Take-Up Technical Annex: Methodology for Estimating Take-Up of Income Related Benefits

February 2012

DWP Department for Work and Pensions

Contents

1 Introduction

1.1 Purpose of document

This document describes the methodology used when producing the National Statistics publication: Income Related Benefits: Estimates of Take-up. The latest publication is available on the internet at: http://research.dwp.gov.uk/asd/index.php?page=irb.

This document is designed to be an in-depth methodology guide to producing and assuring the statistics. It may also be appropriate for other interested analysts, for example specialist users or producers of other statistics. Basic familiarity with the publication and methodology itself is assumed.

For a less detailed statement of our methodology one might also refer to the Methods and Data Sources Chapter of the publication itself and to the Appendix of the publication which provides a worked example of calculating take-up ranges.

The chapters of this document cover the methodology applied in broadly chronological order in terms of how the publication is produced. The document does not have specific references to SAS techniques, programs or datasets, so that it can be widely understood. Details of the programs used are available on request from irb.takeup@dwp.gsi.gov.uk.

2 Administrative data

We use administrative data to tabulate the number of benefit recipients and the average amounts claimed for each benefit estimates are produced for, split by family type, tenure type and employment status (for Housing Benefit only).

In each case we make exclusions from the data to be compatible with the scope of the Family Resources Survey (FRS). There are benefit specific considerations which are described in the relevant section below.

The general style in the notation below is to use capital letters for quantities from 100 per cent administrative data, and to use lower case letters for 5 per cent administrative data, 1 per cent administrative data or survey data.

2.1 Income Support (IS) & Employment and Support Allowance (Income-Related) (ESA (IR))

There are five splits of interest for these benefits (eight including totals and sub-totals):

- All families.
  - Families of working-age with children.
    - Couples with children.
    - Singles with children.
  - Families of working age without children.
    - Couples.
    - Single males.
    - Single females.

The benefit specific considerations are:

- We only focus on income-related Employment and Support Allowance (ESA (IR)) for the purposes of the publication.
- For the couple family types we use two years' data to get large enough sample sizes.
- We are reporting these two benefits jointly, so the information needs to be combined appropriately.

The exclusions required for IS and ESA (IR) are:

- 16 and 17 year olds without dependents (because the special housing circumstances which would make them eligible for JSA or Income Support cannot be modelled on the FRS).
- Cases with very high housing costs, above the 99th percentile, are excluded (because this is done for FRS entitled non-recipient average amounts to avoid distortion of estimates).
- Residents in non-private households, as the FRS does not cover these.

Table 2.2.1 shows how the caseloads for each family type are obtained for IS and ESA (IR) taking these exclusions and benefit specific considerations into account and Table 2.2.2 shows how the average amounts claimed are obtained. The quantities in the shaded cells (R and £R) are the ones that are taken forward into the calculation and the other quantities are no longer required.

2.1.1 IS and ESA (IR) caseload from administrative data

I	IS recipients according to the 100 per cent WPLS data source. Average of 4 quarters for singles but 8 quarters for couples as we need to use two years data to get a sufficient sample size. Note that exclusions on the basis of age have been applied to this quantity, so are not applied again using the QSE.
E	Employment and Support Allowance recipients according to 100 per cent data. Note that exclusions on the basis of age have been applied to this quantity. These figures are available from November 2008 only.
i	Count of IS cases in the 5 per cent QSE data that can be included in the analysis (if not excluded on the basis of age, housing costs, or special circumstances). Average of 4 quarters for singles or 8 quarters for couples (as using two years data for couples).
x	Count of cases in the 5 per cent QSE data that should be excluded because they have exceptionally high housing costs (above the 99th percentile) or because they are in special circumstances. Average of 4 quarters for singles or 8 quarters for couples (as using two years data for couples).
f	**= i / (i + x).** Factor to be applied to the total obtained from 100 per cent WPLS data to further exclude cases with exceptionally high housing costs or 'special circumstances' (non-private households). Note that this factor is entirely based on Income Support, since we do not have the corresponding information for Employment and Support Allowance.
R	**=f * (I + E).** This is the quantity to be taken forward in each case: Total caseload for IS and ESA (IR) based on 4 or 8 quarters as appropriate with exclusions applied, first using WPLS information and then using QSE information. This is taken forward into the calculation.

2.1.2 IS and ESA (IR) average amounts claimed from administrative data

£i	Average amount of Income Support claimed by recipients in the 5 per cent QSE data that can be included in the analysis (not excluded on the basis of age, housing costs, or special circumstances). Average of 4 quarters for singles or 8 quarters for couples.
£E	Average amount claimed by Employment and Support Allowance recipients according to 100 per cent data. Note that exclusions on the basis of age have been applied to this quantity. These figures are available from November 2008 only.
i	Count of IS cases in the 5 per cent QSE data that can be included in the analysis (if not excluded on the basis of age, housing costs, or special circumstances). Average of 4 quarters for singles or 8 quarters for couples (as using two years data for couples).
E	Employment and Support Allowance recipients according to 100 per cent data. Note that exclusions on the basis of age have been applied to this quantity. These figures are available from November 2008 only.

£R	**=(E*£E + i*£i) / (i + E).** Weighted average of the amount IS and ESA claimed. This is taken forward into the calculation. This averages WPLS information for ESA (IR) and QSE information for IS. The QSE information is better because it has all the exclusions applied, but for ESA we have only WPLS available.

2.2 Pension Credit (PC)

At the lowest level there are 9 splits we are interested in for Pension Credit (16 including totals and sub-totals):

- All Pension Credit.
  - All Pension Credit for Couples.
  - All Pension Credit for Single Males.
  - All Pension Credit for Single Females.
- Guarantee Credit only (GC).
  - Guarantee Credit only for Couples.
  - Guarantee Credit only for Single Males.
  - Guarantee Credit only for Single Females.
- Guarantee and Savings Credit (GCSC).
  - Guarantee and Savings Credit for Couples.
  - Guarantee and Savings Credit for Single Males.
  - Guarantee and Savings Credit for Single Females.
- Savings Credit only (SC).
  - Savings Credit only for Couples.
  - Savings Credit only for Single Males.
  - Savings Credit only for Single Females.

The benefit specific consideration here is that we need to make special adjustment for backdated claims.

The exclusions required for PC are:

- Cases with very high housing costs, above the 99th percentile, are excluded (because this is done for FRS entitled non-recipient average amounts to avoid distortion of estimates).
- Residents in non-private households, as the FRS does not cover these.

Table 2.2.1 below shows how the caseloads are obtained for PC taking these exclusions and benefit specific considerations into account. The quantities in the shaded cells (R and BK) are the ones that are taken forward into the calculation, and the other quantities are no longer required.

