



Department for  
Communities and  
Local Government

## English Housing Survey

The weighting methodology introduced in 2013-14







Department for  
Communities and  
Local Government

# English Housing Survey

The weighting methodology introduced in 2013-14

February 2015  
Department for Communities and Local Government

*Crown copyright, 2015*

*Copyright in the typographical arrangement rests with the Crown.*

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

This document/publication is also available on our website at [www.gov.uk/dclg](http://www.gov.uk/dclg)

If you have any enquiries regarding this document/publication, complete the form at <http://forms.communities.gov.uk/> or write to us at:

Department for Communities and Local Government  
Fry Building  
2 Marsham Street  
London  
SW1P 4DF  
Telephone: 030 3444 0000

For all our latest news and updates follow us on Twitter: <https://twitter.com/CommunitiesUK>

February 2015

ISBN: 978-1-4098-4503-4

---

# Contents

---

1. Introduction
  2. Motivation for changing the weighting methodology
  3. The weighting methodology of the 2012-13 data
  4. The new weighting methodology
  5. The changes between the previous and new method
- Annex      Comparison tables of previous and new weighting

---

## 1. Introduction

- 1.1 The English Housing Survey (EHS) is a continuous national survey commissioned by the Department for Communities and Local Government (DCLG). It collects information about people's housing circumstances and the condition and energy efficiency of housing in England.
- 1.2 The EHS consists of two component surveys: an initial interview survey (IS) of households and a follow-up physical survey (PS) of a sub-sample of dwellings.
- 1.3 Both the IS and PS component surveys intentionally include a disproportionate number of properties from renting tenure groups. This over-sampling of the less prevalent tenure groups is important in order to provide sufficient information (larger samples) on less prevalent tenures and facilitate reliable and detailed analysis.
- 1.4 The survey is weighted to take account of the over-sampling of the less prevalent tenure groups in order to provide unbiased national estimates. The weighting methodology uses a sequence of stages which correct for the over-sampling of the less prevalent tenure groups and reduce the bias from differential non-response in both the IS and PS response process. The resulting weights sum to estimated population totals, which enable national estimates of the total population to be obtained from the survey.
- 1.5 The EHS has utilized several weighting methodologies since its inception. This paper reviews the weighting methodology applied to the 2012-13 EHS data and describes/evaluates the current weighting methodology (introduced in 2013-14), which aims to simplify the weighting process and to reduce the year-on-year variation in the weighted totals for both households and dwellings.

## 2. Motivation for changing the weighting methodology

- 2.1 The main aim in changing the weighting methodology was to simplify the process. The previous weighting scheme was complicated because of the legacy arrangements dating back to the time when the EHS was part of the Integrated Household Survey (IHS).
- 2.2 The key difference between the previous and new scheme is in the calibration of the data. Calibration takes an initial weight (e.g. selection weight) and then

---

adjusts (or calibrates) it to given control totals. The process generates a weight which produces survey estimates that exactly match the population for the specific characteristics (control totals) used in the adjustment.

- 2.3 The previous method for weighting the EHS used two sets of control totals for households/dwellings when calibrating the sample to match the population of households/dwellings: DCLG counts of dwellings and household counts estimated from ONS's population estimates and the Labour Force Survey (LFS). This method has several disadvantages:
- The LFS estimates of household counts vary year-on-year and are often subject to revision after the initial release; this resulted in variation in the EHS weighted estimates year-on-year as well. The DCLG counts on the other hand are relatively stable.
  - The DCLG counts and LFS estimates are not consistent with each other (because they were produced for different purposes and from different sources), so complexity was added to the EHS weighting process to try to control for that.
- 2.4 The previous weighting methodology in general was over-complicated, with separate strands of weighting for the different levels (household and dwelling) which stemmed from the need to fit the EHS into the IHS model, However, these complication are not needed anymore because the EHS is no longer part of the IHS.
- 2.5 The new weighting methodology uses one set of control totals for households/dwellings. Given their relative stability and the desire for consistency within DCLG, it uses the DCLG dwelling counts.
- 2.6 In addition, the new methodology seeks to eliminate the previous overlap between the weighting stages for households and dwellings.
- 2.7 The next section presents a summary of the previous weighting methodology. More detail is given in the Technical Report for EHS 2012-13<sup>1</sup>.

---

<sup>1</sup> [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/335421/EHS\\_Technical\\_Report\\_2012-13.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335421/EHS_Technical_Report_2012-13.pdf)

---

### 3. The weighting methodology of the 2012-13 data

3.1 Table 3.1 presents the stages needed to generate the weights for the full household sample (all cases for which an interview was completed).

**Table 3.1: Weighting interviewed households (method applied to the 2012-13 data)**

#### ***Stage 1: Calculating the probability of selection for addresses***

As mentioned earlier, the EHS oversamples the less prevalent tenure groups to provide a larger sample of those groups for more detailed analysis. The oversampling involves identifying a predicted tenure for each of the sampled addresses selected from the Postcode Address File (PAF) based on information derived from Experian<sup>2</sup> and uses the predicted tenure to carry out a 'sift' of the sampled addresses by tenure in order to reduce or 'sift out' a random sample of owner-occupied addresses.

The random sift carried out during sampling has an impact on the probability of selection of an issued address. The probability of selection of an issued address on the EHS was therefore calculated as a combination of:

- a grossing constant ( $w_1$ ) equivalent to the total number of delivery points on the Postcode Address File (PAF) divided by the number of delivery points sampled and
- a correction factor required to take account of the oversampling ( $w_1\_sift$ ) equivalent to the number predicted to be tenure X divided by the number selected in predicted tenure X.

#### ***Stage 2a: Adjustment for not-worked addresses***

Very occasionally, a very small number of the addresses sampled were not issued to interviewers for operational reasons. Not-worked addresses were filtered-out by applying an adjustment to the Stage 1 weights so that the distribution of the worked addresses within region matched all issued addresses.

