

	<h1>Methodology Statement for the Department for Work and Pensions Statistical Summary</h1>
	<p>Author: Chris Douglass</p>

June 2011

DWP Department for
Work and Pensions

Contents

- Introduction..... 4
- WPLS Tabulation Tool..... 6
 - Introduction..... 6
 - Production Stage 1 6
 - Quality Assurance Stage 1 7
 - Production Stage 2..... 7
 - Quality Assurance Stage 2 9
 - Pre-Publication 10
 - Publication 10
- Flows Tabulation Tool 12
 - Introduction..... 12
 - Production Stage 1 12
 - Quality Assurance Stage 1 13
 - Production Stage 2..... 13
 - Quality Assurance Stage 2 14
 - Pre-Publication 15
 - Publication 15
- 5% Sample Tabulation Tool..... 16
 - Introduction..... 16
 - Production Stage 1 17
 - Quality Assurance Stage 1 17
 - Production Stage 2..... 17
 - Quality Assurance Stage 2 18
 - Pre-Publication 19
 - Publication 19
- Neighbourhood Statistics..... 21
 - Introduction..... 21
 - Production 21
 - Quality Assurance 22
 - Pre-Publication 22
 - Publication 23
- New Deal Tabulation Tools..... 24
 - Introduction..... 24
 - Production Stage 1 24
 - Quality Assurance Stage 1 25
 - Production Stage 2..... 25
 - Quality Assurance Stage 2 26

Methodology Statement for the Department for Work and Pensions Statistical Summary

- Pre-Publication 27
- Publication 28

- Sanctions Tabulation Tool 29
 - Introduction..... 29
 - Production Stage 1 29
 - Quality Assurance 29
 - Production Stage 2 30
 - Quality Assurance Stage 2 30
 - Pre-Publication 31
 - Publication 32

- Housing Benefit / Council Tax Benefit Statistics 33
 - Introduction..... 33
 - Production 33
 - Quality Assurance 33
 - Publication 34

- The DWP Statistical Summary 35
 - Introduction..... 35
 - Production 35
 - Quality Assurance 36
 - Pre-Publication 36
 - Publication 36

Introduction

The DWP Statistical Summary brings together key National Statistics on DWP administered benefits, some employment programmes and JSA (Jobseeker's Allowance) sanctions and vacancies. To provide a more complete picture of DWP responsibility, statistics on Housing Benefit and Council Tax Benefit (administered by Local Authorities) and the CSA are also included.

The Summary is published monthly, containing Housing Benefit and Council Tax Benefit National Statistics together with official statistics giving early estimates of numbers of inactive benefit claimants, plus a link to the latest report on Flexible New Deal. Each quarter (in May, August, November and February), a larger summary also contains the detail of DWP administered benefits, some employment programmes, JSA sanctions, vacancies and CSA.

Alongside the Statistical Summary, there are a large number of tables which enable the user to delve in to the detail. These are provided to the user as:

- The DWP Tabulation Tool - for DWP administered benefits and Employment Programmes the Tabulation Tool provides the user with an interactive tool to select one of thousands of possible tabulations. A similar Tabulation Tool, derived from 5% sample data, is also available. Whenever possible, 100% data should be used in preference to 5% estimates, as these are both more accurate and form DWP's headline statistics. Some statistics remain outside the scope of the Tabulation Tool but will continue to be available via the internet as pre-defined summary tables.
- To widen accessibility and understanding of context, DWP also publish statistics via Nomis (<https://www.nomisweb.co.uk/Default.asp>). This provides easy access for users of other labour market statistics.
- Separate detailed tables on Housing Benefit and Council Tax Benefit and Claimants on out of work benefits (as a "one-click" on the Working Age Client Group Tabulation Tool);
- Separate official statistics regarding Flexible New Deal;
- Tables broken down by Lower Super Output Area (LSOA)
- Other tables and background information via links on the Tabulation Tool pages (e.g. links to long time series spreadsheets; descriptions of the benefit.)

Responsibility for the production of the DWP Statistical Summary and all associated products listed above (collectively known as the DWP Statistical Summary or the Summary) is under the control of the DWP Statistical Summary Virtual Team. This virtual team is effectively three teams working together towards the common goal of publication, and consists of several teams from within Information Directorate (IFD),

Methodology Statement for the Department for Work and Pensions Statistical Summary

which is part of the Department for Work and Pensions. The virtual team works to the lead Statistician, whose role it is to oversee the progress being made by the virtual team during the production of DWP Statistical Summary. During the process, the virtual team works together closely and has weekly meetings in order to review progress against the agreed timetables, communicate effectively, and discuss any issues that have arisen during the process. The virtual team works to an agreed timetable each quarter (one team takes sole responsibility for the production of the Monthly Statistical Summary). The timetable is useful to see all tasks which are completed each quarter, at which points in the process, and by which part of the virtual team.

The main purpose of this document is to provide users with information about all the methods used to produce the DWP Statistical Summary and all major associated products in accordance with practices set out in the Code of Practice for Official Statistics.

WPLS Tabulation Tool

Introduction

The Work and Pensions Longitudinal Study (WPLS, http://statistics.dwp.gov.uk/asd/longitudinal_study/index.php?page=ic_longitudinal_study) Tabulation Tool for DWP administered benefits (<http://83.244.183.180/100pc/tabtool.html>) provides the user with an interactive tool to select one of thousands of possible tabulations. For a list of breakdowns available on the WPLS Tabulation tool see the Guide to Sources document (<http://research.dwp.gov.uk/asd/asd1/tabtools/guidance.pdf>). There are two main sources of the number of benefit claimants; the WPLS (100%) Tabulation Tool and the 5% Sample Tabulation Tool (see the [5% Sample Tabulation Tool](#) section). The best statistics to use are sourced from WPLS as these data are based on 100% of claimants and cover information such as age and gender of claimant, duration of their spell on benefit and geographical locations of claimants. The WPLS Tabulation Tool is sourced from WPLS and the following benefit claimant numbers are recorded on the WPLS Tabulation Tool: Attendance Allowance (AA), Bereavement Benefit (BB), Carer's Allowance (CA), Disability Living Allowance (DLA), Employment and Support Allowance (ESA), Incapacity Benefit (IB), Incapacity Benefit/Severe Disability Allowance (IB/SDA), Income Support (IS), Jobseeker's Allowance (JSA), State Pension (SP), Pension Credit (PC), and Widows Benefit (WB).

The WPLS Tabulation Tool is released quarterly (in May, August, November and February) and is produced and released under the control of the DWP Statistical Summary Virtual Team. A substantial part of the DWP Quarterly Statistical Summary relates to the numbers claiming DWP benefits, which are sourced from the WPLS Tabulation Tool.

Production Stage 1

Production begins with the creation of the Frozen Datasets. The Frozen Datasets are a snapshot of the benefit system at a certain point in time, the last day of the quarter. They are ran using a program called SAS (<http://www.sas.com>) and are derived from the National Benefits Database (NBD) and supplemented by point-in-time data from GMSONE. NBD holds claim level information on what benefits a person has claimed past and present, whilst GMSONE holds claim level information on what benefits a person is on at that point in time and contains key characteristics about each benefit claim at that point in time. The data for GMSONE comes from the different benefit systems. It is collected by DWP for administrative purposes, and covers the following benefits - Jobseekers Allowance (JSA), Income Support (IS), Incapacity Benefit (IB), Severe Disability Allowance (SDA), Employment and Support Allowance (ESA), Carer's Allowance (CA), Widows Benefit/Bereavement Benefit (WB/BB), Disability

Methodology Statement for the Department for Work and Pensions Statistical Summary

Living Allowance (DLA), Pension Credit (PC), State Pension (SP) and Attendance Allowance (AA).

