | 13.504 | < 2e-16 | *** |
| (hhtype6)5 | 0.1281 | 0.02494 | 5.136 | 2.85E-07 | *** |
| (hhtype6)6 | 0.07388 | 0.02291 | 3.225 | 0.00126 | ** |
| (accomhh)2 | -0.4688 | 0.01813 | -25.852 | < 2e-16 | *** |
| (accomhh)3 | -0.5915 | 0.01927 | -30.687 | < 2e-16 | *** |
| (accomhh)4 | -1.308 | 0.02432 | -53.803 | < 2e-16 | *** |
| (accomhh)5 | -1.434 | 0.03764 | -38.098 | < 2e-16 | *** |
| (accomhh)6 | -1.376 | 0.1275 | -10.794 | < 2e-16 | *** |
| (accomhh)7 | -0.8639 | 0.2129 | -4.057 | 4.99E-05 | *** |

---
**Significance:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**
Residual standard error: 0.704 on 13631 degrees of freedom
Multiple R-squared:  0.5137,   Adjusted R-squared:  0.5129
F-statistic: 719.8 on 20 and 13631 DF,  p-value: < 2.2e-1

*Source*: own calculations.

The F statistic of 719.8 is also highly significant with a very low p-value nearly equalling zero. This shows that we can reject the proposition that the coefficients are jointly equal to zero in the regression equation of the UN model. We drop accomhh and hhsizex variables and run the restricted regression model. Table 4 shows the RS regression model in detail.

## Table 4. Regression coefficients of RS model

| Coefficients: | Estimate | Std.Error | t value | Pr(>|t|) | |
|---|---|---|---|---|---|
| (Intercept) | 2.203 | 0.02915 | 75.56 | < 2e-16 | *** |
| Joint_income | 0.000009832 | 8.068E-07 | 12.187 | < 2e-16 | *** |
| own | 0.6923 | 0.03007 | 23.019 | < 2e-16 | *** |
| ownL | 0.5773 | 0.03449 | 16.739 | < 2e-16 | *** |
| sixteen | -0.2587 | 0.02517 | -10.28 | < 2e-16 | *** |
| fulltime | -0.09656 | 0.01882 | -5.13 | 2.94E-07 | *** |
| urban | -0.1855 | 0.01797 | -10.325 | < 2e-16 | *** |
| Joint_income*own | 0.000003045 | 9.495E-07 | 3.207 | 0.00134 | ** |
| Joint_income*ownL | 4.453E-07 | 9.149E-07 | 0.487 | 0.62649 | |
| (hhtype6)2 | 0.4642 | 0.02283 | 20.335 | < 2e-16 | *** |
| (hhtype6)3 | 0.4807 | 0.0313 | 15.358 | < 2e-16 | *** |
| (hhtype6)4 | 0.4993 | 0.02911 | 17.153 | < 2e-16 | *** |
| (hhtype6)5 | -0.4347 | 0.02521 | -17.246 | < 2e-16 | *** |
| (hhtype6)6 | -0.3834 | 0.02301 | -16.663 | < 2e-16 | *** |

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.8131 on 13638 degrees of freedom
Multiple R-squared: 0.351,    Adjusted R-squared: 0.3504
F-statistic: 567.4 on 13 and 13638 DF,  p-value: < 2.2e-16

*Source*: own calculations.

Table 4 shows the regression of RS model results that includes only those variables which are considered for the hypotheses. This regression equation results show there is a significant relationship between all the variables on the number of bedrooms actually available in the household at 0.001. However, in this RS regression model, the joint interaction between Joint_income and ownL has no significant relationship. The RS model has an adjusted R-squared value of 0.3504. The F statistic at 567.4 is also highly significant with a very low p-value demonstrating that at least one of the coefficients of the variables being equal to 0 can be ruled out.

ANOVA on UN and RS models are shown in table five. It shows that the F-statistic between the restricted and unrestricted model is highly significant at 0.001. Thus, we can conclude that at least one of the dummy variables in the restricted model (that is accommodation type and hhsizex) is significant for the analysis.

**Table 5. Analysis of variance**

|  | **Res.Df** | **RSS** | **Df** | **Sum of Sq** | **F** | **Pr(>F)** |
|---|---|---|---|---|---|---|
| UN model | 13631 | 6756.4 |  |  |  |  |
| RS model | 13638 | 9015.9 | -7 | -2259.5 | 651.2 | <0.000*** |
| Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 | | | | | | |

*Source*: own calculations.

