The relationship between the built environment and subjective wellbeing – analysis of cross-sectional data from the English Housing Survey

Gesche M. Huebner[1*], Tadj Oreszczyn[1], Kenan Direk[1], Ian Hamilton[1]

[1]UCL Energy Institute, University College London, London, UK.

*Corresponding author:
Gesche M. Huebner
UCL Energy Institute
Central House
14 Upper Woburn Place
London, WC1H 0NN
UK
g.huebner@ucl.ac.uk

Competing interests statement: The authors declare no competing interests.

Funding:

The authors gratefully acknowledge support from UK Research and Innovation through the Centre for Research into Energy Demand Solutions (EP/R035288/). The funder had no role in study design, data analysis, decision to publish, or preparation of the manuscript.

Word count: 7697

CRediT authorship contribution statement:

GMH: Conceptualization; Methodology; Software; Formal analysis; Investigation; Writing-Original Draft; Writing – Review & Editing; Visualization; Supervision, Project administration

TO: Conceptualization; Writing – Review & Editing; Supervision, Funding acquisition

KD: Software; Resources; Data Curation; Writing – Review & Editing.

IH: Conceptualization; Methodology; Resources; Writing – Review & Editing

Abstract

This paper assesses how subjective wellbeing is related to housing and neighbourhood characteristics, controlling for personal variables. The secondary data analysis was based on the *English Housing Survey, 2017: Housing Stock Data* and the *English Housing Survey: Fuel Poverty Dataset, 2017*, collected in the period April 2016 to March 2018 (N = 9205). Subjective wellbeing was measured with four variables - life satisfaction, the perception of things being worthwhile in life, feeling happy and feeling anxious - that were dichotomized into low and high wellbeing. Logistic regression analysis showed that personal variables are most strongly related to wellbeing but that both housing and neighbourhood variables are also significantly related to it. Finding it difficult to keep the living room warm, being in fuel poverty, and finding it difficult to meet heating costs were associated with lower wellbeing. Low area satisfaction and not feeling safe were also significantly associated with lower wellbeing.

The effects of variables are not constant across all four wellbeing measures used which raises the question 'which wellbeing' should be addressed. Results also showed that targeting householders with lowest wellbeing and hence in greatest need of wellbeing interventions based on publicly available data would be challenging.

Finally, the research community needs to address methodological challenges around identifying the most appropriate covariates, defining wellbeing and considering the measurement of key variables.

Keywords: subjective wellbeing; built environment; health; housing; neighbourhood; English Housing Survey

1. Introduction

Wellbeing is a difficult concept to define; whilst we all have an intuitive understanding of what wellbeing is, no unified definition exists (e.g. Dodge, Daly, Huyton, & Sanders, 2012; Hanc, McAndrew, & Ucci, 2018; Pollard & Lee, 2003). Subjective wellbeing refers to how people think and feel about their own wellbeing; objective wellbeing is based on assumptions made about human needs and is often assessed through measures of income, life expectancy and mortality (Western & Tomaszewski, 2016). Health, mental health and wellbeing are related concepts; for example, the WHO states that "Mental health is a state of well-being…" and "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (World Health Organization, 2018). Being in good health matters greatly for wellbeing but wellbeing can also influence health (Department of Health, 2014). Given that in many countries, people spend about 90% of their time in various buildings such as offices, homes, factories, and schools (Klepeis et al., n.d.; Opinium, 2018), the built environment is likely to play a role in wellbeing.

One of the most frequently used surveys to measure wellbeing is the Warwick-Edinburgh Mental Wellbeing Scale that covers subjective wellbeing and psychological functioning; all items focus on positive mental health (*Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) User Guide*, 2008). The General Health Questionnaire which assesses non-psychotic and minor psychiatric disorders (Vieweg & Hedlund, 1983) can be used for evaluation of subjective wellbeing (Dolan et al., 2011).

In national surveys in the UK, wellbeing is usually measured with four items (Office for National Statistics, 2018c). Life satisfaction is considered as an evaluative approach, asking respondents to make a cognitive assessment of how their life is going overall. The question on things being worthwhile sits in the eudemonic approach. It draws on self-determination theory and measures e.g. people's sense of meaning and purpose in life, connections with family and friends, and whether they feel part of something bigger than themselves (Office for National Statistics, 2018b). The questions on happiness and anxiety reflect an affective component of wellbeing, assessing people's positive and negative emotional experiences over a short timeframe (ibid). These questions assess distinct aspects of personal well-being and should not be combined into one composite measure (Office for National Statistics, 2018a).

In this study, the focus is on subjective wellbeing, using the definition and items as defined by the Office for National Statistics (ONS). For the literature review, a somewhat fluid

distinction is made between mental health and wellbeing as some of the literature discusses both concepts together.

1.1. Evidence on wellbeing and built environment factors

A plethora of studies assessed an association between the physical environment and mental wellbeing and mental health; here, only findings from review studies are summarized with a focus on high-income studies (Chu et al., 2004; Clark et al., 2007; Rachel Cooper et al., 2008; Evans et al., 2003; Hunter et al., 2019; Krefis et al., 2018; Moore et al., 2018; Singh et al., 2019) as these are most relevant to the research conducted for this paper. Review studies that only included children were not considered. The review published by (Chu et al., 2004) was included in the review by Clark et al. (2007) and is hence not covered extra here.

Evans (Evans, 2003) provided a detailed summary of direct and indirect effects of the physical environment on mental health. Regarding direct effects, evidence showed that high-rise housing impacts well-being of women with young children negatively, as does crowding in homes, noise and air pollution. The review states that a number of studies suggest that poor housing quality is associated with poor mental wellbeing; however, because of methodological issues, definite conclusions cannot be drawn. Indirect effects are those where the physical environment influences mental health through changing psychosocial processes that then impact on mental health. Examples are personal control, social support, and restoration – e.g. an environment with green, open spaces can allow greater restoration which in turn impacts positively on mental health.

Clark et al (2007) defined mental health to cover psychological wellbeing, symptoms of psychological distress at a level insufficient for the diagnosis of a disorder, diagnoses of psychiatric illness and suicide. Identified studies were evaluated for their quality. Exposure to violence or crime in the neighbourhood was linked to poorer mental health, as was perceived neighbourhood disorder, e.g. vandalism, lack of facilities, vacant housing and litter. There was also evidence for an effect of chronic noise exposure on mental health, in particular in adults. Studies found little association between household density and mental health and household tenure and mental health. Evidence for an effect of housing quality on mental health was mixed with no longitudinal evidence, and the cross-sectional studies that did indicate an association had study design issues. Consistent evidence indicated that housing and neighbourhood regeneration was linked to improved mental health, and cross-sectional

evidence indicated that access to green and open spaces was linked to improved mental health, though in studies with low response rates.

Cooper et al (2008) covered some areas not relevant to this paper, such educational settings, workplaces, and healthcare settings which are not considered here. Outcome variables were mental health, mental wellbeing and mental capital, which stands for someone's total cognitive and emotional resources. The researchers followed a 'snowball method' to identify relevant work. No rating of the quality of the studies had been made which makes it hard to understand where evidence was strong. The evidence on mental capital indicated that poor housing and poor condition of the neighbourhood have negative impacts. Moving to better housing and better environments can improve mental wellbeing. People in high-rise buildings, particularly on higher floor levels suffer greater mental health problems than those in low-rise developments, and detached housing was associated with good mental wellbeing. High-density living, noise overcrowding, mould, damp, too cold or too hot temperatures have negative impacts on wellbeing.

Moore et al (Moore et al., 2018) restricted their review to randomised (or cluster) randomised controlled trials and controlled before-and-after studies of changes to the built environment. Of the 14 identified studies, only four were considered to contain robust data. The authors indicated that overall evidence was weak that built environment interventions improve mental health and quality-of-life estimates. There was no evidence that urban regeneration and green infrastructure improvements impacted on mental health. Some evidence indicated that changes in green infrastructure improved quality-of-life and reduced self-isolation.

Krefis et al (Krefis et al., 2018) restricted their review to European and North American cities to be able to categorize impact factors into a model of urban wellbeing and health for the global North (von Szombathely et al., 2017). Individual housing factors were not considered in this review. Whilst bias in the review as such was considered, individual studies were not systematically assessed for quality and bias. The authors indicate that most evidence suggested for a link between access to green spaces and positive health and wellbeing outcomes.

Hunter et al. (Hunter et al., 2019) conducted a review on urban greenspace interventions. Quality and bias in the evaluated evidence was assessed. Strong evidence exists that the greening of vacant lots improves wellbeing, e.g. through stress reduction; it also has positive

social outcomes, e.g. reduction in crime and greater perceived safety which in turn likely impact on wellbeing.

Singh et al (Singh et al., 2019) focused exclusively on how housing disadvantage impacts on mental health in temporally ordered studies where the exposure to housing disadvantage preceded mental health measures. All 12 reviewed studies, of which five were judged as high quality, showed some relationship between housing disadvantage and mental health. Substandard housing quality was shown to be linked to higher stress levels, anxiety was higher in renters than owners; overcrowding was linked to a measure of mean depressive symptoms but not a depressive disorder. Ige et al (Ige et al., 2019) focused on the link between buildings and physical health but reported that relocation to low-poverty areas was associated with a decrease in depressive symptoms and that fabric improvements were related to mental health though through complex temporal relationships (Curl et al., 2015).

In summary, the reviews considered varied in important aspects; some focused only housing and neighbourhood factors; others only on urban factors (Hunter et al., 2019; Krefis et al., 2018; Moore et al., 2018); one review only on housing disadvantage (Singh et al., 2019); only some reviews assessed the quality of the underlying studies (Clark et al., 2007; Hunter et al., 2019; Moore et al., 2018; Singh et al., 2019); and the outcome measures considered varied substantially. Overall, the evidence seems to indicate that there is a link between urban greenspace and wellbeing outcomes; but conflicting findings on the effect of urban regeneration. Whilst cross-sectional studies suggest a link between housing quality and mental health, there was only scarce longitudinal evidence (Singh et al., 2019). What unites all reviews is the call for more and better studies in this field, particularly those that allow drawing causal conclusions, have large samples and that control for confounders.

1.2. Outline of this study

This study analyses cross-sectional data, i.e. it cannot contribute to the much-needed longitudinal work. However, it uses a large sample with participants drawn from the English population (N = 9205) and so overcomes the issue of small, select samples. It includes a very wide range of variables, spanning personal, housing and neighbourhood level factors. Controlling for personal variables is essential in isolating built environment variables. Most factors discussed in previous reviews were included in some form, with the exception of data on regeneration schemes and green spaces as they were not present in the data (beyond information about garden access). It uses four wellbeing measures that allow studying of

different facets of wellbeing. What further sets our work apart, is a focus on variables related to being able to keep a dwelling warm / being fuel poor that have received little attention in previous studies.

Finally, the study follows good research practices around open, reproducible science – the work was preregistered (*10.17605/OSF.IO/F26ZS*), the code is made available *(https://github.com/Gesche-Huebner/Wellbeing_Repo*), and the RECORD reporting guideline (Benchimol et al., 2015) is followed (Appendix A). The data cannot be shared by the authors but is accessible on a public database. Details on those practices are given in the attached checklist (Appendix B).

2. Methods

No ethics approval was needed since secondary, fully anonymized data were used.

2.1. Data sets

This study uses data from the English Housing Survey (EHS). The EHS is a continuous national survey commissioned by the Ministry of Housing, Communities and Local Government (MHCLG). It collects information about people's housing circumstances and the condition of housing in England. It consists of: (1) a household interview (i.e. self-reported data), and (2) a physical inspection of a subsample of the properties through a surveyor (Department for Communities and Local Government, 2018). The addresses for the initial EHS sample are selected through a systematic random sample design. The response rate for the interview was 58%. For the physical survey, a disproportionate number of social housing renters is included to ensure a large enough sample size (Department for Communities and Local Government, 2018). Since in this study, the interest is not in the state of national wellbeing but rather relationships between wellbeing and other variables, no survey weights have been applied.

Datasets included in the analysis for this study were the *English Housing Survey, 2017: Housing Stock Data* (Department for Communities and Local Government, 2020) and the *English Housing Survey: Fuel Poverty Dataset, 2017* (Department for Business, 2019). These data cover the period April 2016 to March 2018. The files contain the paired sample primary 'raw' interview survey and physical survey data plus associated derived variables for all cases where a physical survey has been completed. All data are available on the UK Data Archive, a national data service that provides research access to a range of social and economic data collections (see for details UK Data Service, n.d.).

2.2. Data selection and sample size

Data sets varied in number of answers (see Figure 1). Some physical survey data had been collected in vacant dwellings; for those there is no corresponding interview data. For other questions, data from all household members was collected; we only use self-reported data from the household reference person (HRP) who is defined as *'The person in whose name the dwelling is owned or rented or who is otherwise responsible for the accommodation. In the*

case of joint owners and tenants, the person with the highest income is taken as the HRP. Where incomes are equal, the older is taken as the HRP.' (MHCLG, 2021, p. 55).

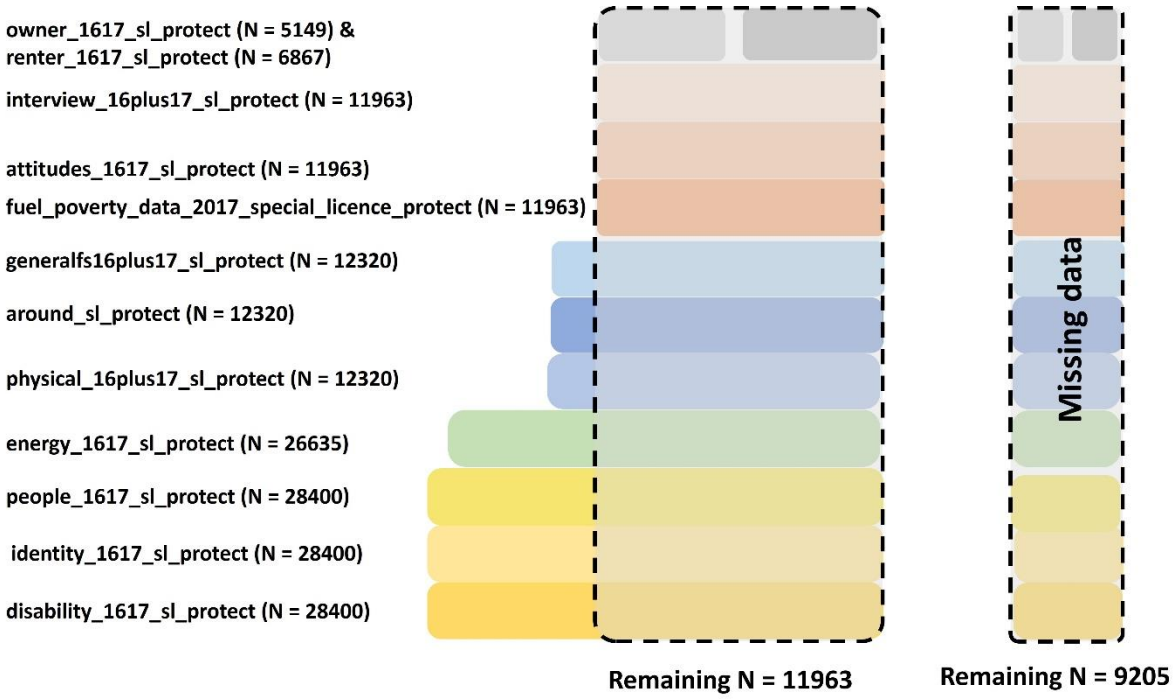

Figure 1. Data sets used and sample sizes at different stages. Note, original file names from the data set are given to allow others to easily identify which files were used.

