Numbers of pension schemes from the Occupational Pension Schemes Survey

A methodology update

Bryan Cotterell, Alexandra Wojewodka-Dunn and Sian Summers

Abstract

This article describes the method currently used by the Office for National Statistics to estimate the number of occupational pension schemes from the Occupational Pension Schemes Survey (OPSS). It then outlines a proposed new methodology and presents a table, using data from 2007-2009, comparing the scheme count estimates using both methodologies. It is proposed to repeat the analysis when the 2010 data becomes available.

Contents

Introduction	2
Current weighting methodology	3
Proposed new methodology	3
Assumptions	4
Comparison.....	4
Estimating standard errors.....	5
Conclusion	6
Annex 1 – Scheme count calculations	7
Annex 2 – Estimating standard errors for counts of schemes in the Occupational Pension Scheme Survey.....	9

Introduction

The Occupational Pension Schemes Survey (OPSS) collects information from pension schemes about scheme membership, number of schemes, benefits, contributions and includes sections on very small schemes (schemes with 2 to 11 members) and those that are winding up. OPSS covers both private and public sector occupational pension schemes registered in the United Kingdom. OPSS includes analysis by scheme status (open, closed, frozen and winding up); by size and by benefit structure (defined benefit and defined contribution) and distinguishes between private and public sector schemes. It does not cover individual or group personal pensions or stakeholder pensions, which are based on individuals entering into a contract with a pension provider; nor does it cover state pensions. Responsibility for running the Occupational Pension Schemes Survey (OPSS) was transferred from the Government Actuary's Department (GAD) to the Office for National Statistics (ONS) in 2005. An annual report presenting the results of the survey and a statistical bulletin are published each year and are available on the ONS website.

In 2008, scheme numbers were not published in the annual report. Figures for scheme numbers have always been weaker than other OPSS estimates as the survey is designed primarily to measure membership numbers. During the 2008 survey we were alerted to the fact that the survey design as it stands was not sufficient to produce robust estimates for scheme numbers. As a consequence it was decided that scheme numbers would not be published until the methodology division in the ONS had investigated the survey design.

The sample each year consists of around 1,650 schemes from the 60,000 to 70,000 on The Pension Regulator's (TPR) register. For schemes with 2 to 11 members only one scheme in 250 is sampled and it was thought that the problem was with these very small schemes.

The very small schemes make up a significant proportion of the total counts for scheme numbers but are, by nature, negligible when considering the membership estimates for all pension schemes, as indicated below:

Table 1: Impact of very small schemes on total counts for scheme numbers

	numbers	
	All schemes	Very small schemes
2007	54,110	44,370
2008	55,290	43,750

Source: ONS analysis based on Occupational Pension Schemes Survey data

Table 2: Impact of very small schemes on total membership for all schemes

	millions	
	All schemes	Very small schemes
2007	8.8	0.11
2008	9.0	0.08

Source: ONS analysis based on Occupational Pension Schemes Survey data

The figures for the very smallest schemes are only used in the calculation of these two aggregates; for other variables, such as contribution rates, the amounts from the very smallest schemes have always been omitted so the quality of the data has not been an issue.

The methodology division in ONS have examined the current weighting methodology and scheme estimates. This paper briefly explains the current methodology, introduces the proposed new methodology and finishes with a table comparing the scheme count estimates.

Current weighting methodology

The existing weighting scheme for OPSS uses an estimator which expands the results for the sample to estimate for the population, taking into account those schemes that did not respond (see Annex 1).

When selecting the sample the population is divided into 6 different size-bands (strata) based on membership from the Pension Regulator's register.

Band A 10,000+
Band B 5,000-9,999
Band C 1,000-4,999
Band D 100-999
Band E 12-99
Band F 2-11

All the weighting calculations take place at a stratum level and for public and private sector schemes.

The design weight (the expansion estimator) for each stratum is calculated by dividing the number of schemes in the population by the number of schemes that were actually sampled. The final weight (the expansion estimator with the non response adjustment) is then calculated by multiplying the design weight by the total number of scheme members in the sample in each stratum divided by the total number of scheme members in the responding sample in the stratum.