The SAS program used to create the datasets is run as soon as the latest NBD and Address Hierarchy are available, where the Address Hierarchy contains an address history for all customers who have come into contact with the Department. The Frozen Datasets contain all customers on benefit at that point in time across all benefits and contain key characteristics about each benefit claim at that point in time. Additionally there is a “Pers” frozen dataset which represents the personal details of each claimant. This contains information such as age, gender, location. The key feature of the Pers dataset is the identification of key groups of people within the benefit system. These are known as client group (e.g. working age, pension age) and statistical group (e.g. jobseeker, the main reason why they are in the benefit system).

Quality Assurance Stage 1

Quality assurance of the Frozen Datasets is an automated process carried out in SAS with manual analysis of output. The automated program aims to check for various features within the datasets compared to the previous version, such as missing variables, new decodes, missing values, and changes to the distribution of variables. The results are stored away in a lookup file for use in future quarters and are presented in the form of an html report. Members of the team assess these reports and look for any potential issues within the datasets. All potential issues and suspect movements in the datasets are raised and investigated. It is often the case that movements can be attributed to change in benefit policy and are fully expected. Once completed, the lead Statistician will ensure all checking has been done and signs off the Frozen Datasets. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Production Stage 2

Once the Frozen Dataset Quality assurance is completed and fully signed off, a process known as “Cube creation” begins. This part of the process is also carried out in SAS and just re-formats the Frozen Datasets into a format compatible with the next step, which is the Tabulation Tool creation process. The cube creation also pre-aggregates each dataset into every unique combination of variables within the dataset. This speeds up the Tabulation Tool production later. The cubes also set-up most of the desired categories for the Tabulation Tool, for example age variables are formatted into age bands and Local Authority variables are numbered to be put into the desired order. There is one Cube for each benefit published on the Tabulation Tool. Each Cube contains data from the quarter in question and no amendments to a previous quarter are made (unless an error has been flagged). The back-series remains the same. There are also Time-series Cubes, which due to no amendments

Methodology Statement for the Department for Work and Pensions Statistical Summary

to the back-series are just a concatenation of all single-date cubes from each quarter, allowing time-series Tabulation Tool pages to be produced.

Before the Tabulation Tool can be run, we update the Modsheets. These are a set of spreadsheets, which contains every single parameter required for the run. All formats, titles, labels, links, and footnotes are defined here as well as which parts of the Tabulation Tool are required from the run. Changes to the Modsheets are often required each quarter. Changes in benefit policy can often require additional footnotes each quarter, or Parliamentary Constituencies may change and require a new format. Once the Modsheets are completed, the Tabulation Tool production can begin. The Tabulation Tool is essentially a series of html pages which contain different cross-tabulations of data alongside appropriate titles, labels, and footnotes for that data. JavaScript code exists which create the “front-end” parts of the Tool. These are the interactive menus which sit on top of the html pages. The user makes selections on these pages and navigates to the appropriate table of data according to the selection made.

The data published on the WPLS Tabulation Tool is also published on the Nomis website (<https://www.nomisweb.co.uk/Default.asp>). The Nomis website has different functionality to the WPLS Tabulation Tool in that different tabulations can be produced, albeit slightly slower than the WPLS Tabulation Tool as tables are created on-the-fly when the user selects them. Not every scheme available on the Tabulation Tool is also available on Nomis, but most are. The Nomis publication process is relatively simple. Once the Cubes have been created above, a SAS program is run to convert each Cube into a text file with particular variable formats. These text files are then transferred to Nomis and uploaded to their pre-release area ready for Quality Assurance and publication.

Several ‘One-Click’ tables are also produced as part of the WPLS Tabulation Tool run. The Key Out Of Work Benefits table is an addition to the Working Age Client Group Tabulation Tool. This table presents annual and full-time series for the four main statistical groups defined to be in the “key out of work benefits” group. Three of the statistical groups are an exact match to the figures published within the working age statistical groups on the Tabulation Tool; however, the Jobseekers aspect is based upon Office for National Statistics (ONS) figures. There are two numbers for Jobseeker’s Allowance claimants given, UK Seasonally Adjusted and GB Non-Seasonally Adjusted. Both sets are simply taken from Nomis. There is no total presented for all groups combined. The production of this table is relatively easy and the table is prepared for publication by simply updating the spreadsheet. As part of the ONS Labour Market release on DWP Statistical Summary release day, they also publish the same figures. This is Table 25 of the ONS Labour Market release. This table is prepared in line with the key out work benefits table and sent to the ONS prior to DWP Statistical Summary release.

In addition, the ‘One-Click’ IS Lone Parents by Ethnicity table is added to the Tabulation Tool for Income Support (IS) as there is no ethnicity breakdown on this Tabulation Tool. This table contains lone parents on Income Support split by ethnicity and a host of other variables. It is run using an individual SAS program.

Methodology Statement for the Department for Work and Pensions Statistical Summary

Overall production of the WPLS Tabulation Tool takes around 6-8 weeks but is also subject to waiting times for retrospection. When data is subject to retrospection, it is collected not only at the reference date but also during further extracts which occur during a following set amount of time, permitting late information to be added to the database (with respect to the reference date) retrospectively. The 100% claimant total therefore represents the number of live claimants at a date more accurately, and is especially useful for benefits such as Incapacity Benefit and Disability Living Allowance where claims are often complicated and may take a long time to process. Retrospection was discussed at some length in the 2004 consultation documents and users were content with the quality/timeliness of publication trade off suggested (http://research.dwp.gov.uk/asd/asd1/stats_consultation/modstats.pdf).

Combined with retrospection this means that the latest data relates to the quarter ending 6 months previously. For example:

May 18 th 2011 Release	– Data relates to Quarter ending November 2010
August 17 th 2011 Release	– Data relates to Quarter ending February 2011
November 16 th 2011 Release	– Data relates to Quarter ending May 2011
February 15 th 2012 Release	– Data relates to Quarter ending August 2011

Quality Assurance Stage 2

Once the Tabulation Tool has successfully been created, the team are required to Quality assure the output. There are two questions to answer during this stage: Are the html pages correct and display all the correct data as expected? Does the tool have all the correct functionality? The first question is addressed by use of an automated SAS checking program which directs attention to those pages showing the most change. The idea behind this program is to compare the html page from the new quarter with the identical html page from the previous quarter. The program checks for things such as missing rows and columns from the table, missing footnotes, new footnotes and changed labels. It also assesses the level of change within the figures of the tables and flags those tables with the largest amount of change. All results are combined in an html report which is then checked by a member of the team. Anything listed in the html report is investigated to see whether it is expected. The second question requires members of the team to use and check the Tabulation Tool to ensure that all tables/breakdowns/footnotes are available and as expected. At any point during the Quality Assurance process if an error is spotted an incident will be raised and will be thoroughly investigated. When any SAS program is ran a log file is stored for future reference and to see where any part of the process may have broken down. If required any part of the process can be re-run and all checking will re-occur. 'One-Click' tables are individually assured and comparison to the information on the WPLS Tabulation Tool can be made.

Once completed, the lead Statistician will ensure everything has been checked and passed, and if happy with everything, will sign off the WPLS Tabulation Tool. All

actions are recorded on a release timetable, and are checked off, once complete, to ensure that no steps are missed.

Pre-Publication

Once the WPLS Tabulation Tool has been signed off it is then sent to the external web host and Nomis in preparation for publication. The external web host currently hosts the various Tabulation Tools, including the WPLS Tabulation Tool. All data from previous quarters resides on the website and is already live on the internet. When new data is created each quarter it must be sent to the external web host to be uploaded onto the site and made live alongside the existing data.

When the data has been received by the external web host and uploaded for publication it is the responsibility of the team to ensure that the correct data exists on the hosting site as was created in the first place. Further quality assurance must therefore be carried out at this stage. The external web host organises a secured pre-release window to the hosting site for a few days in the week prior to release day and notifies the team of how to access this. A member of the team will then run an automated SAS checking program, which compares the WPLS Tabulation Tool on our system to that on the external web site. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the external website and the pre-release access window is closed.