Akaike's Information Criterion (AIC) is used to judge the adequacy of a regression model. The AIC of the UN model is 29184 and for the RS model AIC is 33108.52. The model with the lowest AIC is preferred while comparing two models. We select the UN model with a lower AIC.

To check the functional adequacy of the regression, we run the RESET test for the UN model. Table 6 shows the results of the RESET test for the UN model.

**Table 6. RESET test**

| RESET = 14.437, df1 = 2, df2 = 13629, p-value = 5.454e-07 |
|---|

*Source*: own calculations.

Table 6 shows the results from the RESET test of the unrestricted model with the p-value nearly equalling zero, so we can conclude that the model does not explain the full variability in NBedsX and is neglecting some systematic information.

We check the robustness of the UN Model by running the heteroskedasticity-robust standard errors of the UN Model. Figure 1 showed that Joint_income is highly skewed to the left. We also run the regression UN by transforming the variable using log (Joint_income) without interaction. The results of both robustness estimators are shown in Table 7.

The regression output of the unrestricted model without interaction and log of income shows that all the variables are highly significant except the age of 16 variable.

We argue that since hypothesis H3 is not rejected in any of the models discussed, we may be able to infer that, from this sample, there is no relationship between the number of bedrooms available in the houses and below 16yrs age members living in the house.

## Table 7. The results of both robustness estimators

| Coefficients: | Heteroskedasticity-robustness | | | Regression with Log of Joint_income | | |
|---|---|---|---|---|---|---|
| | **Estimate** | **Std. Error** | **Pr(>\|t\|)** | **Estimate** | **Std. Error** | **Pr(>\|t\|)** |
| **(Intercept)** | 2.25337 | 0.04021 | 0.000*** | 0.272161 | 0.128372 | 0.034016 * |
| **Joint_income** | 0.00001 | 0.00000 | 0.000*** | - | - | - |
| **logJoint_income** | - | - | - | 0.22307 | 0.012522 | 0.000*** |
| **own** | 0.33227 | 0.02824 | 0.000*** | 0.404696 | 0.018159 | 0.000*** |
| **ownL** | 0.23866 | 0.03085 | 0.000*** | 0.297921 | 0.017919 | 0.000*** |
| **sixteen** | 0.00297 | 0.02239 | 0.894407 | 0.004214 | 0.023236 | 0.856075 |
| **fulltime** | -0.05072 | 0.01636 | 0.0019422* * | -0.060389 | 0.016892 | 0.000*** |
| **urban** | -0.00624 | 0.01724 | 0.7174706 | -0.006321 | 0.016254 | 0.697366 |
| **hhsizex** | 0.26315 | 0.01099 | 0.000*** | 0.257554 | 0.009009 | 0.000*** |
| **Joint_income*own** | 0.00000 | 0.00000 | 0.000*** | - | - | - |
| **Joint_income*ownL** | 0.00000 | 0.00000 | 0.0205567* | - | - | - |
| **(hhtype6)2** | -0.13791 | 0.02907 | 0.000*** | -0.126209 | 0.026993 | 0.000*** |
| **(hhtype6)3** | 0.17304 | 0.02650 | 0.000*** | 0.163491 | 0.028685 | 0.000*** |
| **(hhtype6)4** | 0.34663 | 0.02646 | 0.000*** | 0.356726 | 0.026257 | 0.000*** |
| **(hhtype6)5** | 0.12808 | 0.02619 | 0.000*** | 0.145928 | 0.025681 | 0.000*** |
| **(hhtype6)6** | 0.07388 | 0.02428 | 0.0023475* * | 0.064698 | 0.023274 | 0.005445* * |
| **(accomhh)2** | -0.46879 | 0.02063 | 0.000*** | -0.505345 | 0.018192 | 0.000*** |
| **(accomhh)3** | -0.59148 | 0.02192 | 0.000*** | -0.630058 | 0.019335 | 0.000*** |
| **(accomhh)4** | -1.30823 | 0.02535 | 0.000*** | -1.331547 | 0.024538 | 0.000*** |
| **(accomhh)5** | -1.43384 | 0.03741 | 0.000*** | -1.446455 | 0.038032 | 0.000*** |
| **(accomhh)6** | -1.37620 | 0.08568 | 0.000*** | -1.421267 | 0.128747 | 0.000*** |
| **(accomhh)7** | -0.86391 | 0.30052 | 0.0040501* * | -0.923696 | 0.215056 | 0.000*** |

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.7112 on 13633 degrees of freedom
Multiple R-squared: 0.5036,    Adjusted R-squared: 0.5029
F-statistic: 768.3 on 18 and 13633 DF, p-value: < 0.000

*Source*: own calculations.