#### ***Stage 2: Address-to-dwelling ratio***

---

<sup>2</sup> Experian possess a database that contains information obtained from a number of sources including insurance companies, Census, etc. referred to as 'Residata'. It is from this that information is taken on predominant tenure within a postcode as well as other information. The matching of the EHS sample to 'Residata' is carried out by BRE.

---

**Stage 3: Dwelling-to-household ratio**

These stages corrected for the selection of one dwelling unit at the sampled address, and one household at the selected dwelling. The address-to-dwelling and dwelling-to-household ratios were averaged (smoothed) within weighting (smoothing) classes defined by tenure and region to remove large peaks and troughs within the classes.

**Stage 4a: Filtering out office refusals<sup>3</sup>****Stage 4: Contact at the interview survey****Stage 5: Cooperation at the interview survey**

These stages made a series of adjustments for non-participation. They are based on weighting classes (groups of cases) generated from the CHAID<sup>4</sup> algorithm of the SPSS AnswerTree software. These classes were derived using an unweighted CHAID model (per stage) which partitioned the sample based on predictor variables significantly associated with the propensity to respond at a given fieldwork stage. Once the classes had been generated, the adjustment within each class was made based on the weighted estimates (using the combined weights up to that stage). The main reason for doing the correction in stages was to use any additional information available at each stage.

**Stage 6: Calibrating weighting for the interview survey**

The weighted interview sample was calibrated to match both the distribution of age/sex by region using the mid-year population estimates and the distribution of tenure using the LFS. This stage generated the weights for the full household sample.

- 3.2 The weighting stages for cases which were eligible for PS (the paired sample) are presented in Table 3.2. Cases in the paired sample included addresses which the interviewers had determined to be vacant and where a physical survey

---

<sup>3</sup> Office refusals are refusals the fieldwork agency receives from sampled households after they have received the letter announcing the survey but before the interviewer makes contact.

<sup>4</sup> Chi-squared Automatic Interaction Detector

---

had also been achieved. The inclusion of vacant addresses was thus necessary for the production of the final paired dwelling weights.

**Table 3.2: Weighting cases eligible for the physical survey (method applied to the 2012-13 data)**

<p><b>Stage 1: Calculating the probability of selection for addresses</b></p> <p><b>Stage 2a: Adjustment for not-worked addresses</b></p> <p><b>Stage 2: Address to dwelling ratio</b></p> <p>As per Table 3.1.</p> <p>Note that Stage 3 was not required because the weighting took place at the dwelling level and not at the household level.</p>
<p>The next three stages (7a, 7 and 8) effectively repeated Stages 4a, 4 and 5 described in Table 3.1.</p> <p><b>Stage 7a: Filtering out office refusals (for interview)</b></p> <p><b>Stage 7: Contact at the dwelling (for interview)</b></p> <p><b>Stage 8: Cooperation at the dwelling (for interview)</b></p> <p>In practice, the models themselves were normally exactly the same as for the corresponding IS stages (with the same cases included and identical weighting classes/terms in the model) The difference was that the adjustments were made using slightly different weights (as the household adjustment from Stage 3 was not included).</p> <p>Vacant dwellings were treated as having 'responded' at each of these stages and were assigned a weight adjustment of unity for each stage. In effect their weights were unchanged from Stage 2, and their inclusion had no effect on the models derived for the occupied cases.</p>
<p><b>Stage 9: Sub-sampling by tenure for the physical survey</b></p> <p>The PS sample includes a disproportionately larger number of dwellings from renting tenure groups to enable detailed analysis of these. This was achieved by under-sampling of properties likely to be owner occupied. This stage calculates selection weights for the field sift which was required to achieve a disproportionate number of</p>

---

rented properties.

***Stage 10: Co-operation at the physical survey***

This stage adjusted for non-response to the physical survey using weighting classes generated from a CHAID model. Data collected during the interview survey (for occupied dwellings only) were also used to help determine the weighting classes. Vacant cases were treated separately (using a separate CHAID model) because the process of participation for vacant dwellings is generally different from that for occupied dwellings and this needed to be reflected in the weights.

***Stage 11: Calibration weighting for the physical survey & new-build adjustment***

The dwelling sample was calibrated to DCLG dwelling estimates by tenure and region.

The achieved sample of dwellings did not include any dwellings built since the sample was drawn, so the weights were adjusted using the numbers of new dwellings built between the sampling date and the reference date for weighting<sup>5</sup>.

Because of the small number of recently-built dwellings in the survey, the weights of all cases with a construction date of 1990 onwards were adjusted in this process (re-calibrated to the same DCLG dwelling estimates). This was carried out, separately for areas with a high/low rate of new build, and for private/social sector housing (excluding local authority housing because the rate of new building in this sector is negligible).

***Stage 12: Creation of initial paired household weights***

Household weights were derived by calibrating the stage 10 weights to the same population totals as for the full household sample (i.e. to ONS population projections by sex and age group by region, and to tenure proportions from the LFS, as in Stage 6).

***Stage 13: Creation of interim paired household weights***

---

<sup>5</sup> For EHS 2012-13, sampling date was May 2011 and reference date October 2012.

---

A second set of household weights were also derived using the stage 11 weights and the household-to-dwelling ratio. These weights gave total numbers of households, by tenure and region which were consistent with the DCLG dwelling estimates. However, these totals would differ slightly from those given by the LFS and population estimates.

***Stage 14: Creation of final paired household weights***

In order to reconcile these differences, the final paired household weights were derived by scaling the Stage 13 weights to produce the same weighted totals as the Stage 12 calibrated weights.

***Stage 15: Creation of final paired dwelling weights***

The final paired dwelling weights were derived to be consistent with the final paired household weights. These were achieved by applying the scaling factors derived at Stage 14 to the weights derived at Stage 11 for both the occupied and vacant dwellings. As a result, the weighted dwelling totals (unlike those produced at stage 11) no longer matched the DCLG dwelling estimates.

The adjustments in Stages 14 and 15 ensured that the final paired sample dwelling and household weights were consistent with each other and with the LFS and population estimates data.