2.2.1 PC caseload from administrative data

P	Pension Credit recipients according to 100 per cent WPLS data. Note that this is not from the commonly available National Statistics frozen datasets but rather from a special version of the dataset based on the entitlement start date rather than the claim start date. Based on 4 quarters. There are no exclusions applied to this quantity at this stage.
i	Pension Credit recipients who are not excluded for statistical purposes, that is, not excluded on the basis of housing costs or special circumstances. Based of 4 quarters of the 5 per cent QSE.

x	Pension Credit recipients who are excluded for statistical purposes, that is, excluded on the basis of housing costs or special circumstances. Based of 4 quarters of the 5 per cent QSE.
f	**= i / (i + x).** Proportion of recipients who do not need be excluded, estimated from the QSE
R	**= P * f.** This is the quantity to be taken forward. It is the total number of Pension Credit recipients according to WPLS, using entitlement start date, adjusted for high rents and special circumstances.
Q	Pension Credit recipients according to the commonly available National Statistics frozen datasets based on claim start date rather than entitlement start date. Based on 4 quarters. There are no exclusions applied to this quantity at this stage.
BK	**= f * (P – Q) / R.** The gap between caseload including backdaters, and the caseload excluding backdaters, expressed as proportion of the caseload which includes backdaters and adjusted for high rents and special circumstances.

The average amounts of PC claimed (£P) are drawn on exactly the same basis as the initial number of cases P. So they are based on entitlement start date, but are not adjusted for exclusions.

2.3 Housing Benefit (HB) and Council Tax Benefit (CTB)

For both these benefits – which can be claimed by any qualifying adult - there is a detailed split by family type, and a split by tenure type:

- Total.
  - Pensioners.
  - Non-Pensioners.
    - Couples with Children.
    - Singles with children.
    - Non-pensioners without children.
- Total.
  - Social Rented.
  - Private Rented.
  - Owner Occupier (CTB only).

For Housing Benefit (HB) there is a further detailed split of Non-Pensioners:

- All Non-pensioners.
  - In Employment.
  - Not in employment.

There are no benefit specific considerations for HB or CTB.

The exclusions required for HB and CTB are:

- 16 and 17 year olds without dependents (because the special housing circumstances which would make them eligible for some benefits cannot be modelled on the FRS).
- Cases with very high housing costs, above the 99th percentile, are excluded (because this is done for FRS entitled non-recipient average amounts to avoid distortion of estimates).
- Full-time self-employed as their income on the FRS is very difficult to assess.
- Cases claiming Second-Adult Rebate are not included for Council Tax Benefit only.

- Owner-occupiers are excluded from HB only as they are not assessed on the FRS because they should not be entitled according to benefit regulations.
- Cases receiving less than 50p of Housing Benefit are excluded from Housing Benefit only.

The FRS does not cover residents in non-private households but it is no longer possible to exclude this group for HB and CTB as with the other benefits, as changes to the data sources mean it is no longer possible to identify non-private households. The effect of this change is estimated to be very small.

Table 2.3.1 shows how the caseloads for each split are obtained for HB taking these exclusions and benefit specific considerations into account and Table 2.3.2 shows how the average amounts claimed are obtained. The quantities in the shaded cells (R and £R) are the ones that are taken forward into the calculation and the other quantities are no longer required.

The method is identical for Council Tax Benefit.

2.3.1 HB caseload from administrative data

H	Housing Benefit recipients according to the 100 per cent Single Housing Benefit Extract (SHBE). Based on an average of 4 quarters. Does not include cases where rent exceeds the 99th percentile of rents, or cases where all the adults are under 18 and have no dependents
n	The number of HB recipients on the Family Resources Survey in the given year who are not in full-time self-employment.
s	The number of HB recipients on the Family Resources Survey in the given year who are in full-time self-employment.
f	**= n / (s + n).** The proportion of HB recipients on the Family Resources Survey not in full-time self-employment.
R	**= H * f .** This is the value to be taken forward in each case. It is the number of benefit recipients in the year, averaged over 4 quarters, and adjusted for high rents, under 18s, the full-time self-employed and special circumstances not covered by the FRS.

2.3.2 HB average amounts claimed from administrative data

£H	The average amount claimed by HB recipients according to the 100 per cent Single Housing Benefit Extract (SHBE). Based on an average of 4 quarters. Does not include cases where rent exceeds the 99th percentile of rents, or cases where all the adults are under 18 and have no dependents.
£n	The average amount claimed by HB recipients on the Family Resources Survey in the given year who are not in full-time self-employment.
£t	The average amount claimed across all HB recipients on the Family Resources Survey in the given year, both self-employed and not.
a	**= £n / £t.** The relative size of awards to those who are not full-time self-employed as compared to the whole population. This is expected to be close to 1, but could be more or less than 1.
£R	**= H * a .** This is the value to be taken forward in each case. It is the average amount claimed according to SHBE adjusted slightly for the effect of excluding the full-time self-employed. The adjustment factor is based on the FRS.

2.4 Jobseeker's Allowance (JSA)

There are only three working-age splits of interest for this benefit (four including totals and sub-totals):

- All Jobseeker's Allowance.
  - Single males without children.
  - Single females without children.
  - Couples with children.

Estimates of take-up by couples without children and singles with children are not been presented as they are not statistically robust.

The benefit specific considerations are:

- We focus only on JSA Income-Based (JSA (IB)) for the purposes of the publication.
- For couples with children we need to combine two years' data to get large enough sample sizes.

The exclusions required for JSA (IB) are:

- 16 and 17 year olds without dependents (because the special housing circumstances which would make them eligible for JSA IB, Income Support or Employment and Support Allowance cannot be modelled on the FRS).
- Cases with very high housing costs, above the 99th percentile, are excluded (because this is done for FRS entitled non-recipient average amounts to avoid distortion of estimates).
- Residents in non-private households, as the FRS does not cover these.

Table 2.4.1 below shows how the caseloads are obtained for JSA (IB) taking these exclusions and benefit specific considerations into account. The quantities in the shaded cells (RC, RM and RF) are the ones that are taken forward into the calculation, and the other quantities are no longer required.

2.4.1 JSA (IB) caseload from administrative data

J1	Total JSA caseload (all families, no exclusions) based on 1 year Work and Pensions Longitudinal Study (WPLS) 100 per cent data (4 scans)
J2	Total JSA caseload (all families, no exclusions) based on 2 years WPLS 100 per cent data (8 scans)
j1	Total JSA caseload (all families, no exclusions) based on 1 year Quarterly Statistical Enquiry (QSE) 5% data (4 scans)
j2	Total JSA caseload (all families, no exclusions) based on 2 years QSE 5 per cent data (8 scans)
c2	JSA(IB) couples, exclusions removed, based on 2 years QSE
m1	JSA(IB) males, exclusions removed, based on 1 year QSE
f1	JSA(IB) females, exclusions removed, based on 1 year QSE
RC	**=(c2/j2) * J2.** Total number of couple recipients of JSA (IB) for take-up purposes. This is calculated by multiplying the proportion of the QSE total caseload that are estimated to be couples on JSA (IB) who are not excluded (**c2/j2**) by the total caseload on the 100 per cent WPLS data (**J2**).
RM	**=(m1/j1) * J1.** Total number of single male recipients of JSA (IB) for take-up purposes. This is calculated by multiplying the proportion of the QSE total caseload that are estimated to be single males on JSA (IB) who are not excluded (**m1/j1**) by the total caseload on the 100 per cent WPLS data (**J1**).
RF	**=(f1/j1) * J1.** Total number of single female recipients of JSA (IB) for take-up purposes. This is calculated by multiplying the proportion of the QSE total caseload that are estimated to be single females on JSA (IB) who are not excluded (**f1/j1**) by the total caseload on the 100 per cent WPLS data (**J1**).