The WPLS data is also sent to Nomis (<https://www.nomisweb.co.uk/Default.asp>) in preparation for publication. When the data has been received by Nomis and uploaded for publication it is the responsibility of the team to ensure that the correct data exists on the website. Further quality assurance must therefore be carried out at this stage. Nomis organises a secure pre-release window prior to release day and notifies the team of how to access this. A member of the team will then carry out some manual checks. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the Nomis website.

Publication

The WPLS Tabulation Tool is added to the website at 09:30am on Release Day. The data available on the WPLS Tabulation Tool is also added to the Nomis website at 09:30am on Release Day. Users can access the external web site via the DWP Internet pages. The WPLS Tabulation Tool is accessible via the Create Your Own Statistics landing page which itself is available via the DWP Statistical Summary Landing Page.

The DWP Statistical Summary landing page can be accessed here:
http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries

The Create Your Own Statistics landing page can be accessed here:

Methodology Statement for the Department for Work and Pensions Statistical Summary

<http://research.dwp.gov.uk/asd/index.php?page=tabtool>

Additionally, the data published on the WPLS Tabulation Tool is available via the Nomis website:

<https://www.nomisweb.co.uk/Default.asp>

Flows Tabulation Tool

Introduction

The Flows Tabulation Tool for DWP administered benefits is the collective name of the following Tabulation Tools:

On Flows Tabulation Tool: http://83.244.183.180/flows/flows_on/tabtool.html

Off Flows Tabulation Tool: http://83.244.183.180/flows/flows_off/tabtool.html

The Flows Tabulation Tool provides the user with an interactive tool to select one of thousands of possible tabulations. The Flows Tabulation is the sole source of benefit flows numbers and is from the same source benefit data as used for the WPLS Tabulation Tool. Flows data is very useful as it can be used to closely track changes in benefit levels and the nature of benefit claimant cohorts. The following benefit flows are recorded on the Flows Tabulation Tool: Income Support (IS), Incapacity Benefit/Severe Disability Allowance (IB/SDA), Widow's Benefit/Bereavement Benefit (WB/BB), Employment and Support Allowance (ESA), Pension Credit (PC) and State Pension (SP).

The Flows Tabulation Tool is released quarterly (in May, August, November and February) and is produced and released under the control of the DWP Statistical Summary Virtual Team. A substantial part of the DWP Quarterly Statistical Summary relates to DWP Benefit Flows, which are sourced from the Flows Tabulation Tool.

Production Stage 1

Similar to for the WPLS Tabulation Tool, production begins with the creation of the Frozen Datasets. The Frozen Datasets created for the Flows process aim to capture all on-flows and off-flows occurring in the quarter. They are run using SAS and are derived from the National Benefits Database (NBD) and supplemented by point-in-time data from GMSONE. NBD holds claim level information on what benefits a person has claimed past and present, whilst GMSONE holds claim level information on what benefits a person is on at that point in time and contains key characteristics about each benefit claim at that point in time. The data for GMSONE comes from the different benefits systems and is collected by DWP for administrative purposes. It covers the following benefits - Jobseekers Allowance (JSA), Income Support (IS), Incapacity Benefit (IB), Severe Disability Allowance (SDA), Employment and Support Allowance (ESA), Carer's Allowance (CA), Widow's Benefit/Bereavement Benefit (WB/BB), Disability Living Allowance (DLA), Pension Credit (PC), State Pension (SP) and Attendance Allowance (AA).

The SAS program used to create the Flows datasets is run as soon as the latest NBD and Address Hierarchy are available, where the Address Hierarchy contains an

address history for all customers who have come into contact with the Department. The Frozen Datasets contain all customers who have flowed on or off benefit in the quarter.

Quality Assurance Stage 1

Quality assurance of the Frozen Datasets for Flows consists of several stages, beginning with an automated process carried out in SAS with manual analysis of output. The automated program aims to check for various features within the datasets compared to the previous version, such as missing variables, new decodes, missing values, and changes to the distribution of variables. The results are stored away in a lookup file for use in future quarters and are presented in the form of various HTML reports. Members of the team assess these reports and look for any potential issues within the datasets. All potential issues and suspect movements in the datasets are raised and investigated. It is often the case that movements can be attributed to change in benefit policy and are fully expected.

The outputs are also compared with the benefit claimant numbers obtained when running the WPLS Tabulation Tool to see that they 'balance' i.e. that adding on flows from the previous month and taking away off flows should equal the new claimant totals.

Once completed, the lead Statistician will ensure all checking has been done and signs off the Frozen Datasets. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Production Stage 2

Similar to for the WPLS Tabulation Tool, once the Frozen Dataset Quality assurance is completed and fully signed off a process known as Cube creation begins. This part of the process is also carried out in SAS and just re-formats the Frozen Datasets into a format compatible with the next step, which is the Tabulation Tool creation process. The cube creation also pre-aggregates each dataset into every unique combination of variables within the dataset. This speeds up the Tabulation Tool production later. The cubes also set-up most of the desired categories for the Tabulation Tool, for example, age variables are formatted into age bands and Local Authority variables are numbered to be put into a desired order. There is one Cube for each benefit published on the Tabulation Tool. Each Cube contains data from the quarter in question and no amendments to a previous quarter are made (unless an error has been flagged). The back-series remains the same. There are also a series of Time-series Cubes, which due to no amendments to the back-series, are just a concatenation of all single-date cubes from each quarter, allowing time-series Tabulation Tool pages to be produced.

Before the Tabulation Tool can be run we update the Modsheets. These are a set of spreadsheets, which contains every single parameter required for the run. All

Methodology Statement for the Department for Work and Pensions Statistical Summary

formats, titles, labels, links, and footnotes are defined here as well as which parts of the Tabulation Tool are required to be run. Changes to the Modsheets are often required each quarter. Changes in benefit policy can often require additional footnotes each quarter, or Parliamentary Constituencies may change and require a new format. Once the Modsheets changes are completed, the Tabulation Tool production can begin. The Tabulation Tool is essentially a series of html pages which contain different cross-tabulations of data alongside appropriate titles, labels, and footnotes for that data. JavaScript code exists which create the “front-end” parts of the Tabulation Tool. These are the interactive menus which sit on top of the html pages. The user makes selections on these pages and navigates to the appropriate table of data as per the selection made.

There is an extra table which is prepared and released alongside the Flows and WPLS Tabulation Tools for Pension Credit (PC) referred to as the ‘One-Click’ Pension Credit On Flows by Month Time series. This table relates to monthly on-flows to Pension Credit (PC) and is run using a single SAS program.

Quality Assurance Stage 2

Once the Tabulation Tool has successfully been created, the team are required to quality assure the output. There are two questions to answer during this stage: Are the html pages correct and display all the correct data as expected? Does the tool have all the correct functionality? The first question is addressed by use of an automated SAS checking program which directs attention to those pages showing the most change. The idea behind this program is to compare the html page from the new quarter with the identical html page from the previous quarter. The program checks for things such as missing rows and columns from the table, missing footnotes, new footnotes and changed labels. It also assesses the level of change within the figures of the tables and flags those tables with the largest amount of change. All results are combined in an html report which is then checked by a member of the team. Anything listed in the html report is investigated to see whether it is expected. The second question requires members of the team to use and check the Tabulation Tool to ensure that all tables/breakdowns/footnotes are available and as expected. The ‘One-Click’ Pension Credit On Flows by Month Time series table is extremely easy to assure as it is directly comparable to the quarterly Flows Tabulation Tool for Pension Credit (PC).

Once completed, the lead Statistician will ensure everything has been checked and passed, and if happy with everything, will sign off the Flows Tabulation Tool. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Pre-Publication

Once the Flows Tabulation Tool has been signed off it is then sent to the external web host in preparation for publication. The external web host currently hosts the various Tabulation Tools, including the Flows Tabulation Tool. All tables from previous quarters reside on the internet host and are already live on the internet. When new data is created each quarter it must be sent to the external host to be uploaded onto the site and made live alongside the existing data.