## 4. The new weighting methodology

- 4.1. The previous and new methods have essentially the same steps, except that, for the new method, the stages are re-ordered. The new method is developed to use only one set of control totals for households/dwellings: the DCLG estimates of tenure within region, which are used in conjunction with the age/sex distribution of population estimates. In this, it represents a reversion to the control totals used prior to the inclusion of the EHS in the IHS (from 2008-09 onwards).
- 4.2. Tables 4.1 and 4.2 present the weighting stages for the new method. To avoid confusion, the prefix “N” is used to describe the stages.

---

**Table 4.1: Weighting interviewed households (new method)**

<p><b>Stage N1: Calculating the probability of selection for addresses</b></p> <p>Same as Stage 1 of the previous method.</p>
<p><b>Stage N2a: Adjustment for not-worked addresses</b></p> <p>Same as Stage 2a of the previous method.</p>
<p><b>Stage N2: Filtering out office refusals</b></p> <p><b>Stage N3: Contact at the interview survey</b></p> <p><b>Stage N4: Cooperation at the interview survey</b></p> <p>The non-response models (for the occupied dwellings) are fitted at the address level, rather than at household level as per the previous method. As a result, there is no need to derive/impute the number of dwellings at an address and number of households at a dwelling for the non-productive occupied addresses<sup>6</sup>.</p> <p>These stages are equivalent to Stages 4a, 4 and 5, and Stages 7a, 7 and 8 of the previous method.</p>
<p><b>Stage N5: Address-to-dwelling ratio</b></p> <p>Same as Stage 2 of the previous method, but re-ordered to take place after the non-response weighting; it includes occupied (productive) as well as vacant (productive and unproductive) addresses (see Change 4 below).</p>
<p><b>Stage N6: Dwelling-to-household ratio</b></p> <p>Same as Stage 3 of the previous method.</p>
<p><b>Stage N7: Calibrating weighting for the interview survey</b></p> <p>The interview sample (occupied dwellings only) is calibrated (weighted by the</p>

---

<sup>6</sup> See also “change 4” in the next section.

composite weight from Stages N1 to N5) to:

- population proportions for age/sex (for dwellings with more than one household, the household counts are adjusted using the dwelling-to-household ratio from Stage N6);
- counts of occupied dwellings by tenure for each region (these control totals are estimated by adjusting the DCLG dwelling counts, which include both occupied and vacant dwellings, using estimates of the proportion of dwellings that are occupied from the current and four previous years of the EHS).

Because the pre-calibration weights are at the dwelling level, the calibration can be done directly to the DCLG dwelling counts (i.e. there is no need to weight to LFS estimates of household counts).

The calibration weight from Stage N7 is combined with the dwelling-to-household ratio from stage N6 to give the weights for the full household sample.

**Table 4.2: Weighting surveyed households (new method)**

***Stage N1: Calculating the probability of selection for addresses***

***Stage N2a: Adjustment for not-worked addresses***

***Stage N2: Filtering out office refusals***

***Stage N3: Contact at the interview survey***

***Stage N4: Cooperation at the interview survey***

***Stage N5: Address-to-dwelling ratio***

***Stage N6: Dwelling-to-household ratio***

As per Table 4.1.

***Stage N8: Sub-sampling by tenure for the physical survey***

Same as Stage 9 of the previous method.

***Stage N9: Cooperation at the physical survey***

---

Same as Stage 10 of the previous method.

***Stage N10: Calibrating weighting for the physical survey***

The occupied paired sample (i.e. cases with both IS and PS) is weighted by the composite weight from Stages N1 to N5 and N8 to N9. The vacant sample (i.e. vacant addresses at PS) is weighted by the composite weight from stages N1, N2a, N5, N8, and N9. Both the occupied and vacant samples are then combined and calibrated to:

- population proportions for age/sex (for dwellings with more than one household, the household counts are adjusted using the dwelling-to-household ratio from Stage N6; for vacant dwellings the household counts are all zero);
- DCLG counts of all dwellings (occupied + vacant) by tenure for each region;
- counts of vacant dwellings by tenure (RSL and LA collapsed) for each region (these control totals are estimated by adjusting the DCLG dwelling counts using estimates of vacancy rates from the current and four previous years of the EHS).