The first step consisted of identifying for which respondents' data was available in all relevant data files. Note that the files "owner…" and "renter…" only had data for some respondents depending on their tenure; they were merged to create a new variable with answers for all N = 11963 respondents (see section 3.2.1 for details). Other 2768 respondents were excluded because of missing data on the wellbeing measures (the vast majority) or two predictors for which there was no logical way of merging the missing data with other response categories (General Health, Area Satisfaction; fewer than 30 cases). Hence, the final sample size for all regression analyses was N = 9205.

2.3. Variables

2.3.1. Wellbeing measures

Personal wellbeing was measured using four questions to be answered on a numerical scale (see Table 1). The questions were designed by the ONS to measure distinct aspects of personal wellbeing (Office for National Statistics, 2018c).

Table 1. Measures of personal wellbeing.

Question	**Answer options**
Overall, how satisfied are you with your life nowadays?	0 ("not at all") – 10 ("completely)
Overall, to what extent do you feel that the things you do in your life are worthwhile?	0 ("not at all") – 10 ("completely)
Overall, how happy did you feel yesterday?	0 ("not at all") – 10 ("completely)
Overall, how anxious did you feel yesterday?	0 ("not at all anxious") – 10 ("completely anxious")

2.4. Predictor variables

The variables were classified into three categories by the authors: personal factors, housing factors, and neighbourhood factors (see Tables 4, 5, and 6). The majority of variables were used as given by the EHS but three variables need to be developed for the analysis (see 3.2.1. for details). The first one is a variable indicating difficulty to pay mortgage or rent. The second variable is developed following a factor analysis on the variables indicating problems in the local area indicated by a surveyor. The third variable was a scale on "feeling of safety" based on the variables that measure perceived safety in the home, in the neighbourhood during the day, and the neighbourhood during night. In order to show the frequencies of all predictor variables together, descriptive information has been placed in the results section (3.2). Appendix C shows all variables and which EHS datafile they were derived from.

2.5. Hypotheses and statistical analysis

As per prespecification, ordinary Least Squares Regression Analysis would have been conducted if assumptions for it were met; since that was not the case (see 3.3 for more

details), logistic regression was used with the outcome variable dichotomized into lower and higher wellbeing. A separate analysis is conducted for each wellbeing measures, with the predictors being the same across all our regression models. The following hypotheses are tested for each wellbeing variable:

A. *Wellbeing is lower for occupants in dwellings with worse Energy Performance Certificate (EPC) ratings.*[1]
B. *Wellbeing is highest for occupants in detached homes and lowest in high-rise flats.*
C. *Wellbeing is lower for occupants who find it difficult to meet their heating/fuel costs.*
D. *Wellbeing is lower for occupants in fuel poverty (using the 10%) definition.*
E. *Wellbeing is lower for occupants in overcrowded households.*
F. *Wellbeing is lower for occupants unable to keep their living room at comfortable temperatures.*
G. *Wellbeing is lower for occupants in dwellings with higher repair costs per square meter*.
H. *Wellbeing is lower for occupants who live in areas with more problems.*
I. *Wellbeing is lower for occupants who are less satisfied with their environment.*
J. *Wellbeing is lower for occupants who feel less safe in their local environment.*
K. *Wellbeing is lower for occupants who live in more deprived areas.*
L. *Wellbeing is lower for occupants with damp problems*. (not prespecified, added after the literature review)

Additionally, we build separate regression models to gauge the relative importance of personal factors, housing factors, and neighbourhood factors; i.e., an individual model for each class of predictor. We then create combined models, creating a personal and housing factors model (*Personal&Housing*), and one with personal, housing, and neighbourhood factors (*Personal&Housing&Neighbourhood*). We use ANOVAs to test if the additional variables improve the model. Specifically, we will compare *Personal* against *Personal&Housing*, and *Personal&Housing* against *Personal&Housing&Neighbourhood*. For Life Satisfaction, we also test to what extent publicly available / existing data can explain variance in wellbeing; starting with a model encompassing Energy Performance Certificate

[1] A is the highest EPC rating, indicating the most energy efficient dwelling, and G the lowest EPC rating, indicating the least energy efficient building.

(EPC) data, then adding area level information, and finally Census data; again using ANOVAs to test if additional variables improve the model (not prespecified).

Analyses were conducted using the R Statistical language (R Core Team, 2020), version 4.0.3. *Tidyverse* was used for data wrangling (Wickham et al., 2019); *ggplot2* for basic plotting (Wickham, 2016). The core analysis was done using the stats library. *SjPlot* was used for calculation of odds ratios in the logistic regression and for visualization of the logistic regression results (Lüdecke, 2021), see (*insert link after peer-review)* for any other packages used. For logistic regression, there are several Pseudo R^2 estimates. Here, we used the estimate as developed by Tjur (Tjur, 2009). Minor deviations from the prespecification are noted in Appendix D; the two major changes (added hypothesis L, one exploratory analysis in 3.5) are noted in the main manuscript.

2.6. Bias

Bias can occur at different stages of any research project (Pannucci & Wilkins, 2010). As the study constitutes secondary data analysis, a number of possible biases were outside the control of the researchers, such as around selection bias or interviewer bias (see the EHS report on possible bias: Department for Communities and Local Government, 2014). To mitigate the risk of confirmation bias, i.e. a bias related to researchers searching for and interpreting information in a way that confirm their prior ideas or opinions, all analysis was pre-specified.

3. Results

3.1. Descriptives

3.1.1. Wellbeing levels

Figure 2 shows the distribution of responses to the four wellbeing questions.

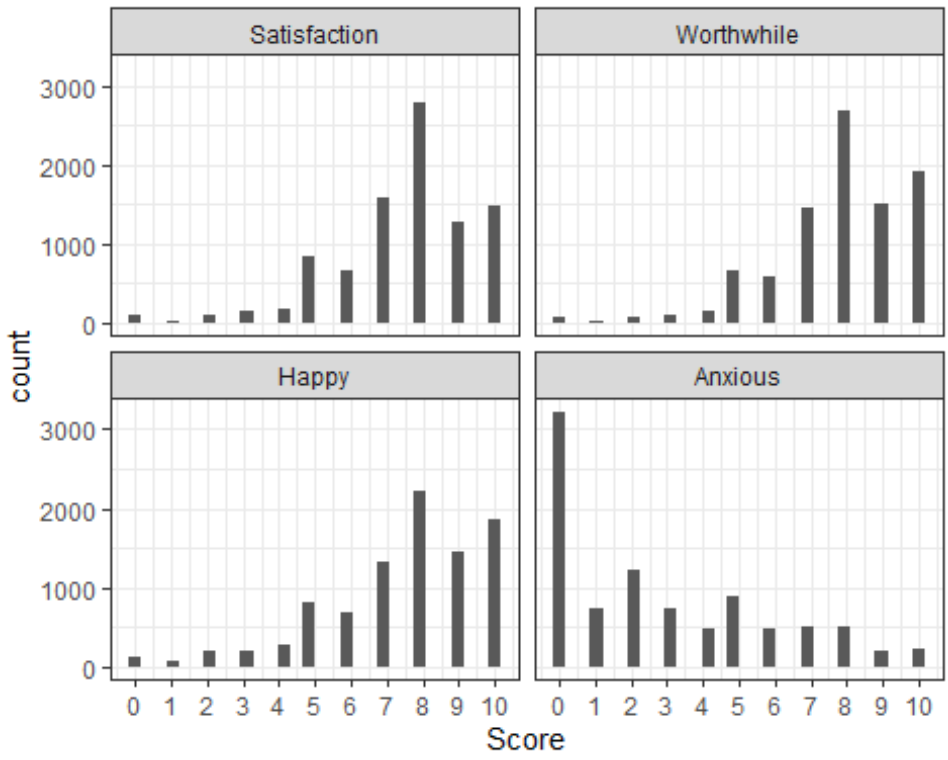

Figure 2. Responses to the four wellbeing measures.

For life satisfaction, worthwhile, and happy, a higher score indicates higher wellbeing; for anxiety, a high score means greater anxiety, and hence lower wellbeing. Across all measures, answers were skewed to higher wellbeing. Mean values were: life satisfaction M = 7.52, worthwhile M = 7.82, happy M = 7.48, anxious = 2.85; the corresponding medians were 8 for the first three measures, and 2 for anxious.

In an independently carried out nationally representative survey covering the same time period (Office for National Statistics, 2019), the mean values were very similar: 7.68 (life satisfaction), 7.87 (worthwhile), 7.52 (happy), and 2.91 (anxious).

All variables correlated weak to moderately, but significantly with each other (Table 2).

Table 2

Means, standard deviations, and correlations with confidence intervals for the wellbeing measures.

Variable	*M*	*SD*	1	2	3
1. Life satisfaction	7.52	1.97			
2. Worthwhile	7.82	1.88	.64**		

			[.63, .65]		
3. Happy	7.48	2.23	.58** [.56, .59]	.52** [.51, .54]	
4. Anxious	2.84	2.95	-.36** [-.37, -.34]	-.29** [-.31, -.28]	-.46** [-.48, -.45]

Note. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates $p < .05$. ** indicates $p < .01$.

3.2. Predictor variables

3.2.1. Developed predictor variables

Three variables were developed as predictors for the subsequent regression analysis.

Arrears

Amongst home owners, 16 respondents *found it very difficult to keep up* and 85 *have found it rather difficult to keep up*. These two categories were combined into one. The answer *have had no difficulty in keeping up* was given by 2200 respondents. For 2848 respondents the question was not applicable or no answer was given. Amongst renters, 365 indicated being behind with rent payments; 2786 were not. For 3716 respondents the question was not applicable or no answer was given. A combined indicator of renters and owners was constructed, coded as “difficulty”, “no difficulty”, “not valid answer”.

Problems in neighbourhood

The second variable is developed following a factor analysis on the 17 variables indicating the extent of problems in the local area as indicated by the surveyor. Bartlett’s test of sphericity, [Bartlett $Chi^2 = 68368.21$ (136), $p < 0.001$] indicated that correlations between variables were sufficient for a PCA. Initially, an exploratory PCA was run. Three eigenvalues were greater than 1, the criterion that had been prespecified to be used for variable selection. However, the fourth-largest eigenvalue was very close to 1, with 0.97, and factor interpretation was simpler for four factors. Hence, four components were extracted.

We then reran the PCA set to four components and with “varimax” rotation. Table 3 shows the factor loadings for the four factors and the label given to them.

Table 3. Factor loadings from the Principal Component Analysis.

	Scruffy	**Traffic**	**Vacant**	**Behaviour**
Litter	0.44	0.16	-0.09	-0.64
Graffiti	0.13	0.18	-0.28	-0.81
Vandalism	0.16	0.16	-0.31	-0.80
DogExcrement	0.36	0.14	-0.09	-0.64
ConditionDwellings	0.66	0.26	-0.20	-0.33
VacantSites	0.13	0.12	-0.74	-0.25
IntrusiveIndustry	0.16	0.31	-0.71	-0.01
NonConformUse	0.19	0.23	-0.69	-0.13
VacantBuildings	0.14	0.02	-0.65	-0.32
AirQuality	0.29	0.63	-0.14	-0.23
HeavyTraffic	0.12	0.82	-0.13	-0.10
IntrusionMotorways	0.15	0.82	-0.14	-0.07
RailAirNoise	0.12	0.49	-0.21	-0.15
NuisanceParking	0.70	0.12	-0.11	-0.05
ScruffyGardens	0.72	0.13	-0.15	-0.39
ScruffyBuildings	0.68	0.24	-0.25	-0.33
ConditionRoads	0.72	0.13	-0.13	-0.12
Cronbach's α	*0.83*	*0.75*	*0.74*	*0.83*

varimax-rotation

The first factor (“Scruffy”,) was related to the general state of the environment. The second related to traffic and air pollution issues (“Traffic”). The third was slightly harder to describe; it covered non-conform building use and vacant sites (“Vacant”). The fourth one showed the extent of which problems due to human behaviour occured (“Behaviour”).

We calculated the mean score for each factor by averaging the value from the individual variables it was composed of.

Feeling of safety

We calculated Cronbach's alpha for the variables *nhhmsf1, nhsfday, nhsfnte* that stood for the perceived feeling of safety when alone at home, outside during the daytime and outside at night. Only answers between 1 and 5 were retained for analysis as the category of "not doing something for other reasons than safety" was not interpretable, neither was missing data.

The number of households with valid data was N = 5057. Cronbach's alpha was 0.7; hence, the internal consistency was exactly at the usually required and here prespecified value of 0.7, and the variables were combined into a scale. The mean value across the safety items was calculated and rounded to the nearest integer to reflect the category labels.

3.2.2. Descriptives of all predictor variables

Table 4, 5, 6 show the frequencies for the predictor variables, categorized into the three classes of personal factors, housing factors, and neighbourhood factors. The reference category was either the largest category or had been prespecified in case of particular theoretical interest *(link inserted after review).*

We checked for categories with a count of less than <30. For ethnicity, this meant that Chinese was merged with the second-smallest category of ethnicity, 'other Asian'. For Marital Status, the category of "current or former same-sex civil partnership" was merged with the second-smallest category of "separated but still legally married" to an "OtherStatus" category.

For all variables, the first category in Table 4 is the reference category for the subsequent regression.

Table 4. Description of personal variables. First category is the reference category in subsequent regression analysis (printed in italics).

Variable	Categories	Freqs (% of Valid)
GeneralHealth	*Very_Good*	2515 (27.3%)
	Good	3443 (37.4%)
	Fair	2229 (24.2%)
	Bad	777 (8.4%)
	Very_Bad	244 (2.6%)

MaritalStatus	*Married*	2988 (32.5%)
	Single	2954 (32.1%)
	Divorced	1591 (17.3%)
	Widowed	1190 (12.9%)
	Other_Status	485 (5.3%)
HighestQual	*Degree_Level*	2038 (22.1%)
	Other_Qual	4008 (43.5%)
	Not_Asked	3162 (34.3%)
AgeHRP	*45_54*	1643 (17.8%)
	16_24	316 (3.4%)
	25_34	1294 (14.1%)
	35_44	1483 (16.1%)
	55_64	1529 (16.6%)
	65OrOver	2943 (32.0%)
SexHRP	*Female*	4693 (51.0%)
	Male	4515 (49.0%)
EmploymentHRP	*Full-Time_Work*	3377 (36.7%)
	Retired	3002 (32.6%)
	Other_Inactive	1232 (13.4%)
	Part-Time_Work	1121 (12.2%)
	Unemployed	346 (3.8%)
	Full-Time_Education	130 (1.4%)
AHCeqvIncome	*5th_(Highest)*	1500 (16.3%)
	4th	1609 (17.5%)
	3rd	1754 (19.0%)
	2nd	2194 (23.8%)
	1st_(Lowest)	2151 (23.4%)
Househ_Type	*Single=>60yrs*	2079 (22.6%)
	Couple_W/_Dep	1418 (15.4%)
	Single<60yrs	1409 (15.3%)
	Couple>=60yrs	1319 (14.3%)
	Single_Parent_W/_Dep	1119 (12.2%)
	Couple<60yrs	1062 (11.5%)
	Other_Multiperson	802 (8.7%)
EthnicityHRP	*White*	8133 (88.3%)
	Black	395 (4.3%)
	Indian	154 (1.7%)
	Pakist_Bangla	152 (1.7%)
	Mixed	139 (1.5%)
	Other	131 (1.4%)
	Chinese_Other_Asian	104 (1.1%)

Table 5 shows the frequencies or descriptive statistics, respectively, for housing variables.