Notice that for JSA the factors (**c2/j2**), (**m1/j1**) and (**f1/j1**) are not only removing these specific exclusions from the 100 per cent WPLS total, but are also removing Contributory JSA cases as it is not possible to distinguish between Income Based and Contributory Based JSA from WPLS data.

In terms of average amounts claimed, these are drawn from 5 per cent QSE data using 8 quarters for couples and 4 quarters for singles. In other words the average amounts are on the same basis as **c2**, **m1** and **f1**.

3 Family Resources Survey data

3.1 Identifying Entitled Non-Recipients

The survey data are used specifically to tell us about those who are not receiving benefit but who may be entitled to do so. These are referred to as 'entitled non-recipients' (ENRs) throughout this report. As they are not receiving the benefit, this population cannot be obtained using administrative data which is why we need to rely on survey data.

In order to obtain an estimate of the number of ENRs we use a customised version of the Policy Simulation Model, a DWP static micro-simulation model. The model looks at the reported information about incomes and personal circumstances for each respondent of the FRS. The aim is to mimic the benefit regulations and identify the income-related benefits the respondent would be entitled to if they were to apply.

This modelled entitlement information can then be compared to the reported information about the benefits that they are actually receiving, and an initial estimate of ENRs can be obtained.

The estimates we obtain from survey data are described in Table 3.1.1 below. It is important to consider whether the quantities are grossed up to the population using the survey grossing factors, or whether they are based just on the number of sample cases. The gross column in the table below indicates where quantities are grossed up to population level.

The style is to denote all these quantities with lower cases letters, for example **r,** because they are obtained from sample data, rather than 100 per cent administrative data which is denoted with capital letters, for example **R**.

3.1.1 Table of estimates obtained from the Family Resources Survey

Estimate	Description	Gross?
enr	'entitled non-recipient' - modelled as entitled to the benefit of interest, but not reporting receipt thereof.	☑
nenr	'non-entitled non-recipient' - modelled as not entitled to the benefit of interest and not reporting receipt thereof.	☑
ner	'non-entitled recipient' - modelled as not entitled to the benefit of interest but reporting receipt thereof.	☑
er	'entitled recipient' - modelled as entitled to the benefit of interest, and reporting receipt thereof.	☑
r	**= er + ner** Total of those reporting receipt of the benefit in the survey.	☑
w	Total of those who report to the survey that they are awaiting the outcome of a claim to the benefit, and who on the basis of modelling appear as if they are indeed entitled.	☑
spd	'single person discount' - applies only to Council Tax Benefit. Those cases who report an amount of CTB that is within 10 pence of 25 per cent of their Council Tax liability for the property. There is a chance that these cases are misreporting the 25 per cent single person discount as Council Tax Benefit.	☑
v	This is the proportion of sample cases who report receipt of the benefit for which the modelled level of entitlement exceeds the reported level of receipt by more than 10 pence. That is, the proportion cases who are **over modelled**.	☑
u	This is the proportion of sample cases who report receipt of the benefit for which the reported level of receipt exceeds the modelled level of entitlement by more than 10 pence. That is, the proportion cases who are **under modelled**.	☑
iro	This relates to benefit confusion. It is the number of cases who appear to be entitled to but not receiving the benefit in question, but could be confusing it with another benefit they are receiving. This is expressed as a proportion of **R** (the number of recipients of the benefit according to admin data). For example, the benefits we imagine might be confused with Jobseekers Allowance are: Attendance Allowance, Disability Living Allowance, Incapacity Benefit, Carer's Allowance, Severe Disablement Allowance, Maternity Allowance, Guardian's Allowance, Industrial Injuries Disablement Benefit, and Government Training Scheme payments.	☑
tu	**= R / (R + enr)** An initial point estimate of caseload **take-up**, at this point unadjusted for non-sampling or sampling error.	
n	Sample cases of entitled recipients and entitled non-recipients i.e. those modelled as entitled to the benefit of interest.	
x	Sample cases modelled as entitled non-recipients.	
£enr	Mean of the values of modelled benefit entitlement among those who are modelled as entitled but not reporting receipt.	☑
sd(£enr)	This is the standard deviation of the values of modelled benefit entitlement among those who are modelled as entitled but not reporting receipt.	

4 Special Case: Pension Credit

4.1 Data linking and hidden receipt

A particular limitation we are aware of in terms of the survey data we are using is that the grossed up number of people who claim Pension Credit is lower than the number that we know are claiming according to the administrative sources. Comparisons of FRS and administrative data show a 32 per cent under count in 2009-10[1].

This is one of the worst effected benefits captured by the Family Resources Survey. As such we seek to address this at a micro level by matching survey respondents to their administrative records, and in so doing identifying so called 'hidden recipients' among the pensioners who have not reported receipt. The proportions of cases that fall into this category are then taken forward into the error framework below.

A benefit unit is classed as a 'hidden recipient' if they do not report receipt of Pension Credit on the FRS but when linked to administrative date are found to have either:

- Started receiving Pension Credit before the FRS interview; or
- Started entitlement to Pension Credit before the FRS interview and since had payment backdated.

There are, however, some extra methodological steps involved in this because we do not always have the explicit consent that we need from the survey respondents in order to be able to match their records.

4.2 Imputation for non consenters

We cannot directly identify 'hidden receipt' of Pension Credit where we do not have the consent to match. This section describes how we impute 'hidden recipients' onto the non-matched proportion of the survey respondents.

We subdivide the non-recipient population into 3 x 4 = 12 groups

- Three family types (single male, single female, couple)
- Four non-recipient categories (ENR of GC, ENR of GCSC, ENR of SC or an NENR)

Within each of the 12 non-recipient groups we look at those benefit units who consented to matching and establish what proportion of these turn out to be recipients (i.e. what proportion were revealed as hidden recipients).

Within each of the 12 groups we also need to calculate how many benefit units did not consent to the matching.

Our intention is to convert some of these non-consenters into hidden recipients. So we need to decide:

- how many cases to convert; and
- which cases to convert

[1] Family Resources Survey 2009/10, Table M.6, available at the URL: http://research.dwp.gov.uk/asd/frs/

4.2.1 How many cases to Impute

To decide how many cases to convert our agreed methodology is that within each of the 12 groups we should assume that hidden receipt occurs at the same rate in the consenting and non-consenting portions of the non-recipient populations.

So within each group we multiply the rate of hidden receipt among the consenters by the number of non-consenters.

4.2.2 Which cases to Impute

To decide which cases we have created a statistical model that assigns everyone a probability of being a hidden recipient.

We then create a new variable (called ORDER) where we subtract a random number between 0 and 1 from the modelled probability. We then select the benefit units with the highest value of this new ORDER variable to be converted from non-recipients to recipients.

This leads us to new variables which reflect

- hidden receipt revealed by the datamatching (among consenters) and
- imputed hidden receipt (among non-consenters)

Table 4.2.3 shows the estimates obtained for hidden recipients among both consenters and non-consenters, the shaded cells are the ones that are taken forward into the error framework later.