***Stage N11: Adjustment for new build (final paired dwelling weights)***

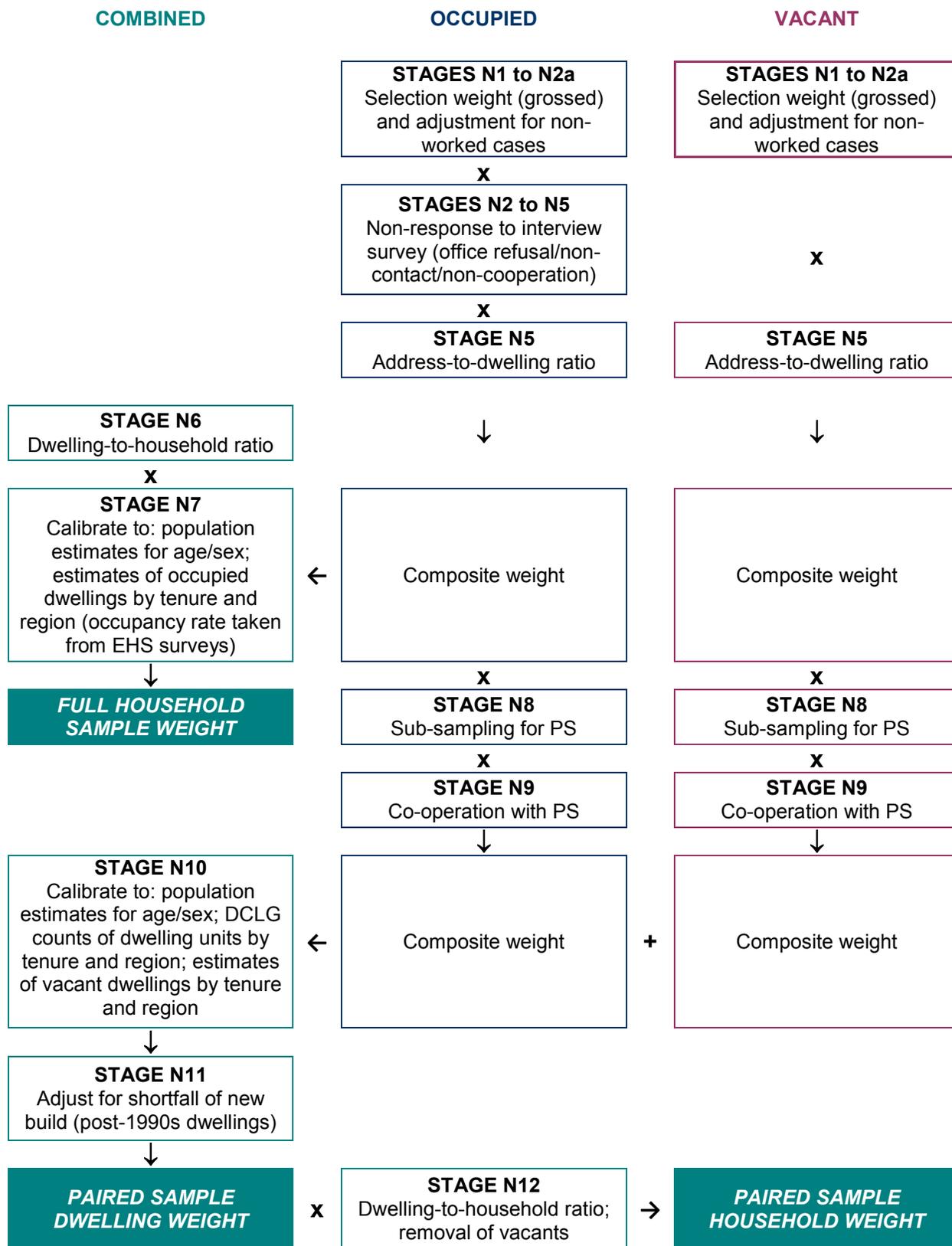
Post-1990 dwellings are weighted-up to cover for new addresses on the PAF by re-calibrating to the same overall control totals, but with re-estimated counts of post-1990 build (same process used in Stage 11 of the previous method). This generates the final paired sample dwelling weights.

***Stage N12: Creation of final paired household weights***

The final paired dwelling weight from Stage N11 (after removing the vacant dwellings) is adjusted using the dwelling-to-household ratio from stage N6 to generate the final paired sample household weights.

4.3. The following flow chart gives a visual representation of the new method, Figure 4.1.

**Figure 4.1: Weighting surveyed households (new method)**



---

## 5. The changes between the previous and new method

- 5.1 In addition to re-organising the weighting stages, the new method introduces a number of small changes to some of the stages. This section lists all the differences between the previous and new method and discusses the impact of each. The most significant change is listed first.

***Change 1: Calibrate to one set of control totals for households/dwellings (estimated DCLG counts of dwellings) rather than two (estimated DCLG counts of dwellings and LFS counts of households)***

- 5.2 As already discussed, the previous method used two sets of control totals for household/dwellings counts by tenure: household counts from the LFS and DCLG-estimated dwelling counts. Using two sets of estimates for households/dwellings acknowledged the fact that the EHS is both a survey of households as well as a survey of dwellings, and enabled the survey to be part of the IHS. However, it resulted in a complicated weighting strategy, and one of the aims of the new method was to use only one set of control totals for households/dwellings. The weighting methodology was therefore amended to use only the DCLG dwelling counts by tenure and region (plus the adjustment to the population distribution by age/sex).
- 5.3 Both the full interview sample and the paired cases interview samples are calibrated to estimated DCLG counts of occupied dwellings (these are derived from the DCLG dwelling counts and estimates of vacancy from the EHS) and population proportions of age/sex using ONS population estimates (see Table 4.1). The calibration weights are at the dwelling level, so when combined with the dwelling-to-household ratio they give the household level interview survey weights.
- 5.4 The physical survey sample is calibrated to estimated DCLG counts of all dwellings (i.e. occupied plus vacant), estimates of the number of vacant dwellings (using estimates of vacancy from the EHS), and population proportions of age/sex using ONS population estimates (see Table 4.2). The calibrated weights (including the stage for new build addresses) are the paired sample dwelling weights. These are then combined with the dwelling-to-household ratio (after removing the vacant dwellings) to give the paired sample household weights.
- 5.5 The new method removes a lot of the complexity and produces weights that are very highly correlated with those using the previous method. Using the EHS

2012-13 data, the correlations between the previous and new method are 0.967 for the IS household weights and 0.938 for the PS dwelling weights. There is also a minimal impact on efficiency: the average design factor (deft) for IS was 1.222 for the previous method and 1.225 for the new method; for PS the deft was 1.686 for the previous method and 1.713 for the new method.

- 5.6 The tables in the Annex compare a range of key estimates from IS and PS (using 2012-13 data) for the two sets of weights from the two methodologies (i.e. the previous and new full household sample and paired sample dwelling weights). The most notable difference seems to be in the estimates for tenure. Specifically (see Table 5.1), the use of the new weighting method increases the number of households by about 600,000, and reduces the estimate of the percentage of owner-occupier households from 65.2% to 63.4%, while the number of private rented sector (PRS) households are increased by about 10%. These differences (which will have an impact on the EHS time-series for household tenure estimates), reflect the lack of consistency between the DCLG and LFS estimates.