Table 5. Description of housing variables. For categorical variables, the first category is the reference category in subsequent regression analysis (printed in italics).

Variable	Categories /	Freqs (% of Valid) / Statistics
BedroomStandard	*At_Standard*	3388 (36.8%)
	1Below	407 (4.4%)
	=>2Below	66 (0.7%)
	1Above	2825 (30.7%)
	=>2Above	2522 (27.4%)
Arrears	*No*	3624 (39.4%)
	NA	5224 (56.7%)
	Yes	360 (3.9%)
Tenure	*Own_Outright*	2223 (24.1%)
	Rent_Housing_Association	2091 (22.7%)
	Rent_Private_Unfurn.	1628 (17.7%)
	Own_Mortgage	1503 (16.3%)
	Rent_Local_Authority	1492 (16.2%)
	Rent_Private_Furn.	271 (2.9%)
LRWarm	*Yes*	8051 (87.4%)
	No	1016 (11.0%)
	Don’t_Know	141 (1.5%)
HeatingCost	*Very_Easy*	2633 (28.6%)
	Fairly_Easy	3532 (38.4%)
	Neither	1499 (16.3%)
	Fairly_Difficult	1058 (11.5%)
	Very_Difficult	405 (4.4%)
	Dont_Know	81 (0.9%)
FuelPovertyLIHC	*Not_In_FP_LIHC*	8129 (88.3%)
	In_FP_LIHC	1079 (11.7%)
FuelPovertyIncome	*Not_In_FP*	8457 (91.8%)
	In_FP	751 (8.2%)
DwellingType	*Detached_House*	1041 (11.3%)
	Semi-Detached_House	2046 (22.2%)
	Purpose_Built_Flat_Low_Ri	1923 (20.9%)
	Medium/Large_Terraced_Hou	1571 (17.1%)
	Small_Terraced_House	1042 (11.3%)
	Bungalow	949 (10.3%)
	Converted_Flat	389 (4.2%)
	Purpose_Built_Flat_High_R	247 (2.7%)
EPC	*D*	4429 (48.1%)
	B	135 (1.5%)
	C	3218 (34.9%)
	E	1071 (11.6%)
	F	272 (3.0%)
	G	83 (0.9%)
DecentHome	*Decent*	7571 (82.2%)
	Non-Decent	1637 (17.8%)
CostUrgentRepair	na	Mean (SD): 8.9 (23.7) Median (IQR): 0 (8.1)
CostBasicRepair	na	Mean (SD: 14.3 (30.6) Median (IQR): 2.3 (16.1)

CostComprRepair	na	Mean (SD): 45.5 (76.3) Median (IQR): 12.7 (62.6)
DwellingAge	*1965_To_1980*	2086 (22.7%)
	Pre_1919	1591 (17.3%)
	1919_To_1944	1187 (12.9%)
	1945_To_1964	2108 (22.9%)
	1981_To_1990	783 (8.5%)
	1991_To_2002	715 (7.8%)
	Post_2002	738 (8.0%)
FloorArea	*50_To_69_Sqm*	2830 (30.7%)
	110_Sqm_Or_More	1263 (13.7%)
	90_To_109_Sqm	1043 (11.3%)
	70_To_89_Sqm	2437 (26.5%)
	Less_Than_50_Sqm	1635 (17.8%)
Garden	*Private_Plot*	7021 (76.2%)
	Shared_Plot_Only	2060 (22.4%)
	Neither	127 (1.4%)
Damp	*No_(+)*	6203 (67.4%)
	Yes_All_year	1484 (16.1%)
	Yes_Winter	1323 (14.4%)
	Yes_Other	198 (2.2%)

Table 6. Description of neighbourhood variables. For categorical variables, the first category is the reference category in subsequent regression analysis (printed in italics).

Variable	**Categories**	**Freqs (% of Valid) / Statistics**
Area_Satisf	*Very_Satisfied*	5192 (56.4%)
	Fairly_Satisfied	2748 (29.9%)
	Neither	520 (5.6%)
	Slightly_Dissatisfied	485 (5.3%)
	Very_Dissatisfied	260 (2.8%)
IMDDeciles	*1st_(Most)*	1416 (15.4%)
	2nd	1171 (12.7%)
	3rd	1055 (11.5%)
	4th	968 (10.5%)
	5th	917 (10.0%)
	6th	843 (9.2%)
	7th	755 (8.2%)
	8th	740 (8.0%)
	9th	723 (7.9%)
	10th_(Least)	617 (6.7%)
GOREHS	*South_East*	1437 (15.6%)
	North_West	1315 (14.3%)
	East	1169 (12.7%)
	London	1158 (12.6%)
	Yorkshire_And_The_Humber	1068 (11.6%)
	West_Midlands	872 (9.5%)

	South_West	841 (9.1%)
	East_Midlands	774 (8.4%)
	North_East	571 (6.2%)
Morphology	*Urban_>_10k*	7646 (83.1%)
	Town_Fringe	870 (9.5%)
	Village	466 (5.1%)
	Hamlet	223 (2.4%)
Scruffy1	na	Mean (SD): 1.7 (0.6)
		Median (IQR): 1.6 (0.8)
Traffic2	na	Mean (SD): 1.5 (0.5)
		Median (IQR): 1.2 (0.8)
Vacant3	na	Mean (SD): 1.1 (0.3)
		Median (IQR): 1.0 (0.0)
Behaviour4	na	Mean (SD): 1.4 (0.5)
		Median (IQR): 1.2 (0.5)
Safety_All	*Very_Safe*	2727 (29.6%)
	Fairly_Safe	1810 (19.7%)
	A_Bit_Unsafe	447 (4.9%)
	Very_Unsafe	71 (0.8%)
	No_Answer	4150 (45.1%)

As a control variable the year of data collection was coded as categorical variable. In 2016, 2937 interviews took place, in 2017 (reference category), 5202, and in 2018, 1066.

3.3. Explaining wellbeing

Initially four ordinary least squares linear regressions were run, one for each of the outcome variable. All VIFs were less than 10 and all Cook's distances were less than 1.0. However, inspection of the plot of fitted values against residuals and the Breusch-Pagan test (all $p < .001$) indicated substantial hetereoskedasticity; i.e. the variance of the residuals varied across the values of wellbeing. Hence, we used logistic regression instead, dichotomizing the outcome variable into "low/medium" wellbeing and "high / very high" wellbeing. A dichotomy was created instead of using the four possible outcome categories of low, medium, high and very high, given the paucity of data points for low wellbeing (see Figure 2). There is debate on whether dichotomizing continuous data is adequate (Kuss, 2013; Maccallum et al., 2002; Snijders & Bosker, 2011); however, given that ONS presents the variables routinely as categories and the ease of interpretation of a logistic regression compared to multiple transformed variables, we proceeded with logistic regression[2]. As per ONS guidance, 0-4 are

[2] Log-transforming the outcome variable, here, the four wellbeing measures can also be used to reduce heteroskedasticity but here, three out of the four models still had significant heteroskedasticity as established by the Breusch-Pagan-test.

low, 5-6 medium,7-8 high, and 9-10 very high wellbeing scores for Life Satisfaction, Worthwhile and Happy. For anxious, 0-1 are very low, 2-3 low, 4-5 medium, and 6-10 high.

Overall model fit was greatest for life satisfaction (Tjur's R^2 = 0.227), followed by worthwhile (Tjur's R^2 = 0.175), happy (Tjur's R^2 = 0.115) and lowest for anxiety (Tjur's R^2 = 0.100). Figure 3 shows the odds ratios for those variables that were significant in at least one regression model.

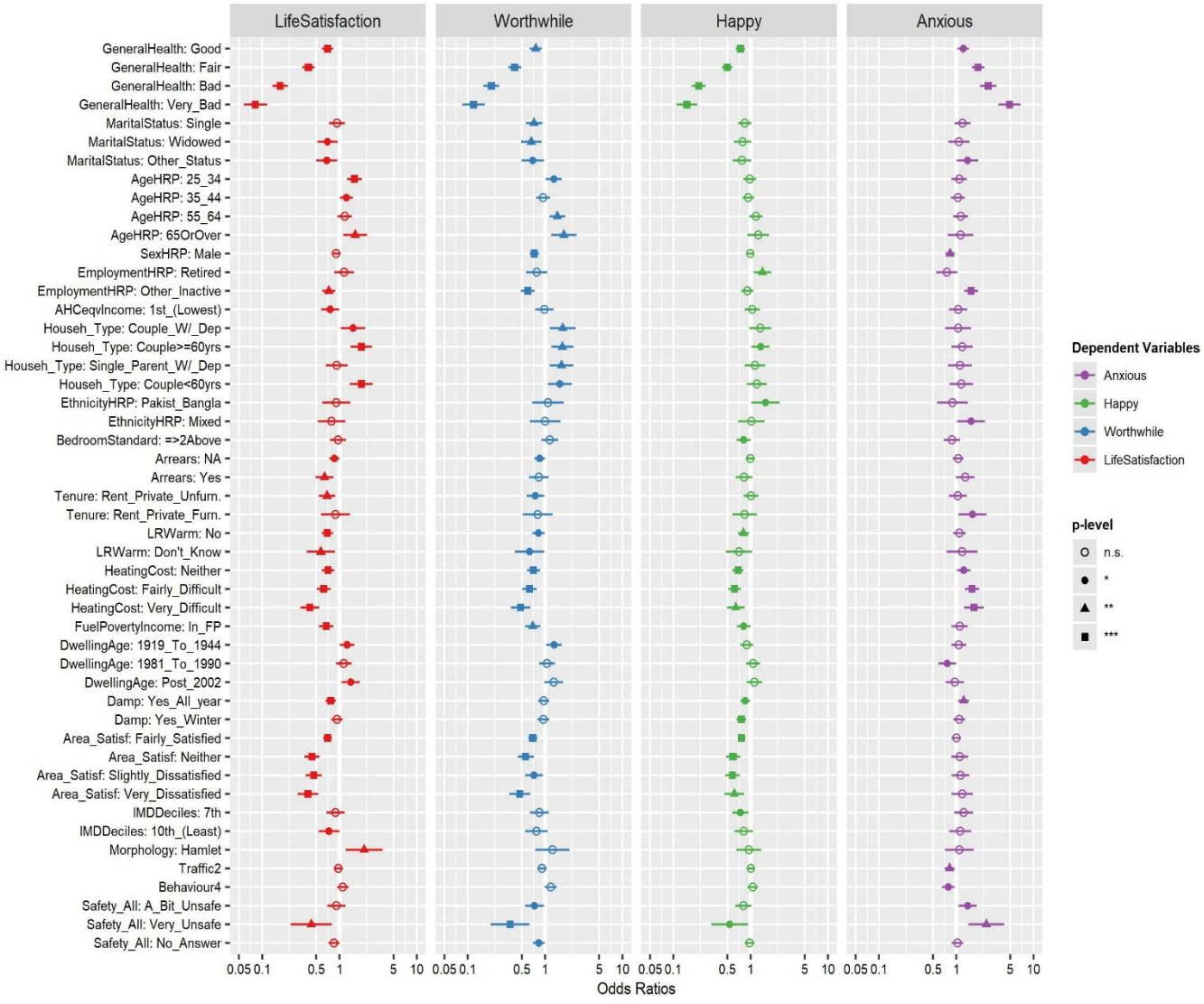

Figure 3. Odds ratio and 95% confidence intervals for those variables significant in at least one regression mode.

The full regression model is presented in Appendix E.

Table 7 show the various hypotheses and whether they are supported for the four outcome variables. Since as per ONS guidance the four outcome variables are to be treated separately, no family-wise error correction was planned. However, given that the four wellbeing measures correlated with each other, we additionally performed Holm-Bonferroni correction of *p*-values. The Holm–Bonferroni method sorts the *p*-values from lowest to highest and then compares them sequentially against adjusted alpha values (Sture Holm, 1979). Here, $\alpha = .05$, and number of hypotheses k = 4; i.e. the adjusted alpha levels are $\alpha_1 = .013$ (.05/4), $\alpha_2 = .017$ (.05/3), $\alpha_3 = .025$ (.05/2) and $\alpha_4 = .05$ (.05/1). Where the Holm-Bonferroni adjustment would lead to the conclusion of a non-significant result, this is noted in the table with 'HB ns'. The table is reproduced in Appendix F with numerical *p* values.

*Table 7. Overview of the hypotheses and whether they were confirmed for the different outcome variables. (asterisks indicate significance level: *<.05; ** < .01; *** < .001)*

	Outcome variable			
Hypothesis	*Life Satisfaction*	*Worthwhile*	*Happy*	*Anxious*
Wellbeing is lower for occupants in dwellings with lower EPC ratings.	ns	ns	ns	ns
Wellbeing is highest for occupants in detached homes and lowest in high-rise flats.	ns	ns	ns	ns
Wellbeing is lower for occupants who find it difficult to meet their heating/fuel costs.				
very difficult vs easy	***	***	***	***
Fairly difficult vs. easy	***	***	***	***
Neither vs. easy	***	***	***	*
Wellbeing is lower for occupants in fuel poverty (using the 10%) definition.	***	**	* (HB ns)	ns
Wellbeing is lower for occupants in overcrowded households.	ns	ns	ns	ns
Wellbeing is lower for occupants unable to keep their living room at comfortable temperatures.				
No vs. yes	***	* (HB ns)	**	ns
Don't know vs. yes	**	* (HB ns)	ns	ns
Wellbeing is lower for occupants in dwellings with higher repair costs per square meter.	ns	ns	ns	ns
Wellbeing is lower for occupants who live in areas with more problems.				
Scruffy	ns	ns	ns	ns
Traffic	ns	ns	ns	* contrary
Vacant	ns	ns	ns	ns
Behaviour	ns	ns	ns	* contrary
Wellbeing is lower for occupants who are less satisfied with their environment.				
Very dissatisfied vs. satisfied	***	***	***	ns
Slightly dissatisfied vs. satisfied	***	***	***	ns
Neither vs. satisfied	***	***	***	ns
Fairly satisfied vs. satisfied	***	***	***	ns
Wellbeing is lower for occupants who feel less safe in their local environment.				
Very unsafe vs. safe	**	***	*	***
A bit unsafe vs. safe	ns	* (HB ns)	ns	* (HB ns)
Wellbeing is lower for occupants who live in more deprived areas.	ns	ns	ns	ns
Wellbeing is lower for occupants with damp problems.				
Year round vs. no	**	ns	*	**
Winter only vs. no	ns	ns	***	ns

3.4. Comparing the effect of the different predictor categories

In a next step, we assessed model fit separately for personal, housing, and neighbourhood variables. Since the year of data selection was not significant in any of the four regression analyses and cannot be classified into the three categories, it was omitted.