4.2.3 Table of estimates obtained from data-linking and imputation

max_enr_hr	This is the measure of the number of hidden recipients, both observed and imputed, among those who are modelled as entitled to but not receiving Pension Credit. The hidden recipients are defined according to the entitlement start date. The measure is the grossed number of such cases expressed as a proportion of **R**, the number of recipients according to administrative data. The minimum is obtained using only the very best matches, and the maximum is obtained by using all acceptable matches.
min_enr_hr	
min_nenr_hr	This is very similar to above; except that it measures hidden receipt among those initially modelled as not entitled to and not receiving Pension Credit (**NENR** as opposed to **ENR**). We do not expect to see many of these cases as it would suggest an error in our modelling of entitlement or an erroneous claim.
max_nenr_hr	

4.3 Special assumptions about capital under-reporting

We have particular evidence from follow-up studies to the Family Resources Survey (FRS) that pensioners can on occasion under-report the amount of capital that they hold[2]. As entitlement to PC takes into account the amount of capital, this can lead to those respondents under-reporting capital being modelled as entitled to Pension Credit when actually they are not.

We use some arbitrary assumptions and a simulation technique to estimate the quantitative effects of this under-reporting of capital in terms of the non-sampling error framework described below.

As informed by the orders of magnitude of misreporting observed in the follow-up study the smallest allowance we made for this effect was to assume that capital was under-reported by a quarter among 9.7 per cent of PC ENR cases. The upper bound to the adjustment allowance was to assume that under-reporting of capital was around a half among 20 per cent of PC ENR cases.

Table 4.3.1 shows the estimates obtained from capital under-reporting, the estimates in the shaded cells (cap_min and cap_max) are the ones that are taken forward into the error framework later, and the other quantities are no longer required

4.3.1 Estimates obtained from capital under-reporting for Pension Credit

shift25	This is the proportion of Pension Credit entitled non-recipients who change status to become a non-entitled non-recipient under the scenario where their reported capital is inflated by either 25 per cent or by 50 per cent
shift50	
xenr_min	**= enr * shift25 * 0.097**
xenr_max	**= enr * shift50 * 0.200** These are minimum and maximum estimates of the numbers of 'false' entitled non-recipients that could have resulted from capital under-reporting.
cap_min	**= xenr_min / (er + enr – xenr_min)**
cap_max	**= xenr_max / (er + enr – xenr_max)** To take forward the assumptions about non-sampling error the methodology expresses these false ENR cases as a proportion of the number of entitled cases. The denominator in each case is itself adjusted for the false ENR cases.

[2] "Entitled but not claiming? Pensioners, the Minimum Income Guarantee and Pension Credit" The report can be found at:
http://research.dwp.gov.uk/asd/asd5/rports2003-2004/rrep197.asp

5 Non-sampling error

5.1 Types of non-sampling error

The methodology requires us to estimate the possible magnitude of any non-sampling error associated with the survey-based estimate of entitled non-recipients (ENR s). The five sources of non-sampling error are described in Table below.

5.1.1 Table of the five sources of non-sampling error

A	**Over-statement of entitlement** - this occurs when a benefit unit that is not truly entitled to the benefit is modelled to be entitled. This could be due to them having under-reported their income or due to some other factor or circumstance that would exclude them from eligibility that is not recorded on the FRS.
B	**Under-reporting of benefit receipt in the FRS** - this occurs when someone receiving the benefit fails to report receipt in the FRS interview.
C	**Under-statement of entitlement** - this occurs when a benefit unit that is truly entitled to the benefit is modelled to be not entitled. This can happen if income is overstated on the FRS or if some factor or circumstance that would make them entitled is not recorded on the FRS.
D1	**Inaccurate grossing-up of FRS counts** - as the FRS is a survey of only a sample of the population, counts derived from the FRS need to be grossed-up - i.e. multiplied up to reflect the true numbers of various family types and people of different ages in the population - to give meaningful estimates of the actual number of recipients or entitled non-recipients in the population. Inaccurate grossing-up, which gives the correct overall population total but inaccurate sub-totals for population subsets, will result in either under or over-estimation of the number of recipients or entitled non-recipients in the population. These are known as error **D1** (under-grossing) and **D2** (over-grossing).
D2	
E	**Payment of benefit to non-entitled benefit units** – may occur from fraudulent claims, or official error on the part of DWP.

5.2 Diagnostic variables

The rules for determining the size of errors A to E depend on a number of diagnostics that can be calculated from the administrative and survey inputs. Table 5.2.1 shows how each of these diagnostics are calculated, refer to Table 3.3.1 for the definition of survey variables used (lower case) and the tables in Chapter 2 for the definition of administrative variables (upper case).

5.2.1 Table of diagnostic variables

p	**= w / R** The number of cases who are awaiting the outcome of a claim to benefit (**w**) and appear to be entitled (based on the grossed-up FRS) expressed as a proportion of the number of recipients according to administrative data (**R**). This is referred to as the pipeline proportion (**p**).
t	**= r / R** for HB In the case of Housing Benefit the **t** statistic is no more than a comparison of the count of recipients obtained from the grossed FRS (**r**) to the count obtained from administrative data (**R**).
	= (r - spd) / R for CTB In the case of Council Tax Benefit there is an adjustment to the calculation for cases who appear to be receiving a single person discount (**spd**) rather than the CTB they are reporting.
	= (r + w) / R for IS, JSA (IB), ESA (IR) In the case of IS, JSA (IB), ESA (IR), this statistic is adjusted for waiting and entitled cases (**w**), as grossed up from the FRS. The adjustment by **w** is not made for HB and CTB because the administrative count **R** does not include such backdated claims, so we are comparing like with like.
t_min	**= ((r + w)/R) + BK + min_enr_hr + min_nenr_hr** for PC For Pension Credit there are three adjustments to the basic definition of **t**: • As for IS, an adjustment is made for waiting and entitled cases (**w**). • The **t** statistic is further boosted to account for a shortfall in **r** caused by hidden recipients among **enr**s (**min_enr_hr**) and hidden recipients among **nenr**s (**min_nenr_hr**). Notice that these extra terms are already expressed as proportions of **R.** • Thirdly the **t** statistic is boosted to account for the fact that **r** will not include backdated claims (**BK**). This is different from hidden recipients because these backdated claims could not possibly have been reported at the FRS interview.
t_max	**= ((r + w)/R) + BK + max_enr_hr + max_nenr_hr** for PC For Pension Credit there is another version of **t** which looks at the maximum possible assumption about hidden receipt. It is otherwise identical to the above.
s	**= ner / r** in general In general the **s** statistic uses the grossed-up FRS to compare the number of non-entitled recipients (**ner**) to the number of recipients (**r**).
	= (ner - spd) / (r - spd) for CTB For Council Tax benefit both the numerator and denominator are adjusted by the number of cases who appear to be receiving single person discount (**spd**) rather than CTB.
X	**= v** **X** is equal to the proportion of FRS recipients whose entitlement has been over-modelled (**v**).
Y	**= 0.5** This is an assumption used in conjunction with error **D** and is always assumed to be 0.5.

5.3 Rules for determining non-sampling error

Having calculated our family of diagnostic variables, these are combined to give our minimum and maximum assumptions about the possible magnitudes of the non-sampling errors A to E.

