

Supplementary Data

Table S1: Potential data sources for property/dwelling information

Data source	Description	Purpose in this study
Definition of households		
Local Land and Property Gazetteer (LLPG) (n = 117,707 residential properties)	Each council is mandated to maintain an LLPG, a list of all addresses in its area (including houses, flats, parking spaces). Addresses are assigned a unique property reference, classified using a nationally standardised list, and collated in a national gazetteer (NLPG). [21]	Derive the sample frame, i.e. all residential properties within a local authority area
Bedroom number data		
Data that could be used		
Islington Council housing database (n = 38,836)	Data on currently and previously owned council stock in Islington: used for allocating properties to residents and maintaining homes.	'Gold standard' data on bedroom number. (A bedroom is defined as rooms designated as bedrooms when property was built)
Commercial property websites: historical data (n = 8,355)	Islington properties advertised on the platform for sale	Historical data on bedroom number (i.e. pre 2018)
Other dwelling characteristics		
Energy Performance Certificates (EPC) (n = 71,826)	Inspections conducted when a property is due to be rented or sold. The requirement for EPCs was introduced in stages from 2007. Apart from a few exempted buildings, a building must have an EPC when constructed, sold or let. Alongside energy performance, they include up to 92 variables related to characteristics of the dwelling and its surroundings. [22]	The following variables were selected to enable imputation of missing bedroom numbers and estimate space within homes: total floor area number of habitable rooms type of property building age tenure
Council-held: Council tax band (n = 100,862)	Each household is liable to pay a fee to their council for local services. The fee is in part determined by the valuation 'band' for their home. The Valuation Office Agency assesses the band of a property. Banding is based on a number of characteristics of the home, include size and layout. [23]	To enable the imputation of missing bedroom numbers.
Council: household tenure (n = 103,212)	Tenure covers whether a household rents or owns their property. For rented property it also covers whether the landlord is the council or a housing association or whether the household rents privately. Tenure is associated with overcrowding, and is associated with the type of properties and characteristics of households within them. [24]	Current data on tenure
Data we investigated but were not able to use or permitted to access		
Nethouseprice website download tool	Properties sold over a number of years.	Current data on bedroom number. Our request to access these data was turned down.
Commercial property websites: current data	Islington properties advertised on commercial property platforms for sale.	Downloading of data on bedroom number is no longer permitted from platforms

Continued

Table S1: Continued

Data source	Description	Purpose in this study
MetaStreet: Private rental sector register [25]	Database of the characteristics of Islington housing licensed for private rental under selective licensing scheme. [26]	Current data on bedroom number in private rental sector properties. Unable to use because bedroom count included living rooms and therefore could not be directly compared with council housing stock data.
Valuation Office Agency	The Valuation Office Agency collects and collates data on properties to provide the UK government with information needed to support taxation and benefits (e.g. council tax band, see above). This includes the numbers of bedrooms in a property. [23]	Current data on bedroom number. Our request to access these data was turned down because VOA were not permitted to disclose this information to local councils unless it is used in connection with a function of VOA or if there is a legal gateway established by an Act of Parliament.



Table S2: Correlation coefficients: independent numerical variables and bedroom number, used to confirm suitability of candidate variables for multiple imputation

	Number of bedrooms	Total floor area	Number of habitable rooms	Council tax band	Property type
Number of bedrooms	1.00				
Total Floor Area	0.80	1.00			
Number of Habitable Rooms	0.91	0.84	1.00		
Council Tax Band	0.56	0.69	0.59	1.00	
Property type	0.56	0.64	0.61	0.58	1.00

(n = 37,450)

Figure S1: Comparison of the distribution of imputed and actual bedroom number data