When the data has been received by the external web host and uploaded for publication, it is the responsibility of the team to ensure that the same data exists on the external host as was created in the first place. So, further quality assurance must be carried out at this stage. The external web host organises a secured pre-release window to the hosting site for a few days in the week prior to release day and notifies the team how to access this. A member of the team will then run an automated SAS checking program, which compares the Flows Tabulation Tool on our system to that on the external web site. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the website and the pre-release access window is closed.

Publication

The Flows Tabulation Tool is added to the external website at 09:30am on Release Day. Users can access the site via the DWP Internet pages. The Flows Tabulation Tool is accessible via the Create Your Own Statistics landing page which itself is available via the DWP Statistical Summary Landing Page.

The DWP Statistical Summary landing page can be accessed here:

http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries

The Create Your Own Statistics landing page can be accessed here:

<http://research.dwp.gov.uk/asd/index.php?page=tabtool>

5% Sample Tabulation Tool

Introduction

The 5% Sample Tabulation Tool for DWP administered benefits (<http://83.244.183.180/5pc/tabtool.html>) provides the user with an interactive tool to select one of thousands of possible tabulations. For a list of breakdowns available on the 5% Sample Tabulation tool see the Guide to Sources document (<http://research.dwp.gov.uk/asd/asd1/tabtools/guidance.pdf>). There are two main sources of benefit claimant numbers; the WPLS Tabulation Tool (see [100% Sample Tabulation Tool](#)) and the 5% Sample Tabulation Tool. The WPLS Tabulation Tool is sourced from WPLS. The best statistics to use are sourced from WPLS as these data are based on 100% of claimants and cover information such as age and gender of claimant, duration of their spell on benefit and geographical locations of claimants.

The reason the 5% Sample Tabulation Tool exists is both historical and coverage related. Before the WPLS (100%) data the 5% Sample data was all that existed in the Department. This data existed first and some series date back to 1995 on a consistent basis via the Tabulation Tool. Longer time series are available in earlier paper based publications from DWP. By the time the 100% extracts enabled the development of the WPLS data, many important time-series had been established, and therefore it was important to keep these series. In addition the retention of these series enabled important comparisons over time. Another reason for the 5% sample data is that this data contains many more variables than the WPLS counterpart. There is often interest in some of the extra variables available in the 5% sample data. When there is a direct comparison available between the 100% benefits data and 5% benefits data it is always recommended to use the 100% benefits data. DWP recommends that, where the detail is only available on the 5% sample data, the proportions derived should be applied to the overall 100% total for the benefit. The basic principle in combining the two, is that the number of claimants should be obtained from the WPLS data. To this, percentages taken from the 5% sample data are applied. The following benefits are recorded on the WPLS Tabulation Tool: Attendance Allowance (AA), Bereavement Benefit (BB), Disability Living Allowance (DLA), Incapacity Benefit (IB), Incapacity Benefit/Severe Disability Allowance (IB/SDA), Employment and Support Allowance (ESA), Income Support (IS), Jobseekers Allowance (JSA), Severe Disability Allowance (SDA), State Pension (SP), Pension Credit (PC), and Widows Benefit (WB).

The 5% Sample Tabulation Tool is released quarterly (in May, August, November and February) and is produced and released under the control of the DWP Statistical Summary Virtual Team.

Production Stage 1

Production begins when the team receive the 5% Samples, which are then loaded into a computer program called ASSIST. There is a separate extract for each benefit apart from Widows Benefit (WB), Bereavement Benefit (BB), and State Pension (SP) which come in on the same extract. The extracts for Attendance Allowance (AA), Disability Living Allowance (DLA), Incapacity Benefit (IB), Income Support (IS), Jobseekers Allowance (JSA) Pension Credit (PC) and Severe Disability Allowance (SDA) come in Quarterly whereas Widows Benefit (WB), Bereavement Benefit (BB), and State Pension (SP) come in six-monthly.

Once the extracts have been loaded into ASSIST they are processed using pre-defined rules, such as postcode matching in order to define geographical variables. The processed extracts are then turned into SAS views, which form the base data for the 5% Sample Tabulation Tool and enables them to be interrogated in the same way as a SAS dataset, enabling the use of SAS in the Cube creation later.

Quality Assurance Stage 1

Quality assurance of the SAS views is an automated process carried out in SAS with manual analysis of output. The automated program aims to check for various features within the datasets compared to the previous version, such as missing variables, new decodes, missing values, and changes to the distribution of variables. All potential issues and suspect movements in the datasets are raised and investigated. It is often the case where movements can be attributed to change in benefit policy and are fully expected. There is also a comparison of the 5% Sample benefit claimant totals with the equivalent WPLS numbers. This quality assurance is important as although the series are not exactly the same, the same trends should be apparent in both. A divergence in the trends here can be reflected in one or both of the series and/or processing of the series and would be investigated further.

Production Stage 2

Once the quality assurance of the SAS views is completed and fully signed off, a process known as Cube creation begins. This part of the process is also carried out in SAS and just re-formats the SAS views into a format compatible with the next step, which is the Tabulation Tool creation process. The cubes set-up most of the desired categories for the Tabulation Tool, for example age variables are formatted into age bands and Local Authority variables are numbered to be put into a desired order. There is one Cube for each benefit published on the Tabulation Tool. Each Cube contains data from the quarter in question and no amendments to a previous quarter are made (unless an error has been flagged). The back-series remains the same. There are also Time-series Cubes, which, due to no amendments to the back-series, are just a concatenation of all single-date cubes from each quarter. These allow time-series Tabulation Tool pages to be produced.

Before the Tabulation Tool can be run we update the Modsheets. These are a set of spreadsheets, which contains every single parameter required for the run. All formats, titles, labels, links, and footnotes are defined here as well as which parts of the Tabulation Tool are required from the run. Changes to the Modsheets are often required each quarter. Changes in benefit policy can often require additional footnotes each quarter, or Parliamentary Constituencies may change and require a new format. Once the Modsheets are completed, the Tabulation Tool production can begin. The Tabulation Tool is essentially a series of html pages which contain different cross-tabulations of data alongside appropriate titles, labels, and footnotes for that data. JavaScript code exists which create the “front-end” parts of the Tab Tool. These are the interactive menus which sit on top of the html pages. The user makes selections on these pages and navigates to the appropriate table of data according to the selection made.

The data published on the 5% Sample Tabulation Tool is also published on the Nomis website (<https://www.nomisweb.co.uk/Default.asp>). The Nomis website has different functionality to the 5% Sample Tabulation Tool in that different tabulations can be produced, albeit slightly slower than the 5% Sample Tabulation Tool as tables are created on-the-fly when the user selects them. Not every scheme available on the Tabulation Tool is also available on Nomis, but most are. The Nomis publication process is relatively simple. Once the SAS views have been created (see Production Stage 1 of [5% Sample Tabulation Tool](#)), a SAS program is run to convert each of the SAS views into a text file with particular variable formats. These text files are then transferred to Nomis and uploaded to their pre-release area ready for quality assurance and publication.

Quality Assurance Stage 2

Once the Tabulation Tool has successfully been created, the team are required to quality assure the output. There are two questions to answer during this stage: Are the html pages correct and display all the correct data as expected? Does the Tool have all the correct functionality? The first question is addressed by use of an automated SAS checking program which directs attention to those pages showing the most change. The idea behind this program is to compare the html page from the new quarter with the identical html page from the previous quarter. The program checks for things such as missing rows and columns from the table, missing footnotes, new footnotes and changed labels. It also assesses the level of change within the figures of the tables and flags those tables with the largest amount of change. All results are combined in an html report which is then checked by a member of the team. Anything listed in the html report is investigated to see whether it is expected. The second question requires members of the team to use and check the Tabulation Tool to ensure that all tables/breakdowns/footnotes are available and as expected. At any point during the quality assurance process if an error is spotted an incident will be raised and will be thoroughly investigated. When any SAS program is ran a log file is stored for future reference and to see where any part of

Methodology Statement for the Department for Work and Pensions Statistical Summary

the process may have broken down. If required, any part of the process can be re-run and all checking will re-occur.