**Table 5.1: LFS and DCLG household estimated counts compared (reference date: October 2012)**

	LFS (households)		DCLG (households)		difference	
	<i>count</i>	<i>percentage</i>	<i>count</i>	<i>percentage</i>	<i>count</i>	<i>percentage</i>
owner occupier	14,336,750	65.2	14,316,824	63.4	-19,926	-0.14
private rented	3,956,092	18.0	4,355,672	19.3	399,580	10.10
social rented	3,683,991	16.8	3,905,056	17.3	221,065	6.00
<b>total</b>	<b>21,976,833</b>	<b>100.0</b>	<b>22,577,552</b>	<b>100.0</b>	<b>600,719</b>	<b>2.73</b>

- 5.7 Table 5.2 shows the estimated LFS household and DCLG occupied dwelling counts for October 2012 (reference date of EHS 2012-13). The two distributions look more similar than one would expect given that the two sets of estimates are actually measuring different things: counts of households for the LFS and counts of dwellings for the DCLG estimates.

**Table 5.2: LFS household estimated counts and DCLG occupied dwellings estimates compared (reference date: October 2012)**

	LFS (households)		DCLG (occupied dwellings)		Difference	
	count	percentage	count	percentage	count	percentage
owner occupiers	14,336,750	65.2	14,295,751	64.6	-40,999	-0.29
private renters	3,956,092	18.0	3,987,974	18.0	31,882	0.81
social renters	3,683,991	16.8	3,854,713	17.4	170,722	4.63
Total	21,976,833	100.0	22,138,438	100.0	161,605	0.74

5.8 In the new method, the dwelling-to-household adjustment is applied post-calibration to derive household counts from the DCLG dwelling counts. When these are compared with the LFS household counts, the two distributions are more different (Table 5.1). This is a result of there being (on average) more households per dwelling for private rented (1.0937) than owner occupied (1.0012) and social rented (1.0089).

### **Change 2: Obtaining estimates for the counts of occupied and vacant dwellings**

5.9 In order to run the calibration weighting for the new method, control totals need to be generated for the number of vacant dwellings by tenure and region. This was done by combining five years of estimates of vacancy rates by tenure and region from the EHS7. We used a weighted average to achieve consistent estimates, with more weight given to the more recent years, in the ratios: 0.5 / 0.25 / 0.125 / 0.0625 / 0.0625.

### **Change 3: Generating the control totals for dwellings**

5.10 The previous method for generating the control totals for dwellings was to use the DCLG counts of dwellings and then add in the estimates of the new dwellings built between their reference date (April) and the control reference date for the EHS (October).

5.11 The problem with this method is that it does not allow for dwellings moving from one tenure to another. It assumes that there is no net change in tenure between the reference date of the counts and that of the EHS, so the counts of dwellings

<sup>7</sup> Combining five years of EHS data gives a more accurate estimate of vacancy rates (compared to using the estimate from the latest survey only).

that are newly built are just added on. However, when comparing the dwelling counts from previous years, the number of dwellings for both owner occupied and local authorities are actually falling. The previous method for generating the dwelling estimates did not take this into account – in fact it was most probably making the dwelling counts for owner-occupied and LA dwellings less accurate.

5.12 Table 5.3 shows the estimates for the dwelling counts since 2005. The coloured estimates are the ones imputed in order to weight the EHS using the previous method. As well as the counts in each year, there is a column showing the change over the previous count.

**Table 5.3: DCLG dwelling count estimates 2005 to 2012 (previous method)**

	owner occupier		private rented		housing association		local authority	
	<i>count</i>	<i>change</i>	<i>count</i>	<i>change</i>	<i>count</i>	<i>change</i>	<i>count</i>	<i>change</i>
Mar-05	15,182		2,720		1,802		2,166	
Mar-06	15,134	-48	2,987	267	1,865	63	2,087	-79
Mar-07	15,168	34	3,182	195	1,951	86	1,987	-100
Mar-08	15,141	-27	3,443	261	2,056	105	1,870	-117
Mar-09	15,042	-99	3,705	262	2,128	72	1,820	-50
Mar-10	14,961	-81	3,912	207	2,180	52	1,786	-34
Mar-11	14,890	-71	4,105	193	2,255	75	1,726	-60
Mar-12	<b>14,874</b>	<b>-16</b>	<b>4,244</b>	<b>139</b>	2,304	49	1,689	-37
Oct-12	<b>14,909</b>	<b>35</b>	<b>4,254</b>	<b>10</b>	<b>2,309</b>	<b>5</b>	<b>1,693</b>	<b>4</b>

5.13 As can be seen, the counts for owner occupied and local authority have been falling consistently since 2005, but the estimates that were produced for the weighting (in red) showed an increase. This clearly demonstrates that the previous methodology was not sensible as it could only add dwellings to a particular tenure, and took no account of dwellings changing tenure.

5.14 Rather than adding the estimates of new build to the most recent counts, the new method adjusts the most recent counts based on the trend in change in counts by tenure (and region) for the previous three years. In order to give most weight to the previous year, the new method uses a weighted average of 0.5 / 0.25 / 0.25 for the change in the previous three years.

5.15 Table 5.4 shows the estimates of dwelling counts using the new method.

**Table 5.4: DCLG dwelling count estimates 2005 to 2012 (new method)**

	owner occupier		private rented		housing association		local authority	
	count	change	count	change	count	change	count	change
Mar-05	15,182		2,720		1,802		2,166	
Mar-06	15,134	-48	2,987	267	1,865	63	2,087	-79
Mar-07	15,168	34	3,182	195	1,951	86	1,987	-100
Mar-08	15,141	-27	3,443	261	2,056	105	1,870	-117
Mar-09	15,042	<b>-99</b>	3,705	<b>262</b>	2,128	72	1,820	-50
Mar-10	14,961	<b>-81</b>	3,912	<b>207</b>	2,180	<b>52</b>	1,786	<b>-34</b>
Mar-11	14,890	<b>-71</b>	4,105	<b>193</b>	2,255	<b>75</b>	1,726	<b>-60</b>
Mar-12	<b>14,810</b>	<b>-81</b>	<b>4,319</b>	<b>214</b>	2,304	<b>49</b>	1,689	<b>-37</b>
Oct-12	<b>14,769</b>	<b>-40</b>	<b>4,426</b>	<b>107</b>	<b>2,332</b>	<b>28</b>	<b>1,668</b>	<b>-21</b>

5.16 For owner occupied, the numbers have been falling since 2007. We assume that between 2011 and 2012, it would fall by the weighted average yearly change between 2008 and 2011:  $0.5 * (-71) + 0.25 * (-81) + 0.25 * (-99) = -81$ .