There is a general set of rules for the benefits where data-linking has not been applied, and a special set of rules for Pension Credit, where data-linking has been applied.

It is very important to remember that these errors are expressed as proportions of the number of benefit recipients.

5.3.1 General Case (Not Pension Credit)

Size of s	Error A min : max	Error C min : max	Error E min : max
< 15%	s/3 : s	s/3 : s	s/3 : s
15% and over	s/3 : s	s/3 : s	s/3 : 15%

Size of t	Error B min	Error B max
<100%	p	(X * (1 - t)) + p
>100%	p	p

Size of t	Error D1 min	Error D1 max
< 90% *for HB, CTB*	Y * (1- (t + Bmax))	(1 - t)
<90% *for IS, JSA, ESA*	Y * (1- (t + Bmax + p))	(1 - t)
90% - 95%	0	(1 - t)
95% - 105%	0	5%
105% - 110%	0	(1 - t) + 10%
>110%	0	0

Size of t	Error D2 min	Error D2 max
< 90%	0	0
90% - 95%	0	(t-1) + 10%
95% - 105%	0	5%
105% - 110%	0	(t-1)
>110%	(t-1) – 10%	(t-1)

5.3.2 Special Case (Pension Credit)

Size of s	Error A min : max	Error C min : max	Error E min : max
All s	s/2+ cap_min : s + cap_max	s/2 : s	5% : s + 5%

Error B		
t	**Lower limit**	**Upper limit**
All t	p + BK + min_enr_hr	p + BK + max_enr_hr

Error D1		
t	**Lower limit (look up using t_max)**	**Upper limit (look up using t_max)**
<95%	Y * (1 – t_max)	(1 – t_min)
95% - 100%	0	(1 – t_min)
100% - 105%	0	5%
>105%	0	0

Error D2		
t	**Lower limit (look up using t_max)**	**Upper limit (look up using t_max)**
<95%	0	0
95% - 100%	0	5%
100% - 105%	0	t_max – 1
105% - 110%	0	t_max – 1
>110%	t_min - 1	t_max – 1 + 2.5%

5.4 Special assumptions about benefit confusion

For some benefits and family types there is evidence to suggest that respondents are getting the benefit being assessed confused with another type of benefit. In this situation we allow for possible benefit confusion when calculating the errors. We never consider this for Pension Credit as we use data-linking to identify any recipients of PC who have not reported receipt.

Agreed factors in deciding whether to allow for possible benefit confusion are:

- Whether benefit confusion assumptions were used in the previous year – if this is the case they should always continue to be used to maintain the time series.
- If they were not used in the previous year then only consider using if t < 0.9 (as this suggests significant under-reporting) and then several other factors should be used to decide:
    - The possible proportion of benefit confusion (if less than 2 percentage points considered as insignificant so do not use benefit confusion assumptions).
    - By how much the t-statistic is below 90%.
    - Increase in the final upper take-up range – if the final upper take-up range shows a large increase from the previous year then consider allowing for benefit confusion.

If it is decided that benefit confusion assumptions should be used then errors B and D1 need to be calculated differently to account for this. The new calculations are shown in Table 5.4.1.

5.4.1 Alternative error calculations when using benefit confusion

Size t	Error B min	Error B max
< 90%	p + (X * iro / 2)	p + (X * iro / 2)

Size t	Error D1 min	Error D1 max
< 90%	Y * (1-(t + Bmax))	1 – t – (X * iro / 2)

5.5 Special assumptions about asymmetry of non-sampling error

In the situation where it seems as if under-modelling occurs with very different frequency to over-modelling, we might change our assumptions about errors A and C.

The agreed factors for deciding to do this are:

- There is a 10 percentage point difference between the level of over/under-modelling (v minus u as obtained from the FRS).
- The s-statistic is larger than 10 per cent.
- These first two signs were also observed in the previous year. The first year we see the signs we do not take action and await confirmation next year.

The changes we make to the assumptions are to shrink or grow our assumptions about the over-modelling error (Error A) according to whether there is evidence that over-modelling occurs more or less frequently than under-modelling error (Error C). Table 5.5.1 shows the new calculations of errors A and C when it is decided that asymmetry is present.

5.5.1 Alternative error calculations when asymmetry is present

Case	Error A min	Error A max
Not Pension Credit	(S/3) * (v / u)	S * (v / u)
Pension Credit	(s/2 + cap_min) * (v / u)	(s/2 + cap_max) * (v / u)

There is a further complication in these circumstances whereby we might want to change whether it is Amin or Amax that is associated with maximum take-up when we come to the stage of calculating the ranges. This possibility is tested using two quantities:

5.5.2 Effect of over- and under-modelling on the ENR count

effectV	**= R * (1 - s) * (Amin + Amax) / 2** This is the extent to which we estimate over-modelling is adding to the **ENR** count. We multiply an average of error **A** by **R** to get an idea of how many falsely entitled cases we might have added in. Most of the cases we have wrongly modelled as entitled are likely not to be receiving. So we reduce the number very slightly by the (**1-s)** factor which is an estimate of the proportion of claims that are genuinely entitled to their claim.
effectU	**= R * (1 – tu + 0.1) (Cmin + Cmax) / 2** This is the extent to which we think under-modelling might be detracting from the ENR. We multiply an average value of error **C** by **R** to get an idea of how many more entitled cases there should be. But then not of all the entitled cases we have missed will be ENRs – some will take-up and some will not. Probably the take-up in this group is lower than usual because they are marginal cases. So we boost the assumed non-take-up rate by 10 percentage points.

Usually **effectV > effectU.** So if errors A and C are at their maximum then overall the ENR count is too high, so we need to reduce it, and in so doing boost the take-up rate.