Personal factors on their own explained most of the variance across the outcome variables; followed by housing factors and neighbourhood factors (see Table 8).

Table 8. R^2 Tjur for the four outcome variables across the three models.

	Model		
	Personal	*Housing*	*Neighbourhood*
Life Satisfaction	0.18	0.12	0.06
Worthwhile	0.14	0.08	0.04
Happy	0.09	0.05	0.03
Anxious	0.07	0.04	0.02

We tested if adding housing to personal variables, and further adding neighbourhood variables significantly improved the model fit using ANOVAs. For all outcome variables, adding housing and neighbourhood variables to the personal model, improved model fit substantially (all $p < .001$); see Appendix G for details.

3.5. Variables to target wellbeing initiatives

The analysis up to this point showed that various variables can explain variation in wellbeing measures, in particular those based on self-report of householders. Whilst these findings can grow our understanding of relationships between built environment and wellbeing, they

would not necessarily help in targeting households most likely to experience low wellbeing as many of the variables are not observable from the outside. Hence, in a final part of the analysis (not prespecified), we identified how well existing data can be used for targeting wellbeing interventions. The first model encompasses EPC data which is publicly available and includes the EPC rating, floor area and dwelling type for a dwelling. The second model consists of Index of Multiple Deprivation (IMD), the Government Office region and the morphology. These data exist on an area basis and apply to any dwelling in that area. The third model consists of the census data which exists on a per-dwelling resolution but is not made publicly available but in theory could be used for analysis and targeting. The variables for this model are: “GeneralHealth”, “MaritalStatus”, “HighestQual”, “AgeHRP”, “SexHRP”, “EmploymentHRP”, “Househ_Type”, “EthnicityHRP”, “Tenure”, “BedroomStandard”. Figure 4 shows how well the models perform to explain variation in Life Satisfaction (analysis not conducted for the other three wellbeing measures).

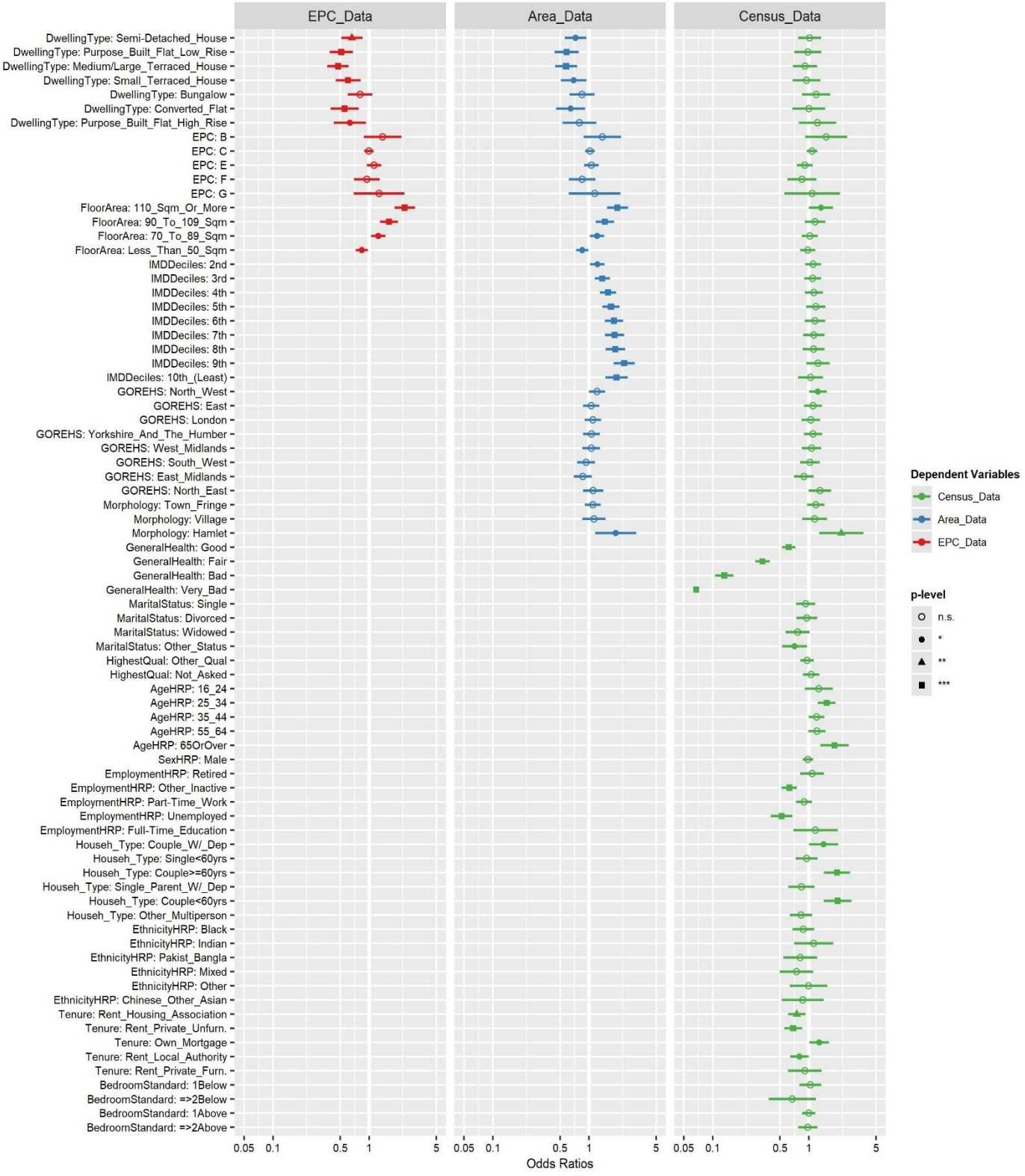

Figure 4. Odds Ratios for an EPC data model, with area and personal data added in model 2 and model 3, respectively.

Adding model 2 and model 3 add significantly to the previous models (Table 9); R^2 Tjur = 0.027 for the EPC model, 0.038 for the area model, and 0.184 when adding census data.

Table 9. Results of model comparison.

		LifeSatisfaction			
Model	Resid Df	Resid Dev	DF	Deviance	p
EPC	9188	9554.1			
EPC & Area	9168	9450.1	20	104.05	<.001
EPC & Area & Census	9126	8199.2	46	1250.89	<.001

The analysis also shows that the effect of a predictor varies depending on which other factors are included. For example, building type and IMD are significant in the EPC & area model but not when adding census variables.

4. Discussion

This study used a large sample (N = 9205) from the English population to understand which personal, housing, and neighbourhood variables are related to subjective wellbeing. The results indicate that across all outcome measures, personal variables play the greatest role for wellbeing. Both housing and neighbourhood variables increase the amount of explained variance but overall, only a modest amount of variance is explained. Hypotheses on lower wellbeing when having greater difficulty in meeting fuel cost, being unable to keep the living room warm, being in fuel poverty (10% definition) were largely supported (though to a lesser extent for anxiousness). For neighbourhood variables, lower satisfaction and lower perceived safety were associated with lower wellbeing.

The findings add to an existing body of literature on the link between the built environment and mental health. However, to our knowledge this was the first study to systematically link built environment and wellbeing as measured by the four questions of the Office for National Statistics.

Previous reviews had indicated conflicting evidence on the link between housing quality and mental health (Clark et al., 2007; R Cooper et al., 2008; Evans, 2003; Singh et al., 2019). A detailed look at the underlying studies shows substantial heterogeneity in how the outcome variable was operationalized and how housing quality was defined. Blair (Blair et al., 2011) reported a link of an objective marker of stress, i.e. cortisol level and a composite definition

of housing quality. Curl et al. (Curl et al., 2015) showed that fabric improvements had positive mental health outcomes. However, Evans (2002) and Clark et al (2007) judge there to be insufficient evidence to assume a causal relationship. Using a very different mental health indicator, here self-reported wellbeing, we also showed a link to housing quality in terms of being able to keep the home warm, including the financial affordability, but only through cross-sectional data. Experiencing damp was also liked to lower wellbeing though not for all operationalizations of wellbeing.

Dissatisfaction with the local environment and feeling very unsafe were strongly linked to lower wellbeing in our study, in line with previous reports (Clark et al., 2007; Hunter et al., 2019); however, it is worth pointing out that surveyor assessed conditions of the local environment were not clearly related to wellbeing (see also 4.2 for further discussion of this point).

Being in arrears with rent or mortgage payment showed a link to life satisfaction only; Alley et al (Alley et al., 2011) had reported a link to depressive symptoms in a sample limited to Americans over 50 years.

Private renters living in unfurnished homes experienced greater anxiety; similar to what Kang et al (Kang et al., 2016) reported generally for renters; though here respondents who lived in social housing or rented privately but furnished did not show a significant association and Kang's study only looked at older citizens.

Overcrowding did not have a significant link to wellbeing in our study whereas e.g. Sadowski et al (Sadowski et al., 1999) reported such an effect; however, they looked at experience of overcrowding in early childhood and its effect on later life whereas our study is cross-sectional.

In summary, our study supports a range of existing findings and adds evidence on a new conceptualization around being able to keep the home warm; however, as discussed in 4.2. there are a number of methodological challenges that limit comparability of different studies.

4.1. Policy implications

This paper highlights that targeting householders based on readily observable characteristics is not an easy task. Given the strong relationship between health and wellbeing (Department of Health, 2014) targeting via medical records might be the most promising avenue; however,

only for those who sought treatment for a medical condition. When not controlling for other factors, smaller dwellings and a lower IMD are associated with lower wellbeing. Wellbeing was highest in detached dwellings. Hence, those variables might be the most suitable for targeting wellbeing interventions in the absence of other information. However, it needs to be emphasized that these variables largely lost their significant association when controlling for other variables which is irrelevant for targeting purposes but indicates careful consideration needs to be given as to what underlies lower wellbeing.

The results also highlight that wellbeing is correlated with personal, building, and neighbourhood variables. Focusing only on one of those areas will hence be limited in how much it can change wellbeing.

Finally, the study showed that variables around homes being financially difficult to heat and keep warm were highly significant predictors. Previous work has linked cold homes to negative health outcomes (Jevons et al., 2016). Fuel poverty is here shown to correlate negatively with wellbeing, supporting the established connection between fuel poverty and mental health (Liddell & Morris, 2010). Interestingly, only the 10% definition, i.e. when a household is unable to obtain adequate energy services for 10% of its income (Boardman, 1991), showed this significant association. The later low income–high costs definition in which a household is seen as fuel poor if it has required fuel costs that are above the national median average and the household would be left with an income below poverty line if it spent that amount on fuel (Hills, 2012), had no significant effect. Fuel poverty is a complex issue (Baker et al., 2018). How to best define it is beyond the scope of the paper; however, it is crucial to note that different definitions identify different households as fuel-poor and hence to avoid generic statements about the effect of fuel poverty.

4.2. Methodological implications

Three important methodological considerations are highlighted by this study.

a) Results differ depending on what covariates are being included, and currently, there is no clear guidance on which ones to include. This is problematic if conclusions differ substantially depending on which other variates are included in analyses. A prominent example from this study is dwelling type. Controlling for other variables, dwelling type has no significant effect on wellbeing; however, in the model with fewer covariates it does and previous research has also indicated a role of dwelling type (Rachel Cooper et al., 2008; Evans, 2003). Greater focus on theoretical models to

underpin links between the built environment and wellbeing would help in choosing appropriate covariates.

b) The same variables are differently associated with different wellbeing measures. As the ONS states, the four measures reflect different facets of wellbeing. Hence to find that different factors are differentially associated is congruent with this view. However, it raises the issue that if we want to gain a comprehensive understanding of relationship of the built environment and wellbeing with the aim of intervening to improve wellbeing, we need to be clear on what wellbeing is and which of its multifaceted aspects is to be addressed. This study focused on subjective wellbeing, i.e. a person's evaluation of their personal situation; however, objective approaches to measuring wellbeing exist such as using GDP or composite indicators encompassing health, job opportunities, socioeconomic development, environment, safety, and politics (Voukelatou et al., 2020). This raises further complexity around the relationship between the two, with subjective wellbeing focused on personal experience and objective measures on societal wellbeing. Linking built environmental measures to objective wellbeing will necessarily lose granularity of findings as not the 1:1 relationship of a person and dwelling is assessed.

c) Overall understanding of what predicts wellbeing is low. This study did not control for all variables shown to play a role in earlier studies (see 4.3) which might partly explain this overall low understanding. However, it also raises the question if important predictors are overlooked, such as personality (Biswas-Diener et al., 2004). It is also possible that some important variables of the built environment have not been included or measured suboptimally. For example, this study has only included proxies for outdoor air pollution. Measuring pollutants might help to explain more variance in wellbeing. Also, overcrowding was assessed using an official indicator and not through subjective perception which might be more relevant (Guite et al., 2006).

4.3. Limitations and future research

This study was cross-sectional in nature and can only show correlational evidence. Future research on causal relationships between the built environment and wellbeing is needed as stated previously (e.g. Clark et al., 2007; Hunter et al., 2019; Moore et al., 2018).
Some variables that might have an impact on wellbeing were not available for this study,

such as neighbourhood violence (Clark et al., 2007) or access to recreational spaces (Guite et al., 2006; Hunter et al., 2019; Krefis et al., 2018). Furthermore, the study had a relatively narrow focus on housing conditions in high-income countries; also in the literature review, evidence e.g. around housing in slums was not considered and the findings cannot contribute to the rich body of literature in that field (e.g. Henson et al., 2020; Turley et al., 2013).

This study relied on self-report and surveyor assessment but did not include measured parameters such as temperatures in dwellings, noise or air pollution which might be important (Hoisington et al., 2019) and should be included in future research.

The data was based on a large national survey. Since rented properties were oversampled, the wellbeing data is not fully nationally representative; however, as shown in 3.1, the average wellbeing ratings in this study were very similar to a fully representative sample. Additionally, results on the links between predictors and wellbeing should be generalizable to other households in England with the same characteristics, given the underlying sample size and initial sampling strategy (Department for Communities and Local Government, 2018). However, the longevity of results is unclear; especially the experience of the Covid-19 pandemic can have substantially altered the relationship between the built environment and subjective wellbeing.

5. Conclusion

This study based on a large sample size showed that personal, housing and neighbourhood factors are all correlated with subjective wellbeing which highlights the need to focus on multiple aspects to promote well-being. It also showed that targeting householders with lowest wellbeing based on publicly available data would be challenging. The effects of variables are not constant across all four wellbeing measures used which raises the question ‘which wellbeing’ should be addressed. Finding it difficult to keep the living room warm, being in fuel poverty and struggling with meeting fuel costs were the housing variables most consistently associated with lower wellbeing. Low area satisfaction and low perceived safety were neighbourhood variables associated with lower wellbeing. Finally, the research community needs to address methodological challenges around identifying the most appropriate covariates, defining wellbeing and considering the measurement of key variables.