So for March 2012, we estimate the number of owner occupied dwellings to be the count for March 2011, adjusted by the expected change =  $14,890 - 81 = 14,810$ .

For October 2012, we estimate the number of owner occupied dwellings to be the count for March 2012 plus half the expected change between March 2011 and March 2012 =  $14,810 - 40 = 14,769$  (numbers are rounded).

5.17 The new method will generate more credible estimates of the dwelling counts as it allows for dwellings to switch tenure.

**Change 4: Moving the address-to-dwelling (stage N5) and dwelling-to-household (stage N6) weighting stages.**

5.18 The address-to-dwelling and dwelling-to-household weighting stages previously took place before the non-response stages. This meant that the estimates for the number of dwellings at an address and number of households at a dwelling were required for non-productive addresses. These rely on the interviewer reports, but in practice this information is often not available and most were imputed to be 1. Therefore, the previous method probably under-estimated the address-to-dwelling and dwelling-to-household ratios for the non-productive addresses.

- 
- 5.19 In the new method, these stages are moved after the non-response stages, so that there is no longer a need to derive/impute these estimates for the non-productive occupied addresses<sup>8</sup>, thus removing any bias resulting from this.
- 5.20 The non-response models fitted in CHAID are not weighted, so changing the way these ratios are calculated makes no difference in terms of the model fitted (the model is fitted at the address-level). The dwelling and household weights are only used when generating the grossing weights within each weighting cell generated by CHAID.

### ***Change 5: Adjusting population estimates (calibration stages)***

- 5.21 For the control totals for the one-year weights, the EHS uses 1 October as its reference date. In the previous method, the LFS household counts were taken from the Oct-Dec release of the LFS and the DCLG estimates were calculated as at 1 October. This was also the case for the population counts: mid-year population estimates were adjusted using population projections so that they were also estimates for 1 October. This was quite a considerable amount of work, as the two sets of population estimates are in completely different formats – in fact it takes a day to generate the control totals, using SPSS code.
- 5.22 However, there are two points to note:
- (a) The population estimates are really 'guesstimates' – the difference between the mid-year population estimates and the October projected estimates is negligible compared with the error around these estimates.
  - (b) The actual raw counts are not being used. This is because the population counts are adjusted to be consistent with the DCLG dwelling counts<sup>9</sup>. What is being used is the age/sex distribution (i.e. the percentage in each age/sex category), which is unlikely to change rapidly in the short term.
- 5.23 It was decided to use the unadjusted mid-year population estimates for the new method. The change in the distribution from including a small contribution from the population projections to the mid-year estimates is negligible, and there is no impact on bias from doing this.

---

<sup>8</sup> It is still necessary to estimate the number of dwellings for non-productive vacant addresses at stage N5. This is because the address-to-dwelling ratio (stage N5) needs to be estimated for all vacant dwellings to be used at the sub-sampling stage (stage N8).

<sup>9</sup> The adjustment ensures that the total population based on ONS mid-year population estimates is consistent with the total population derived from DCLG dwelling counts (when the DCLG dwelling counts are multiplied with the weighted estimate of the average number of people per occupied dwelling).

### **Change 6: Combining categories for the subsampling of the physical survey stage (stage N8)**

5.24 At the physical survey stage, a proportion of dwellings that were identified in the interview survey as being owner-occupied are randomly excluded (this includes some vacant dwellings as well); the sub-sampling rate for inclusion of the owner-occupied dwellings varied by quarter. To adjust for this, a selection weight is generated at Stage N8. In the previous method, this was based on ten categories of tenure or likely tenure, repeated for each quarter separately – in other words for 40 categories in total. However, some of these categories were not that prevalent; in fact 17 of them contained fewer than 50 dwellings (see Table 5.5). Therefore, some of these adjustments (those with selection weight >1) would be adding ‘noise’ (i.e. variance) to the weights.

**Table 5.5 Number of interviews by tenure or likely tenure categories used in PS sub-sampling (previous method)**

<b>category</b>	<b>Q1</b>	<b>Q2</b>	<b>Q3</b>	<b>Q4</b>
Interviewed; owned	1699	2032	2116	2433
Interviewed; private rented	518	502	489	575
Vacant/derelict; previously owned	60	59	46	56
Vacant/derelict; previously probably owned	67	57	67	67
Vacant/derelict; previously probably private rented	59	51	45	40
Vacant/derelict; previously probably social rented	19	21	20	22
Vacant/derelict; previously LA	21	15	13	14
Vacant/derelict; previously RSL	15	14	12	9
Vacant/ derelict; previously private rented or unknown	66	69	41	39
Interviewed: social rented	870	834	761	821

5.25 In 2012-13 (and also in 2013-14), sub-sampling was only carried out for the owner-occupied dwellings, so we would expect the screening rate to be 1 for all other tenures – any deviation from this would be due to errors such as miss-classification. Hence it is only necessary to consider the screening for the owner-occupied dwellings for each quarter separately.

5.26 Therefore, the new method uses only the following nine categories:

- 
- Four categories (one for each quarter) for: occupied, interviewed and owner-occupied;
  - Four categories (one for each quarter) for: vacant and previously (or likely to be previously) owner-occupied;
  - One category for all other tenures and all quarters.

5.27 It is possible that sub-sampling for private rented dwellings might be required in the future, in which case the categories would be increased accordingly.

***Change 7: Using the more accurate variable for tenure in PS co-operation stage (stage N9)***

5.28 In the previous method, the non-cooperation model for the physical survey used the predicted tenure that was derived from Experian based on the postcode of the address. The measure of tenure shown in Table 5.5 above is based on information collected in the interview for occupied dwellings and by the interviewer for vacant dwellings. Therefore, it is a much more accurate measure of tenure for using in the non-response CHAID model<sup>10</sup>.