Table 8 shows the estimated values of NBedsX and residuals of the UN model. We observe that the residual mean is zero and that the distribution of the estimate NBedsX is not similar to the distribution of the NBedsX.

**Table 8. Estimation and residuals of UN model**

|                        | Min     | 1st Q   | Median  | Mean   | 3rd Q   | Max     |
|------------------------|---------|---------|---------|--------|---------|---------|
| **Predicted NBexsX**   | 1.178   | 2.343   | 2.806   | 2.751  | 3.255   | 5.987   |
| **NBedsX**             | 1       | 2       | 3       | 2.751  | 3       | 8       |
| **Estimates Residuals** | -2.5206 | -0.4311 | -0.0231 | 0      | 0.41073 | 5.31705 |

*Source*: own calculations.

Finally, we also use an alternative definition[2] of overcrowding and test the empirical results for the various factors which affect it. Overcrowding is defined as the total number of bedrooms per household member. We generate a new dependent variable 'overcrowding' by scaling the number of bedrooms (NBedsX) by the number of household members (hhsizex) and run the following regression model presented in equation (3) below. The STATA results are provided in the table 9.

$$\text{Overcrowding} = \alpha_{RS} + \gamma_1 (\text{Joint\_income}) + \gamma_2(\text{fulltime}) + \gamma_3(\text{sixteen}) + \gamma_4(\text{urban}) + \gamma_5(\text{tenure1}) + \gamma_6(\text{hhtype6}) + \gamma_7(\text{hhtype6}) + \varepsilon_{RS} \qquad (3)$$

**Table 9. Overcrowding in UK household**

| overcrowding   | Coef.    | Std-error | t      | P>t   | [95% Conf. Interval] |            |
|----------------|----------|-----------|--------|-------|----------------------|------------|
| **Joint_income** | 5.02E-06 | 2.82E-07  | 17.79  | 0     | 4.47E-06   | 5.57E-06   |
| **fulltime**   | 0.008407 | 0.003721  | 2.26   | 0.024 | 0.0011136  | 0.0157001  |
| **sixteen**    | -0.28982 | 0.005746  | -50.44 | 0     | -0.301087  | -0.2785621 |
| **urban**      | 0.101769 | 0.007118  | 14.3   | 0     | 0.0878169  | 0.1157213  |
| **tenure1**    | -0.05163 | 0.001614  | -31.98 | 0     | -0.0547974 | -0.0484686 |
| **hhtype6**    | 0.209678 | 0.003001  | 69.87  | 0     | 0.2037951  | 0.2155599  |
| **constant**   | 0.829103 | 0.022398  | 37.02  | 0     | 0.7851998  | 0.8730059  |

**Number of observations = 13,652 and R-squared = 0.4312**

*Source*: own calculations.

---

[2] We thank the anonymous reviewer for giving us this alternative definition for overcrowding. We are highly indebted to the reviewer for providing this suggestion.

## 4. Discussion of empirical findings

The main objective of this paper was to determine the factors which influence the number of bedrooms that UK household have and to identify whether there is overcrowding in UK households. The UK household survey data was used, and regression models were constructed with dependent variable 'number of bedrooms the household has' which has multiple discrete values. Based on the factors that determine the housing allowance and the legal definition of overcrowding the main factors such as rurality, household type, number of persons under 16, type of accommodation, tenure, annual income of the household and number of persons in the household were identified. The model provides expected value of the dependent variable when changes in independent variables identified occurs. Restricted and unrestricted models were constructed based on the hypotheses and the unrestricted model was finalised based on the AIC. The regression results showed that age group and rurality did not have any significant effect on the expected value of the number of bedrooms a household has. The model was also tested for robustness with heteroscedastic model and an alternative definition of income, which both provided significant effects for a number of factors identified as hypotheses.

Based on the income and wealth demonstration effects, it was expected that any increase in the household income should have a positive relationship with the number of bedrooms in the household. This is supported by the empirical evidence presented in table 3. However, what is interesting is the very low value of the coefficient estimate of the joint income which is almost equal to zero. Thus, though there is a very high positive significance, the economic value of the coefficient is not very encouraging. This low value may indicate that in the UK market, where the general per capita income is higher compared to the other countries, the demonstration effect, though significant, plays a very minor role in the determination of the number of bedrooms in the house.

What is more interesting and contrary to the economic argument that a full-time employee who is financially secured may prefer larger houses with a greater number of rooms is not supported in the empirical analysis. In fact, the evidence, on the contrary, shows a negative relationship between the full-time employability of the household member and the number of rooms. Probably this contrasting evidence may require further investigation in future research.