References

Alley, D. E., Lloyd, J., Pagán, J. A., Pollack, C. E., Shardell, M., & Cannuscio, C. (2011). Mortgage Delinquency and Changes in Access to Health Resources and Depressive Symptoms in a Nationally Representative Cohort of Americans Older Than 50 Years. *American Journal of Public Health*, *101*(12), 2293. https://doi.org/10.2105/AJPH.2011.300245

Baker, K. J., Mould, R., & Restrick, S. (2018). Rethink fuel poverty as a complex problem. In *Nature Energy* (Vol. 3, Issue 8, pp. 610–612). Nature Publishing Group. https://doi.org/10.1038/s41560-018-0204-2

Benchimol, E. I., Smeeth, L., Guttmann, A., Harron, K., Moher, D., Peteresen, I., Sørensen, H. T., von Elm, E., & Langan, S. M. (2015). The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine*, *12*(10), 1001885. https://doi.org/10.1371/journal.pmed.1001885

Biswas-Diener, R., Diener, E., & Tamir, M. (2004). The psychology of subjective well-being. *Daedalus*, *133*(2). https://doi.org/10.1162/001152604323049352

Blair, C., Raver, C. C., Granger, D., Mills-Koonce, R., & Hibel, L. (2011). Allostasis and allostatic load in the context of poverty in early childhood. *Development and Psychopathology*, *23*(3), 845–857. https://doi.org/10.1017/S0954579411000344

Boardman, B. (1991). Fuel poverty : from cold homes to affordable warmth. In *Belhaven*. Belhaven.

Chu, A., Thorne, A., & Guite, H. (2004). The impact on mental well-being of the urban and physical environment: an assessment of the evidence. In *Journal of Public Mental Health* (Vol. 3, Issue 2, pp. 17–32). https://doi.org/10.1108/17465729200400010

Clark, C., Myron, R., Stansfeld, S., & Candy, B. (2007). A systematic review of the evidence on the effect of the built and physical environment on mental health. *Journal of Public Mental Health*, *6*(2), 14–27. https://doi.org/10.1108/17465729200700011

Cooper, R, Boyko, C., & Codinhoto, R. (2008). State-of-Science Review: SR-DR2. The Effect of the Physical Environment on Mental Wellbeing. In *Foresight Mental Capital and …* (Issue February 2014). https://www.researchgate.net/profile/Christopher_Boyko/publication/237422888_State-

of-Science_Review_SR-DR2_The_Effect_of_the_Physical_Environment_on_Mental_Wellbeing/links/0a85e52f8af680e280000000.pdf

Cooper, Rachel, Boyko, C., & Codinhoto, Ri. (2008). *DR2– The effect of the physical environment on mental wellbeing. Mental Capital and Wellbeing Project.* Mental Capital and Wellbeing. https://psycnet.apa.org/record/2009-22507-084

Core R Team. (2019). A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, *2*, https://www.R--project.org. http://www.r-project.org

Curl, A., Kearns, A., Mason, P., Egan, M., Tannahill, C., & Ellaway, A. (2015). Physical and mental health outcomes following housing improvements: evidence from the GoWell study. *J Epidemiol Community Health*, *69*(1), 12–19. https://doi.org/10.1136/JECH-2014-204064

Department for Business, E. and I. S. (2019). English Housing Survey: Fuel Poverty Dataset, 2017: Special Licence Access. [data collection]. *UK Data Service. SN: 8501*. http://doi.org/10.5255/UKDA-SN-8501-1

Department for Communities and Local Government. (2014). *English Housing Survey - Quality Report [online]*. https://webarchive.nationalarchives.gov.uk/20160128201230/http://www.ons.gov.uk/ons/guide-

Department for Communities and Local Government. (2018). *English Housing Survey - Technical Report 2016/17*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/725317/2016-17_EHS_Technical_Report.pdf

Department for Communities and Local Government. (2020). English Housing Survey, 2017: Housing Stock Data: Special Licence Access. [data collection]. 2nd Edition. *UK Data Service. SN: 8546*. https://doi.org/http://doi.org/10.5255/UKDA-SN-8546-2

Department of Health. (2014). The relationship between wellbeing and health. In *A Compendium of Factsheets: Wellbeing Across the Lifecourse*. https://www.gov.uk/government/publications/wellbeing-and-health-policy

Dodge, R., Daly, A., Huyton, J., & Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing*, *2*(3). https://doi.org/10.5502/ijw.v2i3.4

Dolan, P., Layard, R., & Metcalfe, R. (2011). *Measuring Subjective Well-being for Public Policy*. www.opsi.gov.uk/click-use/index.htm

Evans, G. W. (2003). The Built Environment and Mental Health. *Journal of Urban Health*,

80(4), 536–555. https://doi.org/10.1093/jurban/jtg063

Evans, G. W., Wells, N. M., & Moch, A. (2003). Housing and Mental Health: A Review of the Evidence and a Methodological and Conceptual Critique. *Journal of Social Issues*, *59*(3), 475–500. https://doi.org/10.1111/1540-4560.00074

Guite, H. F., Clark, C., & Ackrill, G. (2006). The impact of the physical and urban environment on mental well-being. *Public Health*, *120*(12), 1117–1126. https://doi.org/10.1016/j.puhe.2006.10.005

Hanc, M., McAndrew, C., & Ucci, M. (n.d.). *Conceptual approaches to wellbeing in buildings: a scoping review*. *47*(6). https://doi.org/10.1080/09613218.2018.1513695

Henson, R. M., Ortigoza, A., Martinez-Folgar, K., Baeza, F., Caiaffa, W., Vives Vergara, A., Diez Roux, A. V., & Lovasi, G. (2020). Evaluating the health effects of place-based slum upgrading physical environment interventions: A systematic review (2012–2018). *Social Science & Medicine*, *261*, 113102. https://doi.org/10.1016/J.SOCSCIMED.2020.113102

Hills, J. (2012). Getting the measure of fuel poverty: Final Report of the Fuel Poverty Review. *Case Report 72*.

Hoisington, A. J., Stearns-Yoder, K. A., Schuldt, S. J., Beemer, C. J., Maestre, J. P., Kinney, K. A., Postolache, T. T., Lowry, C. A., & Brenner, L. A. (2019). Ten questions concerning the built environment and mental health. *Building and Environment*, *155*, 58–69. https://doi.org/10.1016/j.buildenv.2019.03.036

Hunter, R. F., Cleland, C., Cleary, A., Droomers, M., Wheeler, B. W., Sinnett, D., Nieuwenhuijsen, M. J., & Braubach, M. (2019). Environmental, health, wellbeing, social and equity effects of urban green space interventions: A meta-narrative evidence synthesis. In *Environment International* (Vol. 130, p. 104923). Elsevier Ltd. https://doi.org/10.1016/j.envint.2019.104923

Ige, J., Pilkington, P., Orme, J., Williams, B., Prestwood, E., Black, D., Carmichael, L., & Scally, G. (2019). The relationship between buildings and health: a systematic review. *Journal of Public Health (Oxford, England)*, *41*(2), E121–E132. https://doi.org/10.1093/PUBMED/FDY138

Jevons, R., Carmichael, C., Crossley, A., & Bone, A. (2016). Minimum indoor temperature threshold recommendations for English homes in winter – A systematic review. In *Public Health* (Vol. 136, pp. 4–12). https://doi.org/10.1016/j.puhe.2016.02.007

Kang, H. J., Bae, K. Y., Kim, S. W., Shin, I. S., Yoon, J. S., & Kim, J. M. (2016). Anxiety symptoms in Korean elderly individuals: a two-year longitudinal community study.

International Psychogeriatrics, *28*(3), 423–433. https://doi.org/10.1017/S1041610215001301

Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J. V, Hern, S. C., & Engelmann, W. H. (n.d.). *The National Human Activity Pattern Survey (NHAPS) A Resource for Assessing Exposure to Environmental Pollutants*.

Krefis, A., Augustin, M., Schlünzen, K., Oßenbrügge, J., & Augustin, J. (2018). How Does the Urban Environment Affect Health and Well-Being? A Systematic Review. *Urban Science*, *2*(1), 21. https://doi.org/10.3390/urbansci2010021

Kuss, O. (2013). The danger of dichotomizing continuous variables: A visualization. *Teaching Statistics*, *35*(2), 78–79. https://doi.org/10.1111/TEST.12006

Liddell, C., & Morris, C. (2010). Fuel poverty and human health: A review of recent evidence. *Energy Policy*, *38*(6), 2987–2997. https://doi.org/10.1016/j.enpol.2010.01.037

Lüdecke, D. (2021). *sjPlot: Data Visualization for Statistics in Social Science* (R package version 2.8.7). https://cran.r-project.org/package=sjPlot

Maccallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). *On the Practice of Dichotomization of Quantitative Variables*. https://doi.org/10.1037/1082-989X.7.1.19

MHCLG. (2021). *English Housing Survey - Headline Report 2019-20*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/945013/2019-20_EHS_Headline_Report.pdf

Moore, T. H. M., Kesten, J. M., López-López, J. A., Ijaz, S., McAleenan, A., Richards, A., Gray, S., Savović, J., & Audrey, S. (2018). The effects of changes to the built environment on the mental health and well-being of adults: Systematic review. In *Health and Place* (Vol. 53, pp. 237–257). Elsevier Ltd. https://doi.org/10.1016/j.healthplace.2018.07.012

Office for National Statistics. (2018a). *Personal well-being frequently asked questions* . https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingfrequentlyaskedquestions#why-is-there-not-one-composite-indicator-for-personal-well-being

Office for National Statistics. (2018b). *Personal well-being in the UK QMI* -. https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingintheukqmi

Office for National Statistics. (2018c). *Personal well-being user guidance* . https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/personalwellbeingsurveyuserguide

Office for National Statistics. (2019). *Annual personal well-being estimates - .* https://www.ons.gov.uk/file?uri=/peoplepopulationandcommunity/wellbeing/datasets/headlineestimatesofpersonalwellbeing/october2017toseptember2018/personalwellbeingestimatesoctober2017toseptember2018.xls

Opinium. (2018). *Brits spend 90% of their time indoors | Opinium*. https://www.opinium.com/brits-spend-90-of-their-time-indoors/

Pannucci, C. J., & Wilkins, E. G. (2010). Identifying and avoiding bias in research. *Plastic and Reconstructive Surgery*, *126*(2), 619–625. https://doi.org/10.1097/PRS.0b013e3181de24bc

Pollard, E. L., & Lee, P. D. (2003). Child well-being: A systematic review of the literature. *Social Indicators Research*, *61*(1), 59–78. https://doi.org/10.1023/A:1021284215801

Sadowski, H. S., Ugarte, B., Kolvin, I., Kaplan, C. E., & Barnes, J. (1999). Early life family disadvantages and major depression in adulthood. *The British Journal of Psychiatry*, *174*(2), 112–120. https://doi.org/10.1192/BJP.174.2.112

Singh, A., Daniel, L., Baker, E., & Bentley, R. (2019). Housing Disadvantage and Poor Mental Health: A Systematic Review. *American Journal of Preventive Medicine*, *57*(2), 262–272. https://doi.org/10.1016/J.AMEPRE.2019.03.018

Snijders, T. A. B., & Bosker, R. J. (Roel J. . (2011). *Multilevel analysis : an introduction to basic and advanced multilevel modeling* (2nd ed.).

Sture Holm. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.

Tjur, T. (2009). Coefficients of determination in logistic regression models - A new proposal: The coefficient of discrimination. *American Statistician*, *63*(4), 366–372. https://doi.org/10.1198/tast.2009.08210

Turley, R., Saith, R., Bhan, N., Rehfuess, E., & Carter, B. (2013). Slum upgrading strategies involving physical environment and infrastructure interventions and their effects on health and socio-economic outcomes. *The Cochrane Database of Systematic Reviews*, *2013*(1). https://doi.org/10.1002/14651858.CD010067.PUB2

UK Data Service. (n.d.). *UK Data Service Registration* . Retrieved November 17, 2020, from https://www.ukdataservice.ac.uk/get-data/how-to-access/registration.aspx

Vieweg, B. W., & Hedlund, J. L. (1983). The General Health Questionnaire (GHQ): A comprehensive review. *Journal of Operational Psychiatry*, *14*(2).

von Szombathely, M., Albrecht, M., Antanaskovic, D., Augustin, J., Augustin, M., Bechtel, B., Bürk, T., Fischereit, J., Grawe, D., Hoffmann, P., Kaveckis, G., Krefis, A.,

Oßenbrügge, J., Scheffran, J., & Schlünzen, K. (2017). A Conceptual Modeling Approach to Health-Related Urban Well-Being. *Urban Science*, *1*(2), 17. https://doi.org/10.3390/urbansci1020017

Voukelatou, V., Gabrielli, L., Miliou, I., Cresci, S., Sharma, R., Tesconi, M., & Pappalardo, L. (2020). Measuring objective and subjective well-being: dimensions and data sources. In *International Journal of Data Science and Analytics* (pp. 1–31). Springer. https://doi.org/10.1007/s41060-020-00224-2

Warwick-Edinburgh Mental Well-being Scale (WEMWBS) User Guide. (2008). www.healthscotland.com/mental-health-publications.aspx

Western, M., & Tomaszewski, W. (2016). Subjective wellbeing, objective wellbeing and inequality in Australia. *PLoS ONE*, *11*(10). https://doi.org/10.1371/journal.pone.0163345

Wickham, H. (2016). ggplot2 Elegant Graphics for Data Analysis (Use R!). *Springer*.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43). https://doi.org/10.21105/joss.01686

World Health Organization. (2018). *Mental health: strengthening our response*. https://www.who.int/en/news-room/fact-sheets/detail/mental-health-strengthening-our-response

Appendix A.