---

<sup>10</sup> After expanding it so that there are separate categories for 'vacant/derelict and private rented' & 'vacant/ derelict and DK' and also for 'IS and local authority' & 'IS and housing association. This change will be introduced in 2014-15.

## Annex Comparison tables of previous and new weighting

**Table A: Comparison of household estimates from the EHS 2012-13 full household sample using previous and new weighting strategies**

measure	previous weights		new weights		change	
	thousands	percent-ages	thousands	percent-ages	thousands	percent-ages
<b>tenure</b>						
own outright	7,152	32.5	7,188	31.8	36	0.5
buying with mortgage (inc shared ownership)	7,184	32.7	7,129	31.6	-55	-0.8
<b>all owner occupiers</b>	<b>14,336</b>		<b>14,317</b>		<b>-19</b>	<b>-0.1</b>
local authority	1,684	7.7	1,644	7.3	-40	-2.4
housing association	2,000	9.1	2,262	10.0	262	13.1
private renter unfurnished	3,138	14.3	3,457	15.3	319	10.2
private renter furnished	819	3.7	899	4.0	80	9.8
<b>all private renters</b>	<b>3,957</b>		<b>4,356</b>		<b>399</b>	<b>10.1</b>
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>type of accommodation</b>						
detached house or bungalow	4,914	22.4	4,987	22.1	73	1.5
semi-detached	6,618	30.1	6,744	29.9	126	1.9
terrace/end of terrace	5,975	27.2	6,183	27.4	208	3.5
<b>all houses</b>	<b>17,507</b>		<b>17,914</b>		<b>407</b>	<b>2.3</b>
purpose built flat/maisonette	3,423	15.6	3,556	15.8	133	3.9
flat conversion/rooms	967	4.4	1,027	4.6	60	6.2
<b>all flats</b>	<b>4,390</b>		<b>4,583</b>		<b>193</b>	<b>4.4</b>
caravan or boat	54	0.2	54	0.2	0	0.0
other	24	0.1	26	0.1	2	8.3
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>age of household reference person</b>						
16 - 24	857	3.9	897	4.0	40	4.7
25 - 34	3,200	14.6	3,294	14.6	94	2.9
35 - 44	4,065	18.5	4,179	18.5	114	2.8
45 - 54	4,285	19.5	4,401	19.5	116	2.7
55 - 64	3,615	16.4	3,716	16.5	101	2.8
65 or over	5,956	27.1	6,090	27.0	134	2.2
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>household (HRP + partner) receives any housing benefit?</b>						
yes	3,423	44.0	3,681	43.9	258	7.5
no	4,349	56.0	4,710	56.1	361	8.3
<b>all renters</b>	<b>7,772</b>	<b>100.0</b>	<b>8,391</b>	<b>100.0</b>	<b>619</b>	<b>8.0</b>

**Table A: Comparison of household estimates from the EHS 2012-13 full household sample using previous and new weighting strategies (continued)**

measure	previous weights		new weights		change	
	thousands	percent-ages	thousands	percent-ages	thousands	percent-ages
<b>bedroom standard</b>						
two or more below standard	75	0.3	73	0.3	-2	-2.7
one below standard	602	2.7	629	2.8	27	4.5
at standard	5,659	25.8	5,935	26.3	276	4.9
one above standard	7,556	34.4	7,754	34.3	198	2.6
two or more above standard	8,084	36.8	8,187	36.3	103	1.3
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>ethnic origin of HRP</b>						
white	19,685	89.6	20,227	89.6	542	2.8
ethnic minority	2,292	10.4	2,350	10.4	58	2.5
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>if first-time buyer</b>						
first-time buyer	5,266	24.0	5,205	23.1	-61	-1.2
owner but has owned previously	9,069	41.3	9,111	40.4	42	0.5
non-owner	7,636	34.8	8,257	36.6	621	8.1
<b>all households</b>	<b>21,972</b>	<b>100.0</b>	<b>22,572</b>	<b>100.0</b>	<b>600</b>	<b>2.7</b>
<b>area</b>						
North	6,340	28.8	6,527	28.9	187	2.9
Midlands	4,213	19.2	4,231	18.7	18	0.4
London	3,176	14.5	3,333	14.8	157	4.9
Rest of South	8,248	37.5	8,486	37.6	238	2.9
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>household composition</b>						
couple, no dependent child(ren) under 60	4,018	18.3	4,091	18.1	73	1.8
couple, no dependent child(ren) aged 60 or over	3,803	17.3	3,906	17.3	103	2.7
couple with dependent child(ren)	4,579	20.8	4,673	20.7	94	2.1
lone parent with dependent child(ren)	1,538	7.0	1,618	7.2	80	5.2
other multi-person households	1,821	8.3	1,879	8.3	58	3.2
one person under 60	2,854	13.0	2,980	13.2	126	4.4
one person aged 60 or over	3,365	15.3	3,431	15.2	66	2.0
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>household size</b>						
one	6,218	28.3	6,411	28.4	193	3.1
two	7,885	35.9	8,113	35.9	228	2.9
three	3,561	16.2	3,665	16.2	104	2.9
four	2,936	13.4	2,986	13.2	50	1.7
five	980	4.5	1,001	4.4	21	2.1
six or more	397	1.8	402	1.8	5	1.3
<b>all households</b>	<b>21,977</b>	<b>100.0</b>	<b>22,578</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>