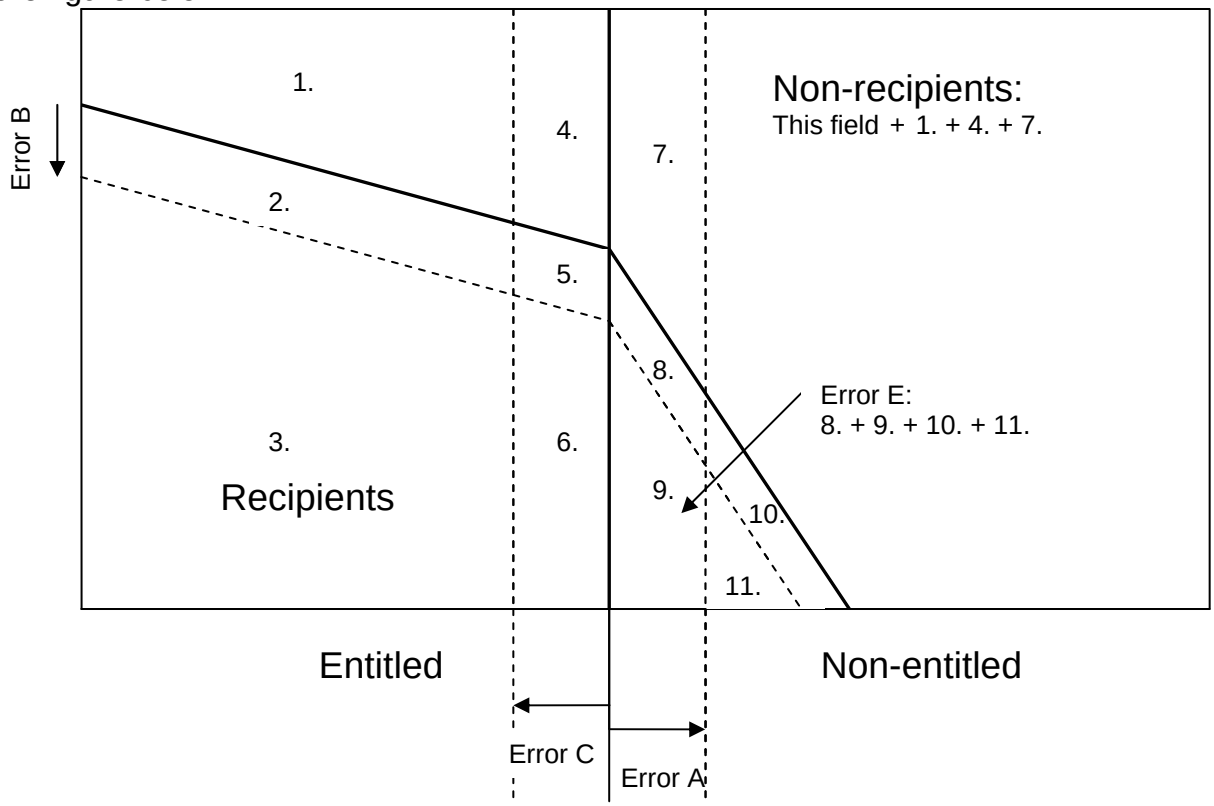
So, usually large values of errors A and C are associated with the maximum possibility for take-up, and vice versa.

In the case of asymmetry where we have shrunk down error A, it may turn out that effectV < effectU in which case we need to associate the maximum values of errors A and C with minimum take-up.

6 Calculation of take-up ranges

6.1 Error Framework

The non-sampling errors we have identified so far can be represented diagrammatically as in the figure below:



In the above figure areas outlined by full lines represent actual take-up, while dashed lines and arrows show the impact each error would have were it to occur.

Our objective is to determine the size of areas 1. and 4. having corrected for the non-sampling errors. This will be the estimate of the true number of entitled non-recipients.

In practice we achieve this by determining all eleven of the areas indicated. To do this we will need to translate the diagram and the assumptions we have made so far about errors A to E as a system of eleven simultaneous equations which can be solved. There will also need to be some further parameters assumed.

Notice that D1 and D2 are not part of the diagram. They feed into the process in a different way.

Aside from errors A to E there are two other parameters we can use. These are the number of ENRs we have initially modelled, and the numbers of recipients R we have observed. For convenience we scale these down such that the total R + ENR = 100.

The details of this scaling are:

- R_{100} = (100 * R) / (R + ENR)
- ENR_{100} = (100 * ENR) / (R + ENR)

This is where error D feeds in. We use it to adjust the ENR parameter for grossing issues:

- For under-grossing: adj(ENR) = ENR_{100} * (1 / (1 - D1));
- For over-grossing: adj(ENR) = ENR_{100} * (1 / (1 + D2)).

The first equation we can derive from our diagram then is that the number of ENRs we have initially modelled will comprise of areas 1. , 2. , 7. and 8.

I. 1. + 2. + 7. + 8. = adj(ENR)

The second thing we can derive is that all the areas except 1., 4. and 7. will be observed as part of the recipient count.

II. 2. + 3. + 5. + 6. + 8. + 9. + 10. + 11. = R_{100}

6.2 Additional Assumptions

In order to be able to write down enough equations to describe the diagram, we are going to need to make more assumptions in addition to those we have already made about errors A, B, C, D and E. Table 6.2.1 below shows the four extra parameters (a, b, c and d) we make assumptions about, parameters c and d are always assumed to be 0.2 and the assumed values of c and d are shown in Table 6.2.2.

6.2.1 Extra parameters needed – a, b, c and d.

a	The assumed take-up rate of those affected by error C, the understatement of entitlement i.e. those who are truly entitled but we mistakenly classify them as non-entitled. **See table 7.2.1 below for which assumption to make**
b	The assumed take-up rate of those affected by error A, the overstatement of entitlement i.e. those who are truly not entitled but we mistakenly classify them as entitled. **See table 7.2.1 below for which assumption to make**
c	The assumed rate of benefit under-reporting among those who are modelled as not entitled but are truly entitled. **Always assume this is 0.2**
d	The assumed rate of benefit under-reporting among those who are modelled as not entitled and are not truly entitled. **Always assume this is 0.2**

6.2.2 Assumed values of parameters a and b

Size of take-up (tu)	**(1 - a)**		**(1 - b)**	
	Max take-up	***Minimum take-up***		**Max take-up**
	Low	***High***	***Low***	**High**
90% < tu < 100%	3 * (100-tu)	5 * (100-tu)	60%	95%
80% < tu < 90%	30%	50%	60%	95%
70% < tu < 80%	40%	60%	70%	95%
60% < tu < 70%	50%	70%	80%	95%
50% < tu < 60%	60%	80%	90%	95%
tu =< 50%	70%	90%	95%	100%

6.3 Equations to be solved

Using these assumptions, we can now describe our diagram using the equations below (the first two are restated from above):

I. 2. + 3. + 5. + 6. + 8. + 9. + 10. + 11. = R

II. 1. + 2. + 7. + 8. = adj(ENR)

III. B * (2. + 3.) = 2.

IV. (1. + 2. + 3. + 4. + 5. + 6.) * C - 4. - 6. = 5.

V. ((2. + 3. + 5. + 6.) * E - 8. - 9.) * d = 10.

VI. (1. + 2. + 3. + 4. + 5. + 6.) * C * (1 - a) = 4.

VII. ((1. + 2. + 3. + 4. + 5. + 6.) * C - 4.) * (1 - B) = 6.

VIII. ((2. + 3. + 5. + 6.) * E - 8. - 9.) * (1 - d) = 11.

IX. (1. + 2. + 3. + 4. + 5. + 6.) * A * (1 - b) = 7.

X. (1. + 2. + 3. + 4. + 5. + 6.) * A * b * c = 8.

XI. (1. + 2. + 3. + 4. + 5. + 6.) * A * b * (1 - c) = 9.

6.4 Minimum and Maximum Scenarios

Recall that for each of the parameters A, B, C, D, E, a, and b, we have made maximum and minimum assumptions about their magnitude. These give rise to maximum and minimum take-up rates in accordance with Table 6.4.1 below.

6.4.1 Error combinations that yield the maximum and minimum limits for true take-up

Error	For minimum true take-up	For maximum true take-up
A	Low	High
B	Low	High
C	Low	High
D1	High	Low
D2	Low	High
E	High	Low
(1-a)	High*	Low*
(1-b)	Low*	High*
c	20%**	20%***
d	20%**	20%***

* See Table 6.2.2 above for high/low values of a and b

** Set to zero if minimum level of Error B is zero

*** Set to zero if maximum level of Error B is zero

6.5 The need for iterations of equation solving

In general there is no need to iterate this process, in the sense of solving the system of equations I to XI, adjusting the parameters A, B, C, D1, D2, E, a, b, c and d.