The RECORD statement – checklist of items, extended from the STROBE statement, that should be reported in observational studies using routinely collected health data**.**

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Title: "cross-sectional" and "survey"	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	1.1. Title: "cross-sectional" and "English Housing Survey"; further details on actuals files in abstract 1.2. The title mentions 'English'; the abstract states "England" and "April 2016 to March 2018" 1.3 Abstract mentions the two data sets linked
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Introduction p 1		

Objectives	3	State specific objectives, including any prespecified hypotheses	Section 2.2		
Methods					
Study Design	4	Present key elements of study design early in the paper			
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Section 2.1 (These data cover the period April 2016 to March 2018; in England; English Housing Survey (EHS); two components: (1) a household interview, and (2) a physical inspection of a subsample of the properties)		
Participants	6	*Cross-sectional study* - Give the eligibility criteria, and the sources and methods of selection of participants	2.2 (description of data sets: complete data on all relevant variables; see Figure 1).	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be	6.1 *systematic random sample design* with link to full technical report. 6.2 none done by the authors of this paper, see

				referenced. If validation was conducted for this study and not published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Technical Report linked in manuscript for details. 6.3 Linked files and overlap of participants are shown in Figure 1. (note, not databases)
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Outcomes: Section 2.3.1 and Section 3.1. Covariates (~predictor & confounders): Section 2.3.2. and 3.2	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	See link to Github and additionally description in the paper (dichotomization of outcome variable described; minimum category count of 30 and hence merging if below)
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is	See Table A1. And methods 2.1		

		more than one group			
Bias	9	Describe any efforts to address potential sources of bias	Described in 2.5 – to avoid confirmation bias, preregistration of study; for other biases link to EHS report on bias		
Study size	10	Explain how the study size was arrived at	2.2 with Figure 1 – any household with complete data on the relevant variables		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	Section 3.3.2 (categories with counts <30 merged as described); reference category either largest or as prespecified		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cross-sectional study* - If	a) Described in 2.4 b) not applicable c) Described in 2.2 (anyone with missing data for wellbeing excluded; other 5		

		applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	cases for missing data excluded as they could not be merged with any other category; for variables with a lot of missing data, this was coded as its own category (details in 3.2.2). d) none taken e) none conducted		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	12.1 not applicable as such as all available data used 12.1 No outlier correction conducted ; for details of merging of categories and dealing with missing data see 3.2.2 and Figure 1.
Linkage		..		RECORD 12.3: State whether the study included person-level,	12.3 not appplicabl e

				institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	a) this is shown in Figure 1 for the available data; for any steps preceding this secondary analysis, a link to the technical report is provided. b) not applicable (numbers reduced only for missing data) c) see Figure 1	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.*, study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	13.1 see Figure 2 and linked technical report
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest	a) this is provided in 3.2 b) this is done throughout 3.2.2 and 2.2		

Outcome data	15	*Cross-sectional study* - Report numbers of outcome events or summary measures	Summary measures of the outcome variables provided in 3.1		
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	a) Figures show OR and include 95% confidence interval for significant effects only; full results in Appendix tables b) only applicable for safety: mean value across the safety items was calculated and rounded to the nearest integer. c) not relevant		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	For the factor analysis, oblique rotation was also carried out leading to the same factor structure; for the regression analysis, OLS was also run, with results roughly the same		

			(results can be reproduced with the linked code)		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Table in 3.3 shows the outcomes of all hypotheses tests; summary at the beginning of 4		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Limitations discussed in 4.3; main limitation is that some variables that might have been important were unavailable	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Some variables would have likely been better as self-reported variables than as surveyor assessed / calculated based on metrics as their subjective perception is likely more pertinent to wellbeing. (see 4.2). Large missing data on safety perception means that this variable is based on

					substantially fewer data points and hence potentially noisier.
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Done throughout the discussion (4.).		
Generalisability	21	Discuss the generalisability (external validity) of the study results	Discussed in 4.3.		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Stated in the Acknowledgement section.		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Link to data, link to preregistration and link to code all stated.

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

Appendix B

'The TReQlist' – A checklist for reporting of tools that promote transparency, reproducibility, and quality of research

Tools	Delete as applicable	Comments
Pre-registration		
This study has pre-analysis plan.	Yes	
If yes		
URL		10.17605/OSF.IO/F26ZS
Was it registered before data collection?	Yes	
Does the paper mention and explain deviations from the PAP?	Yes	*Minor deviations (3) in Appendix; two additional analyses in manuscript body (section 2.5 & 3.5)*
Reporting guidelines		
This paper follows a reporting guideline.	Yes	
If yes		
Which one?		*RECORD (Benchimol, E. I., Smeeth, L., Guttmann, A., Harron, K., Moher, D., Peteresen, I., … Langan, S. M. (2015). The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Medicine, 12(10), 1001885. https://doi.org/10.1371/journal.pmed.1001885)*
Open Data and Code		
Data/code are publicly available	Code only	The code can be directly accessed on Github. The data can not be made publicly available by the authors but can be accessed following registration on UK Data Archive.
Does the paper make a statement on data and code availability?	Yes, on data and code	*Section 1.2*
If yes		
What is / are the link(s)?		*https://github.com/Gesche-Huebner/Wellbeing_Repo*
Have steps been taken to ensure the data are FAIR?	No	Paper constitutes secondary data analysis, i.e. we did not create or share the data.
Has meta-data been uploaded?	No	
Preprints		
Have you uploaded a preprint?	Yes Planned following submission No	
If yes		
What is the link?		*https://osf.io/preprints/socarxiv/t6uxz/*
If planned		
Which preprint server/location?		

Reference: Huebner, G. M., Fell, M. J., & Watson, N. E. (2021). Improving energy research practices: guidance for transparency, reproducibility and quality. *Buildings and Cities*, *2*(1), 1–20. https://doi.org/10.5334/bc.67

Appendix C. Overview of all variables used and their data sources. All variables are from Study Number 8546 except the ones marked with * that come from Study Number 8501.

Dependent variables

Name in paper	*EHS Variable*	*EHS Variable Label*	*Dataset*	*Type of data collection*
LifeSatisfaction	QSatis	Satisfaction with life nowadays	identity_1617_sl_protect	Self-report
Worthwhile	QWorth	Things done in life are	identity_1617_sl_protect	Self-report
Happy	QHappy	How happy yesterday	identity_1617_sl_protect	Self-report
Anxious	QAnxious	How anxious yesterday	identity_1617_sl_protect	Self-report

Identifier variables

Name in paper	*EHS Variable*	*EHS Variable Label*	*Dataset*	*Type of data collection*
na	serialanon	Unique archived identifier	Multiple	Technical
na	persno	Person identifier	Multiple	Technical
YearInterview	fiyear	Interview survey year (interviewed)	general16plus17_sl_protect	Technical

Predictor variables / those for creating predictors

Name in paper	*EHS Variable*	*EHS Variable Label*	*Dataset*	*Type of data collection*
GeneralHealth	QHealth1	General Health	disability_1617_sl_protect	Self-report
MaritalStatus	xMarSta2	Legal marital status	people_1617_sl_protect	Self-report
HighestQual	HiQual	educated to degree level or above	people_1617_sl_protect	Self-report
AgeHRP	agehrp6x	Age of HRP - 6 band	interview_16plus17_sl_protect	Self-report
SexHRP	Sex	Sex	people_1617_sl_protect	Self-report
EmploymentHRP	emphrpx	Employment status (primary) of HRP	interview_16plus17_sl_protect	Self-report
AHCeqvIncome	AHCinceqv5	AHC equivalised income quintiles	interview_16plus17_sl_protect	Self-report

Househ_Type	hhcompx	Household composition	interview_16plus17_sl_protect	Self-report
EthnicityHRP	ethhrp8x	Ethnic origin of HRP - 8 categories	interview_16plus17_sl_protect	Self-report
BedroomStandard	lhastdx	Bedroom standard (2011 definition)	interview_16plus17_sl_protect	Calculated
Arrears	mrgAr21	Any difficulties keeping up with mortgage payments in the last 12 months	owner_1617_sl_protect	Self-report
Arrears	ArrPR2	Fallen behind with rent payments over the last 12 months	renter_1617_sl_protect	Self-report
Tenure	tenure2	Tenure	interview_16plus17_sl_protect	Self-report
Morphology	ru11morph	Rurality classification - morphology (2011 COA)	generalfs16plus17_sl_protect	Surveyor
LRWarm	hmHeatOn	Can you keep living room warm	energy_1617_sl_protect	Self-report
HeatingCost	hmHtCst	How easy is it to meet heating/fuel costs	energy_1617_sl_protect	Self-report
FuelPovertyLIHC	fpLIHCflg*	Fuel poverty flag - low income high costs measure	fuel_poverty_data_2017_special_licence_protect	Calculated
FuelPovertyIncome	Fpflgf*	10% definition fuel poverty flag - full income definition	fuel_poverty_data_2017_special_licence_protect	Calculated
DwellingType	dwtype8x	Dwelling type	physical_16plus17_sl_protect	Surveyor
EPC	EPceeb12e	Energy efficiency rating band (SAP 2012)	physical_16plus17_sl_protect	Surveyor
DecentHome	dhomesy	Decent homes - overall standard (15 hazard HHSRS model)	physical_16plus17_sl_protect	Surveyor
CostUrgentRepair	cststdbx	Basic repair costs (per square metre)	physical_16plus17_sl_protect	Surveyor
CostBasicRepair	cststdux	Urgent repair costs (per square metre)	physical_16plus17_sl_protect	Surveyor
CostComprRepair	Cststdcx	Comprehensive repair costs (per square metre)	physical_16plus17_sl_protect	Surveyor
DwellingAge	dwage7x	Dwelling age	physical_16plus17_sl_protect	Surveyor
FloorArea	floor5x	Useable floor area - original EHS definition	physical_16plus17_sl_protect	Surveyor
Damp	Cdprob	Any problems with condensation, damp or mould in home	damp_1617_sl_protect	Self-report

Area_Satisf	HAS44	Satisfied with area	attitudes_1617_sl_protect	Self-report
SafetyHomeAlone	nhhmsf1	Safety in neighbourhood: at home alone	attitudes_1617_sl_protect	Self-report
SafetyDay	nhsfday	Safety in neighbourhood: outside during the day	attitudes_1617_sl_protect	Self-report
SafetyNight	nhsfnte	Safety in neighbourhood: outside after dark	attitudes_1617_sl_protect	Self-report
IMD	Imd1510	Index of Multiple Deprivation	general16plus17_sl_protect	Surveyor
Garden	Fexpltyp	Type of plot	around_sl_protect	Surveyor
GOR	GorEHS	Government Office Region	general16plus17_sl_protect	Surveyor
na	Farlittr	Litter/rubbish/dumping	around_sl_protect	Surveyor
na	FarGraff	Graffiti	around_sl_protect	Surveyor
na	FarVanda	Vandalism	around_sl_protect	Surveyor
na	FarExcre	Dog/ other excrement	around_sl_protect	Surveyor
na	FarCond	Condition of dwellings	around_sl_protect	Surveyor
na	Farsites	Vacant sites	around_sl_protect	Surveyor
na	FarIndus	Intrusive industry	around_sl_protect	Surveyor
na	FarNocon	Non-conforming uses	around_sl_protect	Surveyor
na	FarVacnt	Vacant/boarded-up buildings	around_sl_protect	Surveyor
na	Farairqu	Ambient air quality	around_sl_protect	Surveyor
na	Fartraff	Heavy traffic	around_sl_protect	Surveyor
na	Farmotor	Intrusion from motorways arterial roads	around_sl_protect	Surveyor
na	Farrails	Railway aircraft noise	around_sl_protect	Surveyor
na	Farparks	Nuisance from street parking	around_sl_protect	Surveyor
na	FarGrdns	Scruffy gardens/landscaping	around_sl_protect	Surveyor
na	FarBldgs	Scruffy/neglected buildings	around_sl_protect	Surveyor
na	FarRoads	Condition of road, pavements and street furniture	around_sl_protect	Surveyor

Appendix C: Notes and deviation from the prespecifaction

Deviations from prespecification

In section 3.2.1

In the prespecification, we had stated to use a factor analysis using the principal factor axis method; however, a PCA was chosen as its solution explained a larger share of the variance with a near identical structure of factor loadings (0.62 versus 0.52 of explained variance). The principal factor method will usually yield results close to the principal component method if either the correlations or the number of variables is large (Rencher, 2002).

In section 3.2.2

For household composition, single person households with an age greater 60 became reference category as opposed to couple without children as it was the larger category contrary to expectations.

For housing factors, some exclusion criteria were changed in deviation from the prespecification to avoid too much data loss. For HeatingCost, those who had answered ‘Don’t know’ or given no answer were combined into a “Don’t know” category. Regarding the ability of keeping one’s living room warm the category of “Don’t know” (N = 141) was retained as its own category instead of being deleted. For damp, 14 respondents had provided no answer or said ‘don’t know’; they were merged with ‘no’ to avoid having to remove them from the analysis.

For all variables on safety, about one third of respondents were coded as “Not applicable”. In order to avoid removing such a large number of respondents, this became its own category ‘No valid answer’. Those who not answered the question though being asked it and those who gave an uninterpretable answer, i.e. not doing something for other than safety reason were also put into that category.

The four variables indicating the factors standing for problems in local neighbourhood were not rounded to the nearest integer to allow keeping greater granularity.

Results of testing models with different versions of repair costs

As per prespecification, we run every model using either CostUrgentRepair or CostBasicRepair and inspected the AIC to understand with which repair variable model fit was better. AIC was very similar across the models. The decision was taken to remove urgent repair costs; AIC was smaller for LifeSatisfaction when removing urgent repair costs; with a magnitude of 1.39 which was the largest of any difference observed. Neither urgent nor basic repair costs were significant. As a previous study had used comprehensive repair costs, we also ran the models using that variable; however, it was likewise non significant and AIC values were very similar.

Appendix E. Full regression models for the four outcome variables.