**Table A: Comparison of household estimates from the EHS 2012-13 full household sample using previous and new weighting strategies (continued)**

measure	previous weights		new weights		change	
	<i>thousands</i>	<i>percent-ages</i>	<i>thousands</i>	<i>percent-ages</i>	<i>thousands</i>	<i>percent-ages</i>
<b>length of residence</b>						
less than 1 year	2,286	10.4	2,434	10.8	148	6.5
one year	1,676	7.6	1,758	7.8	82	4.9
two years	1,440	6.6	1,504	6.7	64	4.4
3-4 years	1,904	8.7	1,985	8.8	81	4.3
5-9 years	3,970	18.1	4,053	18.0	83	2.1
10-19 years	4,691	21.4	4,761	21.1	70	1.5
20-29 years	2,560	11.7	2,594	11.5	34	1.3
30+ years	3,440	15.7	3,481	15.4	41	1.2
<b>all households</b>	<b>21,969</b>	<b>100.0</b>	<b>22,570</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>
<b>previous tenure</b>						
new household	722	3.3	757	3.4	35	4.8
owned outright	487	2.2	497	2.2	10	2.1
buying with a mortgage	925	4.2	944	4.2	19	2.1
owned, unknown if outright or mortgage	10	0.0	11	0.0	1	10.0
local authority	367	1.7	388	1.7	21	5.7
housing association	357	1.6	391	1.7	34	9.5
private rented	2,495	11.4	2,668	11.8	173	6.9
rented, landlord unknown	4	0.0	4	0.0	0	0.0
household reference person resident >= 3 years	16,589	75.6	16,897	74.9	308	1.9
<b>all households</b>	<b>21,955</b>	<b>100.0</b>	<b>22,556</b>	<b>100.0</b>	<b>601</b>	<b>2.7</b>

**Table B: Comparison of previous and new weighting strategies for estimates from the EHS 2012-13 paired dwelling sample<sup>11</sup>**

measure	previous weights		new weights		change	
	thousands	percent-ages	thousands	percent-ages	thousands	percent-ages
<b>dwelling type</b>						
small terraced house	2,111	9.3	2,147	9.3	36	1.7
medium/large terraced house	4,050	17.8	4,064	17.5	14	0.3
semi-detached house	5,906	26.0	5,968	25.7	62	1.1
detached house	4,034	17.8	4,046	17.4	12	0.3
bungalow	2,042	9.0	2,123	9.2	81	4.0
<b>all houses</b>	<b>18,143</b>		<b>18,348</b>		<b>205</b>	<b>1.1</b>
converted flat	918	4.0	991	4.3	72	7.9
purpose built flat, low rise	3,242	14.3	3,431	14.8	189	5.8
purpose built flat, high rise	391	1.7	425	1.8	34	8.6
<b>all flats</b>	<b>4,552</b>		<b>4,847</b>		<b>295</b>	<b>6.5</b>
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>dwelling age</b>						
pre 1919	4,498	19.8	4,623	19.9	124	2.8
1919-44	3,809	16.8	3,880	16.7	71	1.9
1945-64	4,557	20.1	4,701	20.3	144	3.2
1965-80	4,716	20.8	4,730	20.4	14	0.3
1981-90	1,930	8.5	2,002	8.6	72	3.7
post 1990	3,186	14.0	3,260	14.1	74	2.3
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>usable floor area</b>						
less than 50m <sup>2</sup>	2,678	11.8	2,858	12.3	180	6.7
50 to 69m <sup>2</sup>	5,368	23.7	5,520	23.8	152	2.8
70 to 89m <sup>2</sup>	6,209	27.4	6,280	27.1	71	1.1
90 to 109m <sup>2</sup>	3,025	13.3	3,055	13.2	30	1.0
110 m <sup>2</sup> or more	5,415	23.9	5,482	23.6	67	1.2
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>type of boiler</b>						
standard boiler (floor or wall)	5,329	23.5	5,547	23.9	218	4.1
back boiler (to fire or stove)	839	3.7	858	3.7	20	2.3
combination boiler	3,410	15.0	3,444	14.8	35	1.0
condensing boiler	2,994	13.2	3,017	13.0	23	0.8
condensing-combination boiler	7,601	33.5	7,646	33.0	45	0.6
<b>total</b>	<b>20,172</b>	<b>88.9</b>	<b>20,512</b>	<b>88.4</b>	<b>340</b>	<b>1.7</b>
no boiler	2,523	11.1	2,683	11.6	160	6.3

<sup>11</sup> The aim of this analysis was to demonstrate differences in the estimates between the previous and new weighting strategies. For this reason, one-year estimates were used as these were sufficient for this comparison.

**Table B: Comparison of previous and new weighting strategies for estimates from the EHS 2012-13 paired dwelling sample (continued)**

measure	previous weights		new weights		change	
	<i>thousands</i>	<i>percent-ages</i>	<i>thousands</i>	<i>percent-ages</i>	<i>thousands</i>	<i>percent-ages</i>
<b>type of wall and insulation</b>						
cavity with insulation	9,281	40.9	9,498	41.0	217	2.3
cavity uninsulated	6,224	27.4	6,280	27.1	55	0.9
other	7,190	31.7	7,417	32.0	227	3.2
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>main heating system</b>						
boiler system with radiators	20,071	88.4	20,405	88.0	335	1.7
storage radiators	1,407	6.2	1,502	6.5	95	6.7
warm air system	129	0.6	136	0.6	7	5.6
room heater	670	3.0	693	3.0	23	3.4
other systems	21	0.1	20	0.1	0	-1.6
communal	398	1.8	439	1.9	41	10.3
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>energy efficiency (SAP09) rating</b>						
less than 30	579	2.6	604	2.6	25	4.3
30 to 50	3,532	15.6	3,608	15.6	76	2.1
51 to 70	15,194	66.9	15,498	66.8	304	2.0
more than 70	3,390	14.9	3,485	15.0	95	2.8
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>decent homes HHSRS 15 model</b>						
decent	17,999	79.3	18,363	79.2	364	2.0
non-decent	4,696	20.7	4,832	20.8	136	2.9
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>
<b>decent homes HHSRS 26 criterion</b>						
pass	19,793	87.2	20,214	87.1	420	2.1
fail	2,902	12.8	2,981	12.9	80	2.7
<b>all dwellings</b>	<b>22,695</b>	<b>100.0</b>	<b>23,195</b>	<b>100.0</b>	<b>500</b>	<b>2.2</b>