Once completed, the lead Statistician will ensure everything has been checked and passed, and once happy with everything, will sign off the 5% Sample Tabulation Tool. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Pre-Publication

Once the 5% Sample Tabulation Tool has been signed off, it is then sent to the external web host and Nomis in preparation for publication. The external web host currently hosts the various Tabulation Tools, including the 5% Sample Tabulation Tool. All data from previous quarters resides on the external site and is already live on the internet. When new data is created each quarter it must be uploaded onto the hosting site and made live alongside the existing data.

When the data has been received by the external web host and uploaded for publication it is the responsibility of the team to ensure that the same data exists on the hosting site as was created in the first place. So, further quality assurance must be carried out at this stage. The external web host organises a secured pre-release window to the site for a few days in the week prior to release day and notifies the team of how to access this. A member of the team will then run an automated SAS checking program, which compares the 5% Sample Tabulation Tool on our system to that on the external web site. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the external website and the pre-release access window is closed.

The 5% Sample data is also sent to Nomis (<https://www.nomisweb.co.uk/Default.asp>) in preparation for publication. When the data has been received by Nomis and uploaded for publication it is the responsibility of the team to ensure that the correct data exists on the website and so further quality assurance must be carried out at this stage. Nomis organises a secure pre-release window prior to release day and notifies the team of how to access this. A member of the team will then carry out some manual checks. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the Nomis website.

Publication

The 5% Sample Tabulation Tool is added to the external website and Nomis at 09:30am on Release Day. Users can access the website via the DWP Internet pages. The 5% Sample Tabulation Tool is accessible via the Create Your Own Statistics landing page which itself is available via the DWP Statistical Summary Landing Page.

Methodology Statement for the Department for Work and Pensions Statistical Summary

The DWP Statistical Summary landing page can be accessed here:

http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries

The Create Your Own Statistics landing page can be accessed here:

<http://research.dwp.gov.uk/asd/index.php?page=tabtool>

Additionally, the data published on the WPLS Tabulation Tool is available via the Nomis website:

<https://www.nomisweb.co.uk/Default.asp>

Neighbourhood Statistics

Introduction

The Neighbourhood Statistics tables (<http://83.244.183.180/NESS/page1.htm>) contain benefit claimant numbers broken down to small geographical levels such as Lower Layer Super Output Area (LSOA) and Frozen 2003 Ward. These Ward and LSOA geographies are stable and part of a hierarchy of stable geographies in the UK of which the Census Output Area (COA) is the smallest building block. The Ward structure published is the frozen 2003 Ward as opposed to up-to-date electoral wards which are subject to change. Neighbourhood Statistics are available for Attendance Allowance (AA), Disability Living Allowance (DLA), Employment and Support Allowance (ESA), Incapacity Benefit / Severe Disablement Allowance (IB/SDA), Income Support (IS), Jobseekers Allowance (JSA), Pension Credit (PC) and State Pension (SP). A separate table is also available for Working Age Client Group (WACG) and provides counts of working age claimants of key benefits, broken down by their statistical group (their main reason for interacting with the benefit system).

The best statistics to use are sourced from WPLS as these data are based on 100% of claimants and cover information such as age and gender of claimant, duration of their spell on benefit and geographical locations of claimants. Like the WPLS Tabulation Tool, the Neighbourhood Statistics are sourced from WPLS and are released quarterly (in May, August, November and February). They are produced and released under the control of the DWP Statistical Summary Virtual Team. Although sourced from the same location, due to the size of the Neighbourhood Statistics tables, it is not practical to incorporate these into the WPLS Tabulation Tool.

Production

As the Neighbourhood Statistics are derived from the same source data as the WPLS Tabulation Tool we can source the Neighbourhood Statistics from the Frozen Datasets, produced and quality assured as part of the WPLS Tabulation Tool process (see Production Stage 1 of [WPLS Tabulation Tool](#)). The Frozen Datasets carry forward larger geographical variables to the Tabulation Tool such as Local Authority (LA) and Region. However smaller geographical variables such as LSOA exist on the Frozen Datasets. As part of Quality assurance for the Frozen Datasets, we carry out checks on these smaller geographies (outlined below). Once Quality assurance of the Frozen Datasets is complete we can begin with the next stage of Neighbourhood Statistics production.

We run a SAS program which merges the individual benefit Frozen Datasets with the Pers dataset (see Production Stage 1 of [WPLS Tabulation Tool](#)) to create the

Methodology Statement for the Department for Work and Pensions Statistical Summary

Neighbourhood Statistics dataset. This dataset is used to create every single final Neighbourhood Statistics dataset for publication.

It should be noted that figures are subject to rounding techniques for disclosure control purposes. This is in line with DWP's Confidentiality Statement (http://research.dwp.gov.uk/asd/policy/Confidentiality_Statement_final.pdf) and as an additional protective measure; details of the methodology are not published.

There is then a SAS program to create the Neighbourhood Statistics, one for each benefit. The SAS programs draw on the merged dataset above. The published Neighbourhood Statistics take the form of Excel spreadsheets. These spreadsheets exist as templates containing the required dis-aggregations which are populated once the SAS programs are ran. The program must be run for each benefit in turn and is then subject to a significant amount of checking.

Quality Assurance

Due to the hierarchical nature of geography, quality assurance of the Frozen Datasets for the purpose of the Neighbourhood Statistics focuses solely on the smallest published geographical variable, namely LSOA under the assumption that conclusions drawn from quality assurance of the LSOAs can be extended to Wards. The quality assurance performs a comparison between the numbers in each LSOA in the new quarter compared to the previous quarter. Instead of purely looking at the level of change we consider the way in which the numbers have changed. This can be done by considering flows of records between the quarters, i.e. how many people have joined the claimants in each LSOA, how many have moved between areas etc. This is done via a series of automated SAS programs. There is a quality assurance program for each benefit, e.g. so each benefit in turn needs to be analysed. This is to ensure that there are no benefit-specific problems which may be missed running a generic program. Before proceeding to production of the Excel tables, quality assurance must be run and signed off for every single benefit.

Within the SAS run for creating the Excel spreadsheets there are some datasets created to serve as final quality assurance checks. Also, as data is subject to rounding techniques for disclosure control purposes we must ensure that this has worked correctly. Once completed, the lead Statistician will ensure everything has been completed successfully.. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Pre-Publication

Once the Neighbourhood Statistics tables have been signed off they are then sent to Nomis (<https://www.nomisweb.co.uk/Default.asp>) in preparation for publication. When the data has been received by Nomis and uploaded for publication it is the responsibility of the team to ensure that the correct data exists on the website and so

Methodology Statement for the Department for Work and Pensions Statistical Summary

further Quality Assurance must be carried out at this stage. Nomis organises a secure pre-release window prior to release day and notifies the team of how to access this. A member of the team will then carry out some manual checks. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the Nomis website.

We manually add the Neighbourhood Statistics to the external DWP website ready for Release Day.

Publication

The Neighbourhood Statistics tables are manually added to the external website at 9:30am on Release Day by a member of the team. The team have access to the site and can manually update/amend content when necessary. The Neighbourhood Statistics are also added to the Nomis website at 09:30am on Release Day. Users can access the site via the DWP Internet pages. The Neighbourhood Statistics tables are accessible via the Neighbourhood Statistics Data landing page which itself is available via the DWP Statistical Summary Landing Page.

The DWP Statistical Summary landing page can be accessed here:

http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries

The Neighbourhood Statistics Data landing page can be accessed here:

<http://83.244.183.180/NESS/page1.htm>

Additionally, the Neighbourhood Statistics are available via the Nomis website:

<https://www.nomisweb.co.uk/Default.asp>

Periodically, updates are also made to the Office for National Statistics Neighbourhood Statistics pages at:

<http://www.neighbourhood.statistics.gov.uk/dissemination/>

New Deal Tabulation Tools

Introduction

The New Deal Tabulation Tools for DWP employment programmes (http://research.dwp.gov.uk/asd/asd1/tabtools/index.php?page=tabtool_nd) provide the user with an interactive tool to select one of thousands of possible tabulations. These Tabulation Tools provide up to date numbers on starters, leavers, participants and jobs gained for the various programmes. There is a Tabulation Tool for each of the following New Deal schemes: NDLP (Lone Parents), NDYP (Young People), ND25+ (Twenty-Five Plus), NDDP (Disabled People), ND50+ (Fifty Plus), NDP (Partners), and EZ (Employment Zones). However, the Tools for NDYP and ND25+ are no longer updated.