However there are specific situations where this may be necessary. After solving the system we will have estimates of areas 1. and 4. on the diagram, which can be used to calculate another estimate of take-up. In particular the new estimate would be: R / (R + 1 + 4). Remember that here all the quantities have been scaled down to the scale from 1 to 100.

Now it may happen that this new estimate of take-up is not consistent with our assumptions about **a** and **b** as determined by the table in Section 6.2. The new estimate may have jumped into a different 10 percentage point band.

In that circumstance it is necessary to adjust our assumption about **a** and **b** and to resolve the system accordingly.

In the case where take-up exceeds 90 per cent, the assumptions about **a** and **b** depend in some detail on the take-up estimate. In that instance, it is necessary to iterate repeatedly until the estimate of take-up and the assumptions about **a** and **b** are consistent with the table, allowing a tolerance of one hundredth of one percent.

6.6 Result of calculating take-up ranges

The result of these calculations is to produce for each split two point estimates of caseload take-up which reflect the two extremes of the assumptions we are willing to make about non-sampling error.

These can be translated into two extreme assumptions about the number of entitled non-recipients, according to Table 6.6.1 below.

6.6.1 Calculation of take-up ranges

tu_min	**= R_{100} / (R_{100} + 1. + 4.)** The min and max versions are obtained by solving our framework system using the different bundles of assumptions about the system parameters.
tu_max	
enr_max	**= (R / tu_min) - R**
enr_min	**= (R / tu_max) - R**

Subsequently these quantities can be aggregated to form totals and subtotals, adjusted for sampling error, and used as a basis for calculating expenditure take-up.

6.7 Derivations of totals and subtotals

In general the subtotals and totals for enr_max and enr_min can be obtained by simple addition. However, for Housing Benefit and Council Tax benefit all totals are pro-rated to the tenure type total in order to ensure that the totals and subtotals obtained through different methods are internally consistent. This is not necessary for other benefits as they are only split by family type.

The procedure for pro-rating Housing Benefit and Council Tax Benefit is shown in Table 6.7.1 below; the quantities to take forward are in the highlighted cells. Note:

- In the case of HB there are no owner occupiers who are entitled but not receiving and so C = 0.
- In the case of Council Tax Benefit, splits by employment type are not provided, so round 2 of the pro-rating is not required.
- This table is re-using letters and symbols that are used elsewhere in the document. The table should however be considered in isolation rather than trying to read across to other parts of the calculation.

6.7.1 Pro-rating procedure for Housing Benefit and Council Tax Benefit

Group	Initial estimate of enr_min or enr_max	Round 1 of pro rating	Round 2 of pro rating (applies to HB only)
Social rented	A		
Private rented	B		
Owner Occupier **(CTB only)**	C		
Tenure Total	**T1 = A + B + C**		
Pensioner	E	J = E * (T1/T2)	
Couples with Children	F	K = F * (T1/T2)	
Singles with children	G	L = G * (T1/T2)	
Non-pensioner without children	H	M = H * (T1/T2)	
Non-pensioner family total		**T3 = K + L +M**	
Family Total	**T2 = E + F + G + H**	**T4 = J + K + L + M = T1**	
Working	N		Q = N * (T3/T5)
Not-working	P		R = P * (T3/T5)
Non-pensioner employment total	**T5 = N + P**		**T6 = Q + R = T3**

7 Sampling errors

7.1 Caseload take-up

The sampling errors for the estimated numbers of entitled non-recipients rely on the Normal approximation to the binomial distribution. Table 7.1.1 shows how one can move from the enr_max or enr_min estimates - which we arrived at from the error framework - to the sampling error adjusted estimates of entitled non-recipients and caseload take-up that can be published. The estimates in the highlighted cells are those that are published.

7.1.1 Calculation to adjust caseload take-up ranges for sampling error.

n	= (sample entitled cases)
x	= (sample ENR cases)
q	**= x / n** On an ungrossed, sample basis, the proportion of entitled cases who do not receive benefit
z	**= √(q * (1 - q) * (1 / n)) * (er + enr) * 1.96** This is a formula for the width of the 95 per cent confidence interval associated with the raw unadjusted grossed estimate of ENRs obtained from the survey data. We assume that **x**, the sample number of ENRs, follows a binomial distribution where the number of events is **n**, the number of cases modelled as entitled and where each case will receive their entitlement with a certain probability, estimated by our observed **x/n**. So the first part of the formula is an estimate for the standard deviation of our estimate **x/n** of the probability of success. We next think of the grossed number of entitled non-recipients as being this probability of success applied to the grossed number of entitled cases. So to obtain the standard deviation of this grossed number we multiply through by the **(er + enr)** term. Finally we are assuming that for the size of n that we have, the sampling distribution of the sample proportion **(q)** is approximately normal. For a normal distribution we know that 95 per cent of the distribution is within 1.96 standard deviations of the mean, hence the final term in calculating the width of the confidence interval.
CI%	**= z / ENR** This expresses the width of the confidence interval associated with the raw gross estimate of ENRs as proportion of that raw gross estimate.
ENR_max	**= enr_max * (1 + CI%)** This is the result of applying the same proportional adjustment we have made to the raw gross estimate of ENRs to non-sampling error adjusted estimate of ENRs.

ENR_min	**= enr_min * (1 - CI%)**
TU_min	**= R / (R + ENR_max)**
TU_max	**= R / (R + ENR_min)** This is the minimum and maximum publication caseload take-up rate, fully adjusted for sampling and non-sampling errors

7.2 Expenditure take-up

Expenditure estimates of take-up are very easily derived from caseload estimates of take-up and a few key variables in terms of claimed and unclaimed amounts. The estimates in the highlighted cells are those that are to be published.

7.2.1 Calculation to adjust expenditure take-up ranges for sampling error.

£enr	As in Table 3.1.1 above that this is the average amount unclaimed according to the survey data, weighted using the survey grossing factors.
£R	As in Chapter 2 that this is the average amount claimed, derived from administrative data. The derivations are slightly different for each benefit.
sd(£enr)	As in Table 3.1.1 above that this is the observed standard deviation of the unclaimed amounts in the sample.
se(£enr)	= **sd(£enr) / √ x** This is the standard error associated with the average unclaimed amount, and as in Table 3.1.1 above **x** refers to the sample number of entitled non-recipients.
£UC_min	**= 52 * ENR_min * (£enr – 1.96 * se(£enr))** This is a minimum estimate of the amount of unclaimed benefit in the year. This takes account of sampling and non-sampling uncertainty in the number of ENRs, and of sampling error in the average amount. The assumption is made that the sampling distribution of the average unclaimed amount is approximately normal.
£UC_max	**= 52 * ENR_max * (£enr + 1.96 * se(£enr))** This is a maximum estimate of the amount of unclaimed benefit in the year.
£C	**= 52 * R * £R** This is the total amount claimed over the year, based on administrative data.
TUX_min	**= £C / (£C + £UC_min)**
TUX_max	**= £C / (£C + £UC_max)** This is the minimum and maximum publication expenditure take-up rate, fully adjusted for sampling and non-sampling errors

8 Comparisons over time

It is often very difficult to assess whether take-up rates have changed over time, because the ranges will often overlap, and we have no concept of a central estimate. This chapter describes ways of assessing the change over time and caveats to consider.