	LifeSatisfaction_Dummy		**Worthwhile_Dummy**		**Happy_Dummy**		**Anxious_Dummy**	
Predictors	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*	*Odds Ratios*	*p*
(Intercept)	13.87 (6.95 – 27.78)	**<0.001**	22.33 (10.79 – 46.37)	**<0.001**	6.46 (3.52 – 11.90)	**<0.001**	0.08 (0.04 – 0.16)	**<0.001**
Very_Good	*Reference*		*Reference*		*Reference*		*Reference*	
Good	0.70 (0.59 – 0.82)	**<0.001**	0.76 (0.63 – 0.91)	**0.003**	0.74 (0.64 – 0.85)	**<0.001**	1.23 (1.04 – 1.47)	**0.016**
Fair	0.40 (0.33 – 0.47)	**<0.001**	0.40 (0.33 – 0.49)	**<0.001**	0.50 (0.42 – 0.58)	**<0.001**	1.92 (1.59 – 2.32)	**<0.001**
Bad	0.17 (0.14 – 0.21)	**<0.001**	0.20 (0.16 – 0.26)	**<0.001**	0.21 (0.17 – 0.26)	**<0.001**	2.59 (2.03 – 3.30)	**<0.001**
Very_Bad	0.08 (0.06 – 0.12)	**<0.001**	0.12 (0.08 – 0.17)	**<0.001**	0.15 (0.11 – 0.20)	**<0.001**	4.89 (3.52 – 6.79)	**<0.001**
Married	*Reference*		*Reference*		*Reference*		*Reference*	
Single	0.93 (0.73 – 1.18)	0.538	0.72 (0.56 – 0.92)	**0.009**	0.84 (0.69 – 1.02)	0.081	1.21 (0.95 – 1.53)	0.120
Divorced	0.94 (0.73 – 1.22)	0.641	0.82 (0.63 – 1.08)	0.150	0.94 (0.76 – 1.17)	0.596	1.14 (0.88 – 1.48)	0.313
Widowed	0.70 (0.52 – 0.94)	**0.018**	0.66 (0.49 – 0.90)	**0.009**	0.79 (0.61 – 1.02)	0.071	1.08 (0.79 – 1.48)	0.614
Other_Status	0.68 (0.50 – 0.93)	**0.016**	0.69 (0.50 – 0.96)	**0.028**	0.77 (0.59 – 1.02)	0.065	1.40 (1.01 – 1.92)	**0.040**
Degree_Level	*Reference*		*Reference*		*Reference*		*Reference*	

Other_Qual	1.04 (0.88 – 1.23)	0.661	1.01 (0.84 – 1.21)	0.923	1.02 (0.88 – 1.17)	0.829	0.93 (0.78 – 1.11)	0.420
Not_Asked	1.13 (0.92 – 1.39)	0.228	0.91 (0.73 – 1.13)	0.382	1.00 (0.83 – 1.20)	0.986	1.00 (0.81 – 1.23)	0.977
45_54	*Reference*		*Reference*		*Reference*		*Reference*	
16_24	1.38 (0.97 – 1.96)	0.074	0.96 (0.67 – 1.38)	0.808	1.24 (0.90 – 1.73)	0.191	1.23 (0.85 – 1.76)	0.272
25_34	1.56 (1.26 – 1.94)	**<0.001**	1.30 (1.03 – 1.65)	**0.031**	0.97 (0.80 – 1.18)	0.790	1.10 (0.87 – 1.38)	0.429
35_44	1.23 (1.01 – 1.49)	**0.039**	0.94 (0.76 – 1.16)	0.574	0.93 (0.78 – 1.11)	0.421	1.05 (0.85 – 1.30)	0.638
55_64	1.17 (0.94 – 1.45)	0.152	1.44 (1.14 – 1.82)	**0.002**	1.17 (0.96 – 1.42)	0.121	1.14 (0.91 – 1.43)	0.242
65OrOver	1.60 (1.12 – 2.28)	**0.010**	1.75 (1.20 – 2.55)	**0.004**	1.26 (0.91 – 1.73)	0.163	1.14 (0.78 – 1.66)	0.507
Female	*Reference*		*Reference*		*Reference*		*Reference*	
Male	0.90 (0.79 – 1.03)	0.127	0.73 (0.64 – 0.83)	**<0.001**	0.99 (0.88 – 1.11)	0.807	0.83 (0.73 – 0.96)	**0.010**
Full-Time_Work	*Reference*		*Reference*		*Reference*		*Reference*	
Retired	1.15 (0.85 – 1.54)	0.361	0.78 (0.57 – 1.07)	0.120	1.42 (1.10 – 1.84)	**0.008**	0.76 (0.55 – 1.03)	0.080
Other_Inactive	0.73 (0.60 – 0.89)	**0.001**	0.60 (0.48 – 0.74)	**<0.001**	0.91 (0.75 – 1.09)	0.287	1.56 (1.27 – 1.92)	**<0.001**
Part-Time_Work	1.02 (0.83 – 1.24)	0.882	0.91 (0.72 – 1.14)	0.387	1.08 (0.90 – 1.29)	0.431	1.12 (0.91 – 1.38)	0.284

Unemployed	0.76 (0.57 – 1.01)	0.056	0.77 (0.57 – 1.04)	0.089	1.01 (0.77 – 1.33)	0.959	0.94 (0.68 – 1.29)	0.713
Full-Time_Education	1.44 (0.84 – 2.54)	0.197	1.11 (0.64 – 2.00)	0.715	1.00 (0.64 – 1.59)	0.987	0.92 (0.52 – 1.56)	0.755
5th_(Highest)	*Reference*		*Reference*		*Reference*		*Reference*	
4th	0.80 (0.63 – 1.01)	0.066	1.18 (0.92 – 1.51)	0.199	1.14 (0.94 – 1.38)	0.198	0.93 (0.73 – 1.18)	0.550
3rd	0.83 (0.65 – 1.06)	0.133	0.98 (0.76 – 1.25)	0.864	0.96 (0.79 – 1.17)	0.672	1.08 (0.85 – 1.37)	0.539
2nd	0.79 (0.61 – 1.01)	0.058	0.99 (0.77 – 1.27)	0.942	0.91 (0.74 – 1.12)	0.394	0.91 (0.71 – 1.17)	0.462
1st_(Lowest)	0.76 (0.58 – 0.99)	**0.042**	0.98 (0.74 – 1.29)	0.881	1.05 (0.84 – 1.32)	0.680	1.06 (0.81 – 1.39)	0.686
Single=>60yrs	*Reference*		*Reference*		*Reference*		*Reference*	
Couple_W/_Dep	1.50 (1.04 – 2.16)	**0.029**	1.69 (1.15 – 2.50)	**0.008**	1.33 (0.96 – 1.84)	0.087	1.06 (0.72 – 1.55)	0.774
Single<60yrs	0.97 (0.74 – 1.28)	0.837	1.08 (0.81 – 1.44)	0.585	1.14 (0.88 – 1.46)	0.325	1.31 (0.98 – 1.76)	0.072
Couple>=60yrs	1.91 (1.39 – 2.64)	**<0.001**	1.68 (1.21 – 2.33)	**0.002**	1.35 (1.03 – 1.76)	**0.030**	1.19 (0.87 – 1.64)	0.278
Single_Parent_W/_Dep	0.92 (0.66 – 1.28)	0.622	1.63 (1.14 – 2.33)	**0.007**	1.14 (0.84 – 1.54)	0.413	1.12 (0.78 – 1.59)	0.543
Couple<60yrs	1.91 (1.36 – 2.69)	**<0.001**	1.55 (1.09 – 2.22)	**0.015**	1.20 (0.89 – 1.61)	0.226	1.16 (0.82 – 1.65)	0.396

Other_Multiperson	0.88 (0.67 – 1.16)	0.362	1.14 (0.85 – 1.53)	0.396	0.94 (0.73 – 1.22)	0.629	1.01 (0.74 – 1.37)	0.936
White	*Reference*		*Reference*		*Reference*		*Reference*	
Black	1.01 (0.77 – 1.33)	0.967	1.01 (0.75 – 1.37)	0.946	1.18 (0.91 – 1.53)	0.214	0.85 (0.61 – 1.15)	0.294
Indian	1.28 (0.80 – 2.12)	0.309	1.73 (1.02 – 3.10)	0.053	1.37 (0.91 – 2.12)	0.141	1.16 (0.72 – 1.81)	0.534
Pakist_Bangla	0.91 (0.60 – 1.39)	0.646	1.09 (0.69 – 1.77)	0.719	1.56 (1.03 – 2.40)	**0.039**	0.89 (0.55 – 1.39)	0.627
Mixed	0.79 (0.53 – 1.20)	0.254	1.00 (0.64 – 1.60)	0.998	1.02 (0.70 – 1.53)	0.913	1.56 (1.02 – 2.33)	**0.036**
Other	1.01 (0.64 – 1.64)	0.966	1.23 (0.74 – 2.12)	0.442	1.01 (0.67 – 1.56)	0.954	0.61 (0.33 – 1.06)	0.099
Chinese_Other_Asian	0.96 (0.57 – 1.64)	0.866	0.98 (0.58 – 1.71)	0.934	0.79 (0.51 – 1.26)	0.315	1.06 (0.60 – 1.77)	0.842
At_Standard	*Reference*		*Reference*		*Reference*		*Reference*	
1Below	1.16 (0.88 – 1.53)	0.285	1.24 (0.92 – 1.70)	0.170	1.09 (0.84 – 1.41)	0.521	0.91 (0.67 – 1.22)	0.539
=>2Below	0.64 (0.36 – 1.15)	0.133	1.49 (0.75 – 3.22)	0.282	0.94 (0.53 – 1.71)	0.836	0.96 (0.48 – 1.81)	0.903
1Above	1.02 (0.87 – 1.20)	0.796	1.11 (0.93 – 1.32)	0.243	0.99 (0.86 – 1.15)	0.917	1.00 (0.85 – 1.19)	0.955
=>2Above	0.96 (0.75 – 1.22)	0.716	1.15 (0.89 – 1.48)	0.278	0.81 (0.66 – 1.00)	**0.048**	0.88 (0.68 – 1.13)	0.313
No	*Reference*		*Reference*		*Reference*		*Reference*	

NA	0.86 (0.74 – 1.00)	**0.044**	0.85 (0.73 – 0.99)	**0.038**	0.99 (0.86 – 1.13)	0.884	1.06 (0.90 – 1.24)	0.498
Yes	0.64 (0.49 – 0.84)	**0.001**	0.83 (0.62 – 1.12)	0.212	0.82 (0.64 – 1.06)	0.125	1.31 (0.99 – 1.73)	0.059
Own_Outright	*Reference*		*Reference*		*Reference*		*Reference*	
Rent_Housing_Association	0.80 (0.64 – 1.01)	0.063	0.83 (0.65 – 1.06)	0.139	1.01 (0.82 – 1.24)	0.930	1.11 (0.86 – 1.42)	0.420
Rent_Private_Unfurn.	0.69 (0.54 – 0.88)	**0.003**	0.75 (0.58 – 0.97)	**0.028**	1.01 (0.81 – 1.26)	0.940	1.05 (0.81 – 1.37)	0.720
Own_Mortgage	1.09 (0.82 – 1.44)	0.565	0.91 (0.68 – 1.23)	0.531	1.11 (0.87 – 1.41)	0.402	1.06 (0.80 – 1.42)	0.677
Rent_Local_Authority	0.91 (0.71 – 1.17)	0.470	0.89 (0.69 – 1.15)	0.367	1.08 (0.87 – 1.35)	0.480	1.02 (0.78 – 1.33)	0.883
Rent_Private_Furn.	0.89 (0.58 – 1.37)	0.584	0.80 (0.52 – 1.25)	0.324	0.84 (0.59 – 1.20)	0.325	1.62 (1.06 – 2.45)	**0.024**
Yes	*Reference*		*Reference*		*Reference*		*Reference*	
No	0.70 (0.59 – 0.83)	**<0.001**	0.82 (0.69 – 0.99)	**0.033**	0.81 (0.69 – 0.95)	**0.009**	1.10 (0.91 – 1.32)	0.319
Don't_Know	0.57 (0.38 – 0.88)	**0.009**	0.63 (0.41 – 0.98)	**0.035**	0.71 (0.48 – 1.05)	0.082	1.19 (0.74 – 1.87)	0.460
Very_Easy	*Reference*		*Reference*		*Reference*		*Reference*	
Fairly_Easy	0.94 (0.80 – 1.10)	0.412	0.99 (0.84 – 1.17)	0.923	0.89 (0.77 – 1.02)	0.092	1.06 (0.90 – 1.26)	0.482
Neither	0.71 (0.59 – 0.86)	**<0.001**	0.71 (0.58 – 0.86)	**0.001**	0.69 (0.58 – 0.81)	**<0.001**	1.25 (1.02 – 1.53)	**0.028**

Fairly_Difficult	0.63 (0.51 – 0.77)	**<0.001**	0.63 (0.51 – 0.78)	**<0.001**	0.62 (0.51 – 0.75)	**<0.001**	1.61 (1.29 – 2.00)	**<0.001**
Very_Difficult	0.41 (0.31 – 0.55)	**<0.001**	0.48 (0.36 – 0.64)	**<0.001**	0.64 (0.49 – 0.84)	**0.001**	1.70 (1.27 – 2.27)	**<0.001**
Dont_Know	0.67 (0.38 – 1.20)	0.167	0.58 (0.33 – 1.05)	0.064	0.84 (0.50 – 1.46)	0.526	0.75 (0.35 – 1.47)	0.426
Not_In_FP_LIHC	*Reference*		*Reference*		*Reference*		*Reference*	
In_FP_LIHC	1.03 (0.84 – 1.26)	0.775	1.12 (0.91 – 1.39)	0.296	1.01 (0.84 – 1.22)	0.883	0.91 (0.74 – 1.13)	0.412
Not_In_FP	*Reference*		*Reference*		*Reference*		*Reference*	
In_FP	0.67 (0.54 – 0.84)	**<0.001**	0.69 (0.55 – 0.87)	**0.002**	0.81 (0.66 – 1.00)	**0.047**	1.11 (0.87 – 1.41)	0.395
Detached_House	*Reference*		*Reference*		*Reference*		*Reference*	
Semi-Detached_House	1.05 (0.78 – 1.39)	0.752	1.01 (0.74 – 1.35)	0.970	0.96 (0.76 – 1.22)	0.754	1.19 (0.89 – 1.59)	0.239
Purpose_Built_Flat_Low_Rise	1.16 (0.79 – 1.69)	0.445	1.11 (0.75 – 1.64)	0.615	0.78 (0.56 – 1.08)	0.134	1.09 (0.74 – 1.63)	0.660
Medium/Large_Terraced_House	0.99 (0.73 – 1.33)	0.927	1.01 (0.73 – 1.38)	0.955	0.88 (0.68 – 1.13)	0.320	1.14 (0.84 – 1.56)	0.411
Small_Terraced_House	0.98 (0.69 – 1.39)	0.910	1.04 (0.72 – 1.50)	0.820	0.89 (0.66 – 1.20)	0.440	1.18 (0.82 – 1.69)	0.375
Bungalow	1.19 (0.84 – 1.68)	0.339	1.43 (0.99 – 2.05)	0.054	0.98 (0.73 – 1.30)	0.867	1.16 (0.82 – 1.65)	0.405
Converted_Flat	1.07 (0.68 – 1.69)	0.764	0.96 (0.60 – 1.53)	0.849	0.88 (0.60 – 1.31)	0.534	1.15 (0.72 – 1.84)	0.560

Purpose_Built_Flat_High_Rise	1.43 (0.85 – 2.41)	0.175	0.94 (0.56 – 1.59)	0.816	0.76 (0.48 – 1.20)	0.237	1.53 (0.90 – 2.58)	0.114
D	*Reference*		*Reference*		*Reference*		*Reference*	
B	1.04 (0.62 – 1.81)	0.872	1.16 (0.68 – 2.06)	0.593	0.78 (0.51 – 1.23)	0.278	0.93 (0.53 – 1.56)	0.780
C	0.93 (0.80 – 1.07)	0.317	1.01 (0.87 – 1.18)	0.862	0.96 (0.84 – 1.09)	0.525	1.08 (0.93 – 1.26)	0.321
E	1.04 (0.85 – 1.28)	0.726	1.06 (0.85 – 1.32)	0.601	1.09 (0.91 – 1.30)	0.351	0.86 (0.69 – 1.07)	0.173
F	1.07 (0.74 – 1.57)	0.737	1.02 (0.69 – 1.53)	0.927	0.99 (0.72 – 1.38)	0.946	0.99 (0.66 – 1.45)	0.945
G	1.99 (0.99 – 4.28)	0.063	1.45 (0.71 – 3.25)	0.333	1.58 (0.88 – 3.00)	0.142	1.19 (0.61 – 2.21)	0.583
Decent	*Reference*		*Reference*		*Reference*		*Reference*	
Non-Decent	0.88 (0.75 – 1.04)	0.131	1.16 (0.97 – 1.38)	0.113	0.99 (0.85 – 1.15)	0.867	1.15 (0.96 – 1.37)	0.116
CostBasicRepair	1.00 (1.00 – 1.00)	0.162	1.00 (1.00 – 1.00)	0.684	1.00 (1.00 – 1.00)	0.972	1.00 (1.00 – 1.00)	0.407
1965_To_1980	*Reference*		*Reference*		*Reference*		*Reference*	
Pre_1919	1.23 (0.98 – 1.54)	0.077	1.26 (0.99 – 1.60)	0.064	1.02 (0.83 – 1.25)	0.854	0.92 (0.72 – 1.16)	0.475
1919_To_1944	1.26 (1.01 – 1.56)	**0.039**	1.31 (1.04 – 1.65)	**0.022**	0.89 (0.74 – 1.08)	0.233	1.08 (0.86 – 1.34)	0.510
1945_To_1964	1.02 (0.86 – 1.20)	0.858	1.04 (0.88 – 1.24)	0.638	0.92 (0.79 – 1.07)	0.298	0.99 (0.83 – 1.18)	0.898