The seven New Deal schemes for which a Tabulation Tool exist each contain a range of different measures of interest. These measures are Starts (both spells and individuals), Jobs Gained (both spells and individuals), Leavers (both spells and individuals), Interviews (both spells and individuals), and Participants. For some of the New Deal Tabulation Tools there are a few measures within some of the schemes which are not published due to poor quality data. These are explicitly mentioned on the homepage of the New Deal Tabulation Tools.

The New Deal Tabulation Tools are released quarterly (in May, August, November and February) and are produced and released under the control of the DWP Statistical Summary Virtual Team. A section of the DWP Quarterly Statistical Summary relates to DWP Employment Programmes, which are sourced from the New Deal Tabulation Tools.

Production Stage 1

Each New Deal Tabulation Tool is inherently different to the WPLS Tabulation Tool. This is due to the fact that the New Deal Tabulation Tools include cumulative figures relating to each scheme since the schemes began rather than the number of participants at the end of each quarter. The New Deal Tabulation Tools are also subject to a full revision of the cumulative figures each quarter in addition to the new data rather than retaining a frozen snapshot. I.e. they are fully retrospective.

In addition, whereas the WPLS Tabulation Tool contains only two measures (number of claimants and average amount), the New Deal Tabulation Tools contain three different measures which are spells, individuals, and participants. A spell is an occurrence of an action whilst on the New Deal scheme such as starting the scheme, gaining a job through the scheme, or leaving the scheme. Every individual can have more than one spell on a New Deal program. The individuals' measure refers to only one spell per individual, the latest spell. A participant refers to a spell on a New Deal

Methodology Statement for the Department for Work and Pensions Statistical Summary

program at a point in time; this measure is more equivalent to the number of benefit claimants.

The process begins with the receipt of the raw New Deal Databases from data supplier who in turn are responsible for their extraction from the Labour Market System (LMS). The extracts relating to the different New Deal schemes are received at different times in a three week period, meaning that the process does not begin for all schemes at the same time. The extracts arrive having already undergone quality assurance by the external data supplier. As mentioned before the extracts contain a complete historical revision to previous extracts as well as entirely new records relating to the latest time period. The extracts are picked up by a member of the DWP Statistical Summary team for further processing, referred to as WPLS Processing. The reason for this stage is to use WPLS information in order to find extra jobs and improve destination information within the extracts. Extracts arrive on a monthly basis, however, only the quarterly extracts are processed through to the New Deal Tabulation Tools.

Quality Assurance Stage 1

Quality assurance of the New Deal Databases is an automated process carried out in SAS and Excel with manual analysis of output. The aim of the quality assurance is to identify areas where there could be issues with the data which has been processed via the pre-processing SAS code. Other than simply observing the numbers, we use Excel to get a graphical representation of the movement each variable has been encountering, so we can easily identify if there are any issues. Issues may include unexpected changes in variable frequencies, so viewing a noticeable drop or increase would flag up a concern. Long term variable trends (whether something is increasing/decreasing/staying roughly the same/some sort of repeating seasonal pattern etc.) and short term volatilities are monitored and all potential issues and suspect movements in the datasets are raised and investigated. It is often the case where movements can be attributed to change in policy and are fully expected. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Production Stage 2

Once the New Deals Datasets Quality assurance is completed and fully signed off a process known as Cube creation begins. This part of the process is also carried out in SAS and just re-formats the New Deals Datasets into a format compatible with the next step, which is the Tabulation Tool creation process. The cubes set-up most of the desired categories for the Tabulation Tool, for example age variables are formatted into age bands and Local Authority variables are numbered to be put into a desired order. There is one New Deal Cube created for each New Deal scheme and for each breakdown available on the Tabulation Tool. There are separate cubes for

Methodology Statement for the Department for Work and Pensions Statistical Summary

starters spells and starters individuals and jobs spells. Each Cube contains a full revision to the entire back-series.

Before the Tabulation Tool can be run we update the Modsheets. These are a set of spreadsheets, which contains every single parameter required for the run. There is an individual Modsheet for each New Deal Tabulation Tool. All formats, titles, labels, links, and footnotes are defined here as well as which parts of the Tabulation Tool are required from the run. Changes to the Modsheets are often required each quarter. Changes in benefit policy can often require additional footnotes each quarter, or Parliamentary Constituencies may change and require a new format. Once the Modsheets are completed, the Tabulation Tool production can begin. The Tabulation Tool is essentially a series of html pages which contain different cross-tabulations of data alongside appropriate titles, labels, and footnotes for that data. JavaScript code exists which create the “front-end” parts of the Tabulation Tool. These are the interactive menus which sit on top of the html pages. The user makes selections on these pages and navigates to the appropriate table of data as per the selection made.

There is an extra table which must be prepared and released alongside the New Deal Tabulation Tools. This table relates to Leavers and Participants statistics for ND50+. These series have had long-term data quality issues with some of the breakdowns and so were subsequently removed from the Tabulation Tool. We now produce headline numbers for this series alongside limited breakdowns to be published as a one-click table. The one-click table shows cumulative leavers and current participants broken down by gender, age, and ethnicity.

Quality Assurance Stage 2

Extensive quality assurance of the Cubes must take place before the Tabulation Tools are run. The nature of the New Deal statistics are that in a new quarter the entire back-series is revised according to the retrospection contained within the latest extract. Therefore the entire history is processed and published each quarter, with the addition of a new quarter. Another feature is that the published statistics are cumulative, showing figures since the scheme began. One problem of these inherent properties is that quality assurance of the final cumulative figures to be published sometimes masks underlying features of the new quarterly data. It is difficult to distinguish the known retrospection within the data from the entirely new quarterly data which has not been seen before. Furthermore, the properties of this new quarterly data are hidden amongst the properties of the cumulative stock. For example, each new starter spell for NDYP may be for an 18 year old which may be taken as an odd occurrence which could be missed when observing how the cumulative age bands change between quarters.

The idea behind the quality assurance is to distinguish between the retrospective aspect of the cube and the new data within the cube. The retrospective aspect examines what is now known about the New Deal scheme based upon more up-to-

Methodology Statement for the Department for Work and Pensions Statistical Summary

date information compared with what was published last quarter. This checks to see that there has not been too much retrospection in the data, and that this retrospection is restricted to more recent time periods. The data within the cube is examined to check the distribution of different variables against what was observed last quarter, under the assumption that roughly the same distribution of new records would be observed. This quality assurance process is performed by the team as soon as the New Deal cubes have been created.

Once the Tabulation Tool has successfully been created, the team are required to quality assure the output. There are two questions to answer during this stage: Are the html pages correct and display all the correct data as expected? Does the tool have all the correct functionality? The first question is addressed by use of an automated SAS checking program which directs attention to those pages showing the most change. The idea behind this program is to compare the html page from the new quarter with the identical html page from the previous quarter. The program checks for things such as missing rows and columns from the table, missing footnotes, new footnotes and changed labels. It also assesses the level of change within the figures of the tables and flags those tables with the largest amount of change. All results are combined in an html report which is then checked by a member of the team. Anything listed in the html report is investigated to see whether it is expected. The second question requires members of the team to use and check the Tabulation Tool to ensure that all tables/breakdowns/footnotes are available and as expected. At any point during the quality assurance process if an error is spotted an incident will be raised and will be thoroughly investigated. When any SAS program is ran a log file is stored for future reference and to see where any part of the process may have broken down. If required any part of the process can be re-run and all checking will re-occur.