8.1 Algorithm for deciding whether take-up has changed

In order to compare change over time where ranges overlap we calculate several alternative point estimates for the two years we are comparing, the aim is to strip out the effect of different levels of modelling error in the different years. Table 8.1.1 below shows how each of the alternative point estimates is calculated. Pension Credit point estimates need to be calculated differently to other benefits as shown in Table 8.1.2. The highlighted rows are the estimates that are used to determine change over time. Refer to Table 3.1.1 for the definitions of **enr**, **w**, **u** and **v** and Table 5.2.1 for the definition of **p**, these are all obtained from the survey data.

8.1.1 Definition of alternative point estimates adj1 – adj5: General Case (Not Pension Credit)

enr1	**= enr – (R * p) = enr – w** This is the raw enr count adjusted for those cases awaiting the outcome of a claim that appear to be entitled (so we expect the claim to be successful). If we take these cases out of the enr count before we compare two years, then we can strip out the effect of different operational conditions affecting the comparison (e.g. if we did not then in a slow processing year we would have an artificially high number of enr cases)
Adj1	**= R / (R + enr1)**
enr2	**= enr * (1 – (v - u))** If we have evidence of over-modelling, and **v** exceeds **u**, then this adjustment will have the effect of the enr count and vice versa if **u** exceeds **v**. This is one way to try and strip out of the effect of different levels of modelling error in different years.
Adj2	**= R / (R + enr2)**
enr3	**= enr * (1 – (v - u)) – w** **= enr * (1 – (v - u)) – R * p** **= enr1 + enr2 – enr** These are three different ways of expressing an estimate that is adjusted for both modelling error and pipeline cases, with the aim being the ability to assess the underlying change in take-up. The raw count of **enr** is adjusted by a factor for modelling error, and then a subtraction is made to account for pipeline cases.
Adj3	**= R / (R + enr3)**

enr4	**= enr / t** This is an attempt to strip out any under- or over-grossing effect. If there is under-grossing of benefit recipients then we expect **t** to be less than one (ignoring misreporting for the moment). In that case we might also assume there is under-grossing among the enr population, and the adjustment specified (division by a factor less than one) will serve to correct the enr count by inflating it.
Adj4	**= R / (R + enr4)**
enr5	**= (enr / t) – w** **= (enr / t) – R * p** **= enr1 + enr4 – enr** This is a version adjusted for both grossing and pipeline cases, to try and allow us to understand the underlying changes.
Adj5	**= R / (R + enr5)**

8.1.2 Definition of alternative point estimates adj2 – adj6: Special Case (Pension Credit)

enr2	**enr2 = enr – (R * p) – (R * min_enr_hr)** **enr3 = enr – (R * p) – (R * max_enr_hr)** This is the raw estimate adjusted both for pipeline cases (as above, cases awaiting the outcome of a claim who appear to be entitled) and for hidden recipients that have been wrongly included in the ENR count. Note that we are multiplying by **R** in both the subtracted terms, because we have initially defined these quantities as proportions of **R**, to use them in the framework for non-sampling errors.
enr3	
Adj2	**= R / (R + enr2)**
Adj3	**= R / (R + enr3)**
enr4	**= enr * (1 – (v - u))** If we have evidence of over-modelling, and **v** exceeds **u**, then this adjustment will have the effect of the enr count and vice versa if **u** exceeds **v**. This is one way to try and strip out of the effect of different levels of modelling error in different years.
Adj4	**= R / (R + enr4)**
enr5	**enr5 = enr * (1 – (v - u)) – (R * p) – (R * min_enr_hr)** **enr6 = enr * (1 – (v - u)) – (R * p) – (R * max_enr_hr)** This pair of estimates attempt to strip out differences in modelling error, in the level of pipeline cases and in the extent of hidden receipt.
enr6	
Adj5	**= R / (R + enr5)**
Adj6	**= R / (R + enr6)**

The following steps are the basic method for determining whether a change in take-up has occurred, or not, from one year to the next or from the beginning of the time-series.

- Do the ranges overlap? If not, then the minimum change that could have happened is used. For instance, if the range in this year is (88 : 90), and the range last year was (94 : 96), then the minimum change is 94 – 90 = 4 percentage points.
- **General Case (not Pension Credit):** If the ranges do not overlap, then the minimum of **adj3** and **adj5** is used, but only if **adj3**, **adj5** and the point estimate all go in the same direction (see Table 8.1.1 for derivation of **adj3** and **adj5**).
- **Special Case (Pension Credit):** If the ranges do not overlap, then the minimum of adj5 and adj6 is used, but only if **adj5**, **adj6** and the point estimate all go in the same direction (see Table 8.1.2 for derivation of **adj5** and **adj6**).
- If the minimum is less than 0.5 percentage points, or the estimates **adj3** and **adj5** (or **adj5** and **adj6** in the special case) go in different directions, then there is no evidence of change.

9 Caveats

9.1 Change in bias caveats

When stating the change from one year to the next or from the beginning of the time-series it needs to be considered if during that period the results have been affected by large changes in the level of bias present in the modelling. If this is the case then we cannot be certain of the change found due to the changes in bias; a caveat is added to this effect.

In order to decide whether to include a caveat to say that the conclusion is affected by a change in bias the following criteria are used:

- Whether there is a change in the use of the benefit confusion assumptions (see section 5.4).
- Whether there is a change in the use of the asymmetrical error assumptions (see section 5.5).
- Changes in the t-statistic:
    - if the t-statistic changes from above 100 per cent to below 90 per cent or opposite (especially if either of the conditions in the first two bullets are met).
    - if change in the t-statistic is large; above 20 percentage points.
- Changes in the s-statistic:
    - If the s-statistic moves above 10 percentage points (especially if asymmetrical error assumptions used).
    - Change in the s-statistic means a change of around 5 percentage points.
- Difference between the proportion of over- and under-modelled cases:
    - Substantial difference means a change from there being 10 per cent or more over-modelled cases than under-modelled cases one year to there being 10 per cent or more under-modelled cases than over-modelled cases the next year, or vice versa.

Which results to apply the change in bias caveat to is slightly subjective but as a rule we only include change in bias caveats for the worst affected groups.

9.2 High Bias Caveats

If there has been only a small change between the years in the biases, but one statistic is very low or very high for certain groups (mainly the t-statistic), instead of changing bias, consistently high bias is highlighted in a caveat.

Which results to apply the high bias caveat to is slightly subjective but as a rule we only include high bias caveats for the worst affected groups.

9.3 Expenditure Caveats

Where there is a large difference in the between the amounts of modelled entitlement and amounts claimed for those in receipt for a particular group we need to caveat the expenditure based results for that group.

In order to decide which groups need caveats we look at the difference between over- and under-modelling for the year in question. Comparing this difference across benefits and splits shows which groups seem to have particularly high differences. Caveats are applied to those where either over-modelling is much larger than under modelling or where under-modelling is much larger than over-modelling. As a rule we only apply expenditure caveats to the worst affected groups.