While in the unrestricted models which control for various other factors apart from those being studied under the hypotheses the number of people below the age of 16 and the rurality factor do not have any significant impact on the number of rooms in the household, these two factors are significant in a restricted model for determining the overcrowding and the UK market. As pointed out previously, the empirical evidence shows that there is overcrowding in the UK housing. The low positive coefficient of the number of people in the house variable implies that any addition to the number of rooms requires a great addition to the number of household

members. This effectively means there are more than 2 people sharing a single room.

## Conclusion

From the multiple regression model, we can see that the gross annual income of the household, the tenure type, employment status (all representing the financial status) have a significant relationship with the proxy for overcrowding. Thus, we can reasonably conclude that overcrowding is affected by the financial status of the household members. We can also see that the type of accommodation and the number of people living in the house positively affect overcrowding. While the policies relating to the local housing allowances consider the age above sixteen as an important factor, we observe that this, along with the rurality of the house location, are not significant for explaining the number of bedrooms households have in the UK. Finally, in this study, we show that there is overcrowding in UK housing, which may have a negative impact on both physical and mental health and development of the household members and thus the government need to take adequate measures for better implementation of housing policies.

## References

Blau, D. M. and Haurin, D. R. (2017), The effect of the price of housing on child and young adult achievement, *Journal of Real Estate Research,* 39(3), pp. 289-318.

Butler, D. C. Petterson, S. Phillips, R. L. and Bazemore, A. W. (2013), Measures of social deprivation that predict health care access and need within a rational area of primary care service delivery, *Health Services Research,* 48(2), pp. 539-559 (doi:10.1111/j.1475-6773.2012.01449.x).

Böheim, R. and Taylor, M. P. (2002), Tied Down or Room to Move? Investigating The Relationships Between Housing Tenure, Employment Status and Residential Mobility In Britain, *Scottish Journal of Political Economy,* 49(4), pp. 369-392 (retrieved from doi:10.1111/1467-9485.00237).

Cage, R. A. and Foster, J. (2002), Overcrowding and Infant Mortality: A Tale of Two Cities, *Scottish Journal of Political Economy,* 49(2), pp. 129-149 (doi:10.1111/1467-9485.00225).

Cassidy, P. (2013), Overcrowded Hyannis house shuttered, in: Chen, J. (2014). *The New Chinese Model of Public Housing: A Step Forward or Backward?*, *Housing Studies,* 29(4), pp. 534-551 (doi:10.1080/02673037.2013.873392).

Ermisch, J. F. and Jenkins, S. P. (1999), Retirement and housing adjustment in later life: evidence from the British Household Panel Survey, *Labour Economics,* 6(2), pp. 311-333 (retrieved from doi:10.1016/S0927-5371(99)00018-40).

Ferraro, C. F. Trotter, C. L. Nascimento, M. C. Jusot, J.-F. Omotara, B. A. Hodgson, A. . . . Stuart, J.M. (2014), Household Crowding, Social Mixing Patterns and Respiratory Symptoms in Seven Countries of the African Meningitis Belt, *PLoS ONE,* 9(7).

Gao, X. and Asami, Y. (2011), Preferential size of housing in Beijing, *Habitat International,* 35(2), pp. 206-213 (doi:10.1016/j.habitatint.2010.09.002).

Ha, S.-K. (2002), The Urban Poor, Rental Accommodations, and Housing Policy in Korea, *Cities,* 19(3), pp. 195-203 (doi:10.1016/S0264-2751(02)00016-1).

Ha, S.-K. (2005), The role of state-developed housing and housing poverty in Korea, *International Development Planning Review,* 27(2), pp. 227-245.

Huang, Z. Du, X. and Yu, X. (2015), Home ownership and residential satisfaction: Evidence from Hangzhou, China, *Habitat International,* 49, pp. 74-83 (doi:10.1016/j.habitatint.2015.05.008).

Harkness, J. and Newman, S. (2003), Differential effects of homeownership on children from higher-and lower-income families, *Journal of Housing Research, 2003*, pp. 1-19.

Kim, K.-H. (1993), Housing Prices, Affordability, and Government Policy in Korea, *Journal of Real Estate Finance & Economics,* 6(1), pp. 55-72.

Kimhy, D. Harlap, S. Fennig, S. Deutsch, L. Draiman, B. G. Corcoran, C. . . . Malaspina, D. (2006), Maternal household crowding during pregnancy and the offspring's risk of schizophrenia, *Schizophrenia Research,* 86(1), pp. 23-29 (doi:10.1016/j.schres.2006.04.017).