Once completed, the lead Statistician will ensure everything has been checked and passed, and if happy with everything, will sign off the New Deal Tabulation Tools. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Pre-Publication

Once the New Deal Tabulation Tools have been signed off they are then sent to the external web host in preparation for publication. The external web host currently hosts the various Tabulation Tools, including the New Deal Tabulation Tools.

When the data has been received by the external web host and uploaded for publication it is the responsibility of team to ensure that the same data exists on the external site as was originally created. So, further quality assurance must be carried out at this stage. The external web host organises a secured pre-release window to the hosting site for a few days in the week prior to release day and notifies the team how to access this. A member of the team will then run an automated SAS checking program, which compares the New Deal Tabulation Tools on our system to that on

Methodology Statement for the Department for Work and Pensions Statistical Summary

the external site. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the external website and the pre-release access window is closed.

Publication

The New Deal Tabulation Tools are added to the external website at 09:30am on Release Day. Users can access the site via the DWP Internet pages. The New Deal Tabulation Tools are accessible via the Create Your Own Statistics landing page which itself is available via the DWP Statistical Summary Landing Page.

The DWP Statistical Summary landing page can be accessed here:

http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries

The Create Your Own Statistics landing page can be accessed here:

<http://research.dwp.gov.uk/asd/index.php?page=tabtool>

The New Deal Tabulation Tools landing page can be accessed here:

http://research.dwp.gov.uk/asd/asd1/tabtools/index.php?page=tabtool_nd

Sanctions Tabulation Tool

Introduction

The Sanctions Tabulation Tool for Jobseekers Allowance (JSA) Sanctions and Disallowances (<http://83.244.183.180/sanction/sanction/LIVE/tabtool.html>) provide the user with an interactive tool to select one of thousands of possible tabulations. This Tabulation Tools provide six breakdowns: Referrals for fixed length, Referrals for varied length, Referrals for entitlement decisions, Opinions on fixed length, Opinions on varied length, and Entitlement Opinions. Each of these has a referrals and individuals measure.

The Sanctions Tabulation Tool is released quarterly (in May, August, November and February) and is produced and released under the control of the DWP Statistical Summary Virtual Team. A section of the DWP Quarterly Statistical Summary relates to Jobseekers Allowance (JSA) Sanctions and Disallowances, which are sourced from the Sanctions Tabulation Tool.

Production Stage 1

The Sanctions Tabulation Tool is inherently different to the WPLS Tabulation Tool and rather similar to the New Deal Tabulation Tools. This is due to the fact that the Sanctions Tabulation Tool includes cumulative figures relating to each type of sanction/disallowance since April 2000 rather than the number currently receiving a sanction at the end of each quarter. There are two different measures within the Sanctions Tabulation Tool: Referrals and Individuals. Every individual can have more than one referral for each type of sanction. In this way the referrals are similar to New Deal spells. The individuals' measure refers to only one referral per individual, the latest referral. The Sanctions Tabulation Tools is also subject to a full revision of the cumulative figures each quarter in addition to the new data rather than retaining a frozen snapshot i.e. they are fully retrospective.

Production begins with the creation of the Sanctions Cumulative Datasets. The Cumulative Datasets created for the Sanctions process aim to capture each type of sanction/disallowance occurring since April 2000. They are run using SAS and are derived from merging data from the Labour Market System (LMS) and Decision Making and Appeals System (DMAS).

Quality Assurance

Quality assurance of the Cumulative Datasets for Sanctions is an automated process carried out in SAS with manual analysis of output. All potential issues and suspect movements in the datasets are raised and investigated. It is often the case where

movements can be attributed to change in benefit policy and are fully expected. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Production Stage 2

Once the Cumulative Dataset quality assurance is completed and fully signed off a process known as Cube creation begins. This part of the process is also carried out in SAS and just re-formats the Cumulative Datasets into a format compatible with the next step, which is the Tabulation Tool creation process. The cubes set-up most of the desired categories for the Tabulation Tool, for example age variables are formatted into age bands and Local Authority variables are numbered to be put into a desired order. The Cubes contains a full revision to the entire back-series.

Before the Tabulation Tool can be run we update the Modsheets. These are a set of spreadsheets, which contains every single parameter required for the run. All formats, titles, labels, links, and footnotes are defined here as well as which parts of the Tabulation Tool are required from the run. Changes to the Modsheet are often required each quarter. Changes in policy can often require additional footnotes each quarter, or Parliamentary Constituencies may change and require a new format. Once the Modsheet is completed, the Tabulation Tool production can now begin. The Tabulation Tool is essentially a series of html pages which contain different cross-tabulations of data alongside appropriate titles, labels, and footnotes for that data. JavaScript code exists which create the “front-end” parts of the Tabulation Tool. These are the interactive menus which sit on top of the html pages. The user makes selections on these pages and navigates to the appropriate table of data as per the selection made.

Quality Assurance Stage 2

Quality assurance of the Cubes must take place before the Tabulation Tools are run and is done through a series of Excel spreadsheets. The nature of the Sanctions statistics are that in a new quarter the entire back-series is revised according to the retrospection contained within the latest extract. Therefore the entire history is processed and published each quarter, with the addition of a new quarter. Another feature is that the published statistics are cumulative, showing figures since April 2000. One problem of these inherent properties is that quality assurance of the final cumulative figures to be published sometimes masks underlying features of the new quarterly data. It is difficult to distinguish the known retrospection within the data from the entirely new quarterly data which has not been seen before. Furthermore, the properties of this new quarterly data are hidden amongst the properties of the cumulative stock. The idea behind the Quality assurance is to distinguish between the retrospective aspect of the cube and the new data within the cube. The retrospective aspect examines what is now known about the Sanction based upon more up-to-date information compared with what was published last quarter. This

Methodology Statement for the Department for Work and Pensions Statistical Summary

checks to see that there has not been too much retrospection in the data, and that this retrospection is restricted to more recent time periods. The data within the cube is examined to check the distribution of different variables against what was observed last quarter, under the assumption that roughly the same distribution of new records would be observed.

Once the Tabulation Tool has successfully been created, the team are required to Quality assure the output. There are two questions to answer during this stage: Are the html pages correct and display all the correct data as expected? Does the tool have all the correct functionality? The first question is addressed by use of an automated SAS checking program which directs attention to those pages showing the most change. The idea behind this program is to compare the html page from the new quarter with the identical html page from the previous quarter. The program checks for things such as missing rows and columns from the table, missing footnotes, new footnotes and changed labels. It also assesses the level of change within the figures of the tables and flags those tables with the largest amount of change. All results are combined in an html report which is then checked by a member of the team. Anything listed in the html report is investigated to see whether it is expected. The second question requires members of the team to use and check the Tabulation Tool to ensure that all tables/breakdowns/footnotes are available and as expected. At any point during the quality assurance process if an error is spotted an incident will be raised and will be thoroughly investigated. When any SAS program is ran a log file is stored for future reference and to see where any part of the process may have broken down. If required any part of the process can be re-run and all checking will re-occur.

Once completed, the lead Statistician will ensure everything has been checked and passed, and if happy with everything, will sign off the Sanctions Tabulation Tool. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Pre-Publication

Once the Sanctions Tabulation Tool has been signed off it is then sent to the external web host in preparation for publication. The external web host currently hosts the various Tabulation Tools, including the Sanctions Tabulation Tools.

When the data has been received by the external web host and uploaded for publication it is the responsibility of team to ensure that the correct data exists on the hosting site. So, further quality assurance must be carried out at this stage. The external web host organises a secure pre-release window to the hosting site for a few days in the week prior to release day and notifies the team of how to access this. A member of the team will then run an automated SAS checking program, which compares the Sanctions Tabulation Tool on our system to that on the external site. Provided there are no missing tabulations, no changes and no problems then we sign off the content on the external website and the pre-release access window is closed.

Publication

The Sanctions Tabulation Tool is added to the external site at 09:30am on Release Day. Users can access the site via the DWP Internet pages. The Sanctions Tabulation Tool is accessible via the Create Your Own Statistics landing page which itself is available via the DWP Statistical Summary Landing Page.