1981_To_1990	1.13 (0.90 – 1.43)	0.287	1.06 (0.84 – 1.34)	0.644	1.08 (0.87 – 1.33)	0.492	0.76 (0.59 – 0.99)	**0.043**
1991_To_2002	1.21 (0.95 – 1.54)	0.131	1.24 (0.96 – 1.61)	0.097	1.02 (0.82 – 1.27)	0.844	1.09 (0.84 – 1.39)	0.519
Post_2002	1.40 (1.07 – 1.83)	**0.015**	1.29 (0.98 – 1.72)	0.071	1.12 (0.89 – 1.41)	0.356	0.96 (0.73 – 1.26)	0.778
50_To_69_Sqm	*Reference*		*Reference*		*Reference*		*Reference*	
110_Sqm_Or_More	1.24 (0.92 – 1.68)	0.158	1.11 (0.81 – 1.53)	0.510	1.01 (0.79 – 1.30)	0.914	1.26 (0.94 – 1.70)	0.121
90_To_109_Sqm	1.13 (0.88 – 1.45)	0.357	0.98 (0.75 – 1.28)	0.897	1.08 (0.87 – 1.36)	0.473	1.12 (0.86 – 1.46)	0.387
70_To_89_Sqm	1.02 (0.84 – 1.24)	0.825	0.96 (0.78 – 1.18)	0.701	0.97 (0.81 – 1.15)	0.690	1.02 (0.83 – 1.26)	0.833
Less_Than_50_Sqm	0.99 (0.82 – 1.19)	0.884	0.93 (0.77 – 1.13)	0.484	0.99 (0.83 – 1.17)	0.884	0.94 (0.77 – 1.15)	0.577
Private_Plot	*Reference*		*Reference*		*Reference*		*Reference*	
Shared_Plot_Only	0.86 (0.68 – 1.10)	0.235	0.97 (0.76 – 1.25)	0.823	1.01 (0.80 – 1.26)	0.957	1.13 (0.87 – 1.47)	0.368
Neither	1.04 (0.64 – 1.72)	0.867	0.85 (0.52 – 1.39)	0.501	0.77 (0.50 – 1.19)	0.232	1.10 (0.64 – 1.82)	0.729
No_(+)	*Reference*		*Reference*		*Reference*		*Reference*	
Yes_All_year	0.77 (0.66 – 0.90)	**0.001**	0.96 (0.81 – 1.13)	0.617	0.85 (0.74 – 0.98)	**0.024**	1.25 (1.06 – 1.47)	**0.008**
Yes_Winter	0.93 (0.78 – 1.10)	0.370	0.95 (0.80 – 1.14)	0.575	0.76 (0.65 – 0.88)	**<0.001**	1.09 (0.92 – 1.30)	0.316

Yes_Other	0.79 (0.55 – 1.13)	0.190	0.82 (0.56 – 1.21)	0.310	0.94 (0.67 – 1.33)	0.718	0.91 (0.59 – 1.36)	0.655
Very_Satisfied	*Reference*		*Reference*		*Reference*		*Reference*	
Fairly_Satisfied	0.70 (0.61 – 0.80)	**<0.001**	0.69 (0.61 – 0.80)	**<0.001**	0.76 (0.68 – 0.86)	**<0.001**	1.01 (0.87 – 1.16)	0.920
Neither	0.44 (0.35 – 0.55)	**<0.001**	0.56 (0.44 – 0.72)	**<0.001**	0.59 (0.48 – 0.73)	**<0.001**	1.11 (0.86 – 1.43)	0.420
Slightly_Dissatisfied	0.46 (0.37 – 0.59)	**<0.001**	0.72 (0.55 – 0.94)	**0.014**	0.58 (0.47 – 0.73)	**<0.001**	1.13 (0.87 – 1.47)	0.351
Very_Dissatisfied	0.39 (0.29 – 0.53)	**<0.001**	0.47 (0.34 – 0.65)	**<0.001**	0.62 (0.46 – 0.82)	**0.001**	1.19 (0.86 – 1.64)	0.280
1st_(Most)	*Reference*		*Reference*		*Reference*		*Reference*	
2nd	1.03 (0.85 – 1.26)	0.744	0.82 (0.67 – 1.02)	0.069	1.06 (0.88 – 1.27)	0.571	1.03 (0.83 – 1.28)	0.773
3rd	0.98 (0.79 – 1.21)	0.817	0.92 (0.73 – 1.15)	0.448	1.03 (0.85 – 1.26)	0.755	1.12 (0.89 – 1.41)	0.322
4th	1.01 (0.80 – 1.27)	0.947	0.88 (0.70 – 1.12)	0.306	0.93 (0.76 – 1.15)	0.519	1.05 (0.82 – 1.33)	0.719
5th	0.95 (0.74 – 1.21)	0.661	0.91 (0.71 – 1.18)	0.494	0.87 (0.70 – 1.09)	0.223	1.01 (0.78 – 1.30)	0.954
6th	0.95 (0.73 – 1.23)	0.674	0.87 (0.66 – 1.14)	0.308	0.95 (0.75 – 1.20)	0.664	1.24 (0.95 – 1.62)	0.108
7th	0.89 (0.68 – 1.16)	0.382	0.84 (0.64 – 1.12)	0.242	0.74 (0.58 – 0.93)	**0.011**	1.25 (0.94 – 1.65)	0.123

8th	0.87 (0.66 – 1.15)	0.313	0.84 (0.63 – 1.13)	0.246	0.96 (0.75 – 1.23)	0.758	0.90 (0.66 – 1.21)	0.475
9th	0.93 (0.69 – 1.25)	0.612	0.91 (0.66 – 1.25)	0.552	0.95 (0.73 – 1.23)	0.678	0.77 (0.55 – 1.06)	0.113
10th_(Least)	0.73 (0.53 – 1.00)	**0.050**	0.77 (0.56 – 1.08)	0.130	0.81 (0.62 – 1.07)	0.141	1.12 (0.81 – 1.56)	0.487
South_East	*Reference*		*Reference*		*Reference*		*Reference*	
North_West	1.22 (0.98 – 1.51)	0.082	1.02 (0.81 – 1.28)	0.884	0.98 (0.81 – 1.18)	0.809	1.07 (0.85 – 1.34)	0.572
East	1.11 (0.89 – 1.38)	0.364	1.03 (0.81 – 1.30)	0.805	1.14 (0.94 – 1.39)	0.193	0.83 (0.65 – 1.05)	0.119
London	1.00 (0.79 – 1.27)	0.996	1.01 (0.78 – 1.30)	0.947	1.00 (0.81 – 1.24)	0.991	0.96 (0.74 – 1.23)	0.734
Yorkshire_And_The_Humber	1.09 (0.87 – 1.37)	0.471	1.00 (0.79 – 1.28)	0.973	1.00 (0.82 – 1.23)	0.966	1.24 (0.98 – 1.56)	0.072
West_Midlands	1.06 (0.83 – 1.35)	0.629	1.06 (0.82 – 1.37)	0.670	0.89 (0.72 – 1.10)	0.292	0.89 (0.68 – 1.15)	0.358
South_West	1.02 (0.80 – 1.31)	0.850	0.83 (0.65 – 1.07)	0.157	1.11 (0.89 – 1.38)	0.353	0.90 (0.70 – 1.16)	0.427
East_Midlands	0.87 (0.68 – 1.10)	0.243	0.81 (0.63 – 1.04)	0.103	0.92 (0.74 – 1.14)	0.452	1.10 (0.85 – 1.43)	0.454
North_East	1.19 (0.90 – 1.57)	0.231	0.89 (0.67 – 1.18)	0.401	1.02 (0.79 – 1.30)	0.901	0.93 (0.69 – 1.25)	0.630
Urban_>_10k	*Reference*		*Reference*		*Reference*		*Reference*	

Scruffy1	0.96 (0.84 – 1.10)	0.582	0.94 (0.81 – 1.08)	0.367	0.99 (0.87 – 1.11)	0.813	1.13 (0.97 – 1.30)	0.106
Traffic2	0.97 (0.85 – 1.10)	0.603	0.91 (0.79 – 1.05)	0.189	1.01 (0.89 – 1.14)	0.915	0.82 (0.71 – 0.95)	**0.008**
Vacant3	0.96 (0.78 – 1.20)	0.733	0.87 (0.70 – 1.09)	0.231	1.14 (0.93 – 1.40)	0.199	1.07 (0.84 – 1.36)	0.560
Behaviour4	1.10 (0.93 – 1.31)	0.267	1.19 (0.99 – 1.42)	0.061	1.06 (0.91 – 1.25)	0.433	0.79 (0.66 – 0.95)	**0.013**
Town_Fringe	1.19 (0.96 – 1.48)	0.108	1.19 (0.95 – 1.50)	0.126	1.17 (0.97 – 1.41)	0.112	0.84 (0.66 – 1.05)	0.127
Village	1.15 (0.85 – 1.57)	0.382	1.00 (0.73 – 1.38)	0.994	0.91 (0.71 – 1.17)	0.456	0.95 (0.70 – 1.29)	0.760
Hamlet	2.08 (1.23 – 3.73)	**0.009**	1.24 (0.76 – 2.12)	0.401	0.95 (0.66 – 1.38)	0.769	1.10 (0.71 – 1.66)	0.663
Very_Safe	*Reference*		*Reference*		*Reference*		*Reference*	
Fairly_Safe	1.01 (0.85 – 1.20)	0.916	0.90 (0.75 – 1.09)	0.273	0.99 (0.85 – 1.15)	0.881	0.96 (0.79 – 1.15)	0.629
A_Bit_Unsafe	0.91 (0.70 – 1.19)	0.493	0.73 (0.55 – 0.96)	**0.024**	0.81 (0.63 – 1.03)	0.080	1.41 (1.07 – 1.84)	**0.014**
Very_Unsafe	0.43 (0.23 – 0.79)	**0.007**	0.35 (0.20 – 0.62)	**<0.001**	0.53 (0.31 – 0.92)	**0.024**	2.45 (1.43 – 4.19)	**0.001**
No_Answer	0.85 (0.72 – 1.00)	0.050	0.83 (0.70 – 0.99)	**0.034**	0.97 (0.84 – 1.12)	0.663	1.04 (0.87 – 1.23)	0.668
YearInterview2016	0.96 (0.83 – 1.10)	0.547	0.92 (0.80 – 1.07)	0.277	1.07 (0.95 – 1.22)	0.263	0.96 (0.82 – 1.11)	0.541

YearInterview2018	0.96 (0.80 – 1.16)	0.681	0.94 (0.77 – 1.14)	0.501	0.93 (0.79 – 1.10)	0.419	0.88 (0.72 – 1.07)	0.212
Observations	9205		9205		9205		9205	
R^2 Tjur	0.227		0.175		0.115		0.089	

Appendix F. Results of hypothesis testing. HB stands for Holm-Bonferroni correction.

	Outcome variable			
Hypothesis	Life Satisfaction	Worthwhile	Happy	Anxious
Wellbeing is lower for occupants in dwellings with lower EPC ratings.	ns	ns	ns	ns
Wellbeing is highest for occupants in detached homes and lowest in high-rise flats.	ns	ns	ns	ns
Wellbeing is lower for occupants who find it difficult to meet their heating/fuel costs.				
very difficult vs easy	<.001	<.001	<.001	<.001
Fairly difficult vs. easy	<.001	<.001	<.001	<.001
Neither vs. easy	<.001	<.001	<.001	0.028
Wellbeing is lower for occupants in fuel poverty (using the 10%) definition.	<.001	.002	0.047 (HB: ns)	ns
Wellbeing is lower for occupants in overcrowded households.	ns	ns	ns	ns
Wellbeing is lower for occupants unable to keep their living room at comfortable temperatures.				
No vs. yes	<.001	.033 (HB: ns)	0.009	ns
Don't know vs. yes	.009	.035 (HB: ns)	ns	ns
Wellbeing is lower for occupants in dwellings with higher repair costs per square meter.	ns	ns	ns	ns
Wellbeing is lower for occupants who live in areas with more problems.				
Scruffy	ns	ns	ns	ns
Traffic	ns	ns	ns	0.008
Vacant	ns	ns	ns	ns
Behaviour	ns	ns	ns	0.013
Wellbeing is lower for occupants who are less satisfied with their environment.				
Very dissatisfied vs. satisfied	<0.001	<0.001	<0.001	ns
Slightly dissatisfied vs. satisfied	<0.001	0.014	<0.001	ns
Neither vs. satisfied	<0.001	<0.001	<0.001	ns
Fairly satisfied vs. satisfied	<0.001	<0.001	<0.001	ns
Wellbeing is lower for occupants who feel less safe in their local environment.				
Very unsafe vs. safe	0.007	<.001	0.024	0.001
A bit unsafe vs. safe	ns	.024 (HB: ns)	ns	.014 (HB: ns)

Wellbeing is lower for occupants who live in more deprived areas.	ns	ns	ns	ns
Wellbeing is lower for occupants with damp problems.				
Year round vs. no	0.001	ns	0.024	0.008
Winter only vs. no	ns	ns	<.001	ns

Appendix G (1-4). Deviance analysis used for model comparison.

G.1 Outcome Life Satisfaction.

		LifeSatisfaction			
Model	Resid Df	Resid Dev	DF	Deviance	p
Personal	9167	8261.1			
Personal_Housing	9121	8011.1	46	250.02	<.001
Personal_Housing_Neighbourhood	9089	7850.4	32	160.74	<.001

G.2 Outcome Worthwhile.

		Worthwhile			
Model	Resid Df	Resid Dev	DF	Deviance	p
Personal	9167	7550.7			
Personal_Housing	9121	7397.9	46	152.88	<.001
Personal_Housing_Neighbourhood	9089	7291.8	32	106.07	<.001

G.3 Outcome Happiness.

		Happpiness			
Model	Resid Df	Resid Dev	DF	Deviance	p
Personal	9167	9710.8			
Personal_Housing	9121	9591.1	46	119.643	<.001
Personal_Housing_Neighbourhood	9089	9500.6	32	90.545	<.001

G.4 Outcome Anxious.

		Anxious			
Model	Resid Df	Resid Dev	DF	Deviance	p
Personal	9167	7447.3			
Personal_Housing	9121	7360.0	46	87.355	<.001
Personal_Housing_Neighbourhood	9089	7287.9 32	32	72.010	<.001