Lee, J. and Ngai-ming, Y. (2006), Public Housing and Family Life in East Asia: Housing History and Social Change in Hong Kong, 1953-1990, *Journal of Family History,* 31(1), pp. 66-82 (doi:10.1177/0363199005283008).

Leventhal, T. and Newman, S. (2010), Housing and child development, *Children and Youth Services Review,* 32(9), pp. 1165-1174 (retrieved fro doi:10.1016/ j.childyouth.2010.03.008).

Lim, G.-C. (1987), Housing Policies for the Urban Poor in Developing Countries, *Journal of the American Planning Association,* 53(2), pp. 176-185 (doi:10.1080/01944368708976650).

Lopoo, L. and London, A. (2016), Household Crowding During Childhood and Long-Term Education Outcomes, *Demography,* 53(3), pp. 699-721 (doi:10.1007/s13524-016-0467-9).

Myers, D. Baer, W. C. and Choi, S.-Y. (1996), The Changing Problem of Overcrowded Housing, *Journal of the American Planning Association,* 62(1), pp. 66-84 (doi:10.1080/01944369608975671).

Newman, S. J. (2008), Does housing matter for poor families? A critical summary of research and issues still to be resolved, *Journal of Policy Analysis and Management,* 27(4), pp. 895-925 (doi:10.1002/pam.20381).

Olotuah, A. O. (2010), Housing Development and Environment Degeneration in Nigeria, *The Built & Human Environment Review*, 3, pp. 42-48.

Ormandy, D. (2014), Housing and child health, *Paediatrics and Child Health,* 24(3), pp. 115-117 (doi:10.1016/j.paed.2013.08.009).

Rey-Ares, L., Irazola, V., Althabe, F., Sobrino, E., Mazzoni, A., Serón, P., Lanas, F., Calandreli, M. and Rubinstein, A. (2016), Lower tract respiratory infection in children younger than 5 years of age and adverse pregnancy outcomes related to household air pollution in Bariloche (Argentina) and Temuco (Chile), *Indoor Air,* 26(6), pp. 964-975 (doi:10.1111/ina.12274).

Ruback, R. B. and Pandey, J. (2002), Mental Distress and Physical Symptoms in the Slums of New Delhi: The Role of Individual, Household, and Neighborhood Factors 1, *Journal of Applied Social Psychology,* 32(11), pp. 2296-2320 (doi:10.1111/j.1559-1816.2002.tb01864.x).

Seong-Kyu, H. (2010), Housing Crises and Policy Transformations in South Korea, *International Journal of Housing Policy,* 10(3), pp. 255-273 (doi:10.1080/14616718.2010.506742).

Sinai, I. (2002), The determinants of the number of rooms occupied by compound dwellers in Kumasi, Ghana: does working at home mean more rooms?, *Applied Geography,* 22(1), pp. 77-90 (doi:10.1016/S0143-6228(01)00017-0).

Solari, C. D. and Mare, R. D. (2012), Housing crowding effects on children's wellbeing, *Social Science Research*, 41(2), pp. 464-476.

Centre for Comparative Housing Research and the Health Policy Research Unit *(2004),The Impact of Overcrowding on Health & Education: A Review of Evidence and Literature*, London: Office of the Deputy Prime Minister (retrieved from http://dera.ioe.ac.uk/5073/1/138631.pdf).

Tao, L. Wong, F. K. and Hui, E. C. (2014). Residential satisfaction of migrant workers in China: A case study of Shenzhen, *Habitat International*, 42, pp. 193-202.

Tiwari, P. and Hasegawa, H. (2004), Demand for housing in Tokyo: a discrete choice analysis, *Regional Studies,* 38(1), pp. 27-42.

Van Zandt, S. (2007), Racial/ethnic differences in housing outcomes for first-time, low-income home buyers: Findings from a National homeownership education program, *Housing Policy Debate,* 18(2), pp. 431-474 (doi:10.1080/10511482.2007.9521606).

Ward, D. (1971), *Cities and immigrants: a geography of change in nineteenth-century America*, New York: Oxford University Press.

Yip, N. M. and Chang, C.-O. (2003), Housing in Taiwan: State Intervention in a Market Driven Housing System, *Journal of Comparative Asian Development,* 2(1), pp. 93-113 (doi:10.1080/15339114.2003.9678373).

Yu, X. Ghasemizadeh, R. Padilla, I. Meeker, J. D. Cordero, J. F. and Alshawabkeh, A. (2015), Sociodemographic patterns of household water-use costs in Puerto Rico, *Science of the Total Environment,* 524, pp. 300-309 (doi:10.1016/j.scitotenv.2015.04.043).