The DWP Statistical Summary landing page can be accessed here:

http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries

The Create Your Own Statistics landing page can be accessed here:

<http://research.dwp.gov.uk/asd/index.php?page=tabtool>

Housing Benefit / Council Tax Benefit Statistics

Introduction

Housing Benefit / Council Tax Benefit (HB/CTB) statistics detail the number of people in receipt of Housing Benefit and Council Tax Benefit and the amount of benefit received at a particular point in time. They also include breakdowns of Region, Local Authority, Tenancy type, Passported benefits, Age group and Family Type. HB/CTB statistics are released monthly and are produced and released under the control of the DWP Statistical; Summary team. A substantial part of the DWP Monthly/Quarterly Statistical Summary relates to HB/CTB statistics.

Production

HB/CTB Statistics are sourced from the Single Housing Benefit Extract (SHBE) and are derived and published on a monthly basis from SHBE extracts using a combination of SAS (to process the data) and Excel (to present the statistics). The monthly process of publishing Housing Benefit and Council Tax Benefit National Statistics uses two months of SHBE extracts (i.e. recipient statistics relating to Feb11 use the Feb11 and Mar11 SHBE extracts). These are combined to produce two frozen datasets (one for HB and one for CTB) which contain all live claims on a set date (2nd Thursday of the month). An in depth document outlining the HB/CTB process (and SHBE as a source) is available at:

http://research.dwp.gov.uk/asd/asd1/hb_ctb/caseload_meth.pdf.

Quality Assurance

Quality assurance of SHBE data and recipient statistics takes place in three phases; when the SHBE data is delivered to the team, when the HB/CTB frozen datasets are created, and when the final statistics are produced.

When the SHBE data is delivered to the team, a host of Excel files are made available. The Excel files highlight various features within the SHBE datasets compared to the previous version, such as missing variables, new decodes, missing values, and changes to the distribution of variables. A list of Local Authorities missing in the SHBE extract is also given and any problematic Local Authority data is highlighted. A member of the team will check the Excel files and when happy that the SHBE data is accurate, will commence the HB/CTB statistics production.

When the HB/CTB frozen datasets are created, several stages of quality assurance takes place. The SAS programs used to create the HB/CTB Frozen Datasets are run in

Methodology Statement for the Department for Work and Pensions Statistical Summary

two stages. The first stage doesn't create the Frozen Datasets but identifies all Local Authorities which will be substituted during this run, by outputting two html files (one for HB and one for CTB). As previously mentioned, occasionally Local Authorities will not submit data or will submit data that shows an infeasible recipient load. To avoid publishing incorrect figures we undergo the process of substitution; if a Local Authority's data is missing or looks dubious for the current month this data is deleted and replaced with data from the previous month. This is so that an attempt to contact the Local Authority can be made to see why these changes occurred and whether they were valid. The list of Local Authorities in these html files should include all those listed in the Excel files received from the data loading process. Any substituted Local Authorities will be investigated and resolved in subsequent months. Once all Local Authorities to be substituted have been identified the second stage creates the HB/CTB Frozen Datasets. As part of this process another two html files are output. These files compare the current month's HB/CTB Frozen Datasets with the previous month's and highlight various features within the HB/CTB Frozen Datasets compared to the previous month's, such as missing variables, new decodes, missing values, and changes to the distribution of variables. Occasionally an additional Local Authorities may be identified as dubious and wasn't identified earlier. The HB/CTB Frozen Datasets will be re-run and quality assurance will take place again.

When the final statistics are produced in Excel they are again quality assured to ensure that everything has been populated correctly. This is done through a series of automated Excel macros and is a relatively easy task.

Once all quality assurance is completed, the lead Statistician will ensure everything has been checked and passed, and if happy with everything, will sign off the HB/CTB statistics. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Publication

The HB/CTB tables are added to the DWP website at 09:30am on Release Day. The publication has its own landing page which contains the tables and key information about HB/CTB statistics and also includes a link to the latest Statistical Summary: <http://research.dwp.gov.uk/asd/index.php?page=hbctb>

The DWP Statistical Summary

Introduction

The Statistical Summary brings together all of the relevant information and figures from all aspects of DWP Statistical Summary and where necessary provides links to more detailed information. It can only be completed once all other DWP Statistical Summary products have been completed and quality assured. The Summary is a word document published on the DWP website as a pdf document. The Summary primarily contains headline figures (e.g., national totals, top-level figures) and is published monthly, containing Housing Benefit and Council Tax Benefit National Statistics together with official statistics giving early estimates of numbers of inactive benefit claimants, plus a link to the latest report on Flexible New Deal. Each quarter (in May, August, November and February), a larger summary also contains the detail of DWP administered benefits, some employment programmes, JSA sanctions, vacancies and CSA.

Production

The Statistical Summary is a relatively simple document to produce. The format of the Monthly/Quarterly Summary rarely changes between releases (in accordance with the Code of Practice); therefore production just involves sourcing the relevant updated figures from the new Tabulation Tools and/or other publications.

A section on Vacancies is included in the DWP Quarterly Statistical Summary and is sourced from Nomis (<https://www.nomisweb.co.uk/Default.asp>). The Vacancies data on Nomis are sourced from Jobcentre Plus' Labour Market System (LMS), an administrative computer system covering 100% of vacancies notified to Jobcentre Plus. Figures are updated monthly with no revisions to previous month's data. A section on Industrial Injuries Disablement Benefit (IIDB) is included in the DWP Quarterly Statistical Summary and is taken from a larger publication (<http://statistics.dwp.gov.uk/asd/index.php?page=iidb>). All language, key points, and themes within the Summary are kept consistent. Footnotes and notes are reviewed on a release by release basis to advise on the interpretation of the figures within the Summary. The Notes section is a very useful section within which to reference other National Statistics products, inform users of upcoming changes and issues..

Each quarter, new, changed or interesting series can be highlighted in a "Focus on" article at the front of the Summary. These are ad hoc articles which are likely to be of interest to readers.

Quality Assurance

A member of the team takes responsibility for updating the Statistical Summary and once completed, it will be quality assured by a different member of the team. Production of the Statistical Summary involves sourcing figures from the newly produced Tabulation Tools and/or other publications, and doesn't begin until each of these have been quality assured and signed off, so we can be relatively certain the figures are correct at point of source. As the figures are sourced manually they still need to be quality assured in case of manual error when copying them across to the Statistical Summary. Where figures are sourced from a separate publication (i.e. Vacancies) the original publication will be checked (in the case of Vacancies this is Nomis), as will the back-series to ensure any updates and/or amendments to the back-series are recorded. Once all figures within the Summary have been quality assured they are then sent onto the lead Statistician for further checking and sign off. All actions are recorded on a release timetable, and are checked off once complete to ensure that no steps are missed.

Pre-Publication

This is by far the most important part of the DWP Statistical Summary process as this is the publication which is pre-announced. DWP pre-announce the release date of the Statistical Summary at least 12 months in advance, in accordance with release practices set out in the Code of Practice for Official Statistics. Dates of future publications can be found on the DWP internet via the statistics publication schedule web page (<http://research.dwp.gov.uk/asd/index.php?page=schedule>) and via the National Statistics Publication Hub (<http://www.statistics.gov.uk/hub/index.html>)

As the Statistical Summary release is the National Statistic there is an internal departmental requirement to accompany this release with a Submission. The Submission is circulated to the Minister's office and contains input from policy areas. The Submission is sent one day prior to release day and is fully compliant with the UKSA rules surrounding pre-release access of statistics. It contains a blend of the most notable issues within the latest release, key features, and policy guidance around these things. This document is also quality assured in the same manor as the Statistical Summary, before being sent out.

Publication

The Statistical Summary is added to the DWP website at 09:30am on Release Day. The publication has its own landing page which contains the publication itself as well as key information about the contents of the release. Old versions of the Statistical Summary are archived and still available. The landing page can be accessed here: http://research.dwp.gov.uk/asd/index.php?page=statistical_summaries.