This weighting methodology was designed primarily to estimate membership which is perhaps the key variable from the survey. It has been applied to both membership totals and to counts of schemes. This paper considers separate methodologies for both variables. The proposed new methodology for counts of schemes is intended to produce more robust scheme count estimates and will be discussed in the next section.

Proposed new methodology

The weighting scheme proposed in this section of the report is intended to estimate for scheme counts only while the existing methodology will continue to be used to estimate membership. This new methodology attempts to correct for the identification of out-of-scope schemes at different stages of survey processing. Examples of out-of-scope schemes are a scheme with less than two members, or a scheme that has finished the winding up process. Out-of-scope correction was not

considered in the previous methodology as the weighting did not distinguish between in and out-of-scope schemes but merely looked at response (see Annex 1).

The selection of the OPSS sample is carried out by taking a stratified simple random sample without replacement¹ drawn from the TPR register. This selection is examined and some schemes are telephoned before the questionnaires are sent in a cleaning exercise to verify key variables. Schemes found to be out-of-scope from this exercise are not sent a questionnaire. This is the first stage of the detection of out-of-scope schemes. When the questionnaires are returned to the ONS, a second stage of out-of-scope detection takes place. Finally, during processing of the survey results, some further schemes may be identified as out-of-scope of the survey.

For the first stage of out-of-scope identification, we know exactly what proportion of schemes were found to be out-of-scope. We can therefore estimate the number of schemes in the population that would have been excluded had they been selected for the survey. The estimate of the total number of in-scope schemes in the population is calculated by multiplying, for each stratum, the total number of schemes in the population by the proportion of in-scope schemes found within the sample.

In the case of the second and third stages of out-of-scope identification, the situation is more difficult, because some schemes may not respond. Since both the second and third stages of out-of-scope identification take place after the responses have been received, they may be treated as one step. To estimate the final number of in-scope schemes, it is proposed to use the estimate from the first stage (total number of in-scope schemes in the population) and multiply it by the proportion of in-scope schemes found in the responding sample. This is again done for each stratum.

The final weights are then calculated by dividing this figure by the number of in-scope schemes that were found in the responding sample for the given stratum.

Assumptions

The proposed estimation method for scheme counts assumes:

- the response rate for the second step is equal between the in-scope and out-of-scope units
- the sampling frame for each stratum is complete
- there are no duplicate units on the TPR register.

Comparison

In the case of scheme counts, the current method can produce estimates of scheme counts that exceed the number of schemes on the TPR register; this is clearly not a desirable feature. The new estimator can only produce estimates of counts that are feasible i.e. less than or equal to the register totals. The new estimator also improves the treatment of non-response. The original estimator treats all steps of non-response identification as one step. Therefore, it cannot account for the fact that non-response does not affect the first stage of out-of-scope identification. The new method does account for this.

¹ Cochran, W.G. (1977). *Sampling Techniques*, 3rd Edition. Wiley: New York.

The differences in the scheme counts provided by the existing and new estimators are shown in the table below.

Table 3 Estimates of the number of occupational pension schemes by status and sector 2007 to 2009

Number of occupational pension schemes by status and sector 2007 to 2009

United Kingdom		red= existing methodology		black = new methodology		Numbers	
		2007	2007	2008	2008	2009	2009
		existing methodology	new methodology	existing methodology	new methodology	existing methodology	new methodology
Open		28,810	30,595	34,960	34,262	25,751	25,161
	Public	130	139	240	236	226	231
	Private	28,680	30,456	34,720	34,027	25,525	24,930
Closed		19,110	20,230	10,509	10,104	16,819	16,509
	Public	120	121	20	20	24	25
	Private	18,990	20,109	10,210	10,084	16,795	16,484
Frozen		4,410	4,614	7,224	6,943	4,081	4,109
	Public	60	60	70	72	56	57
	Private	4,350	4,554	6,910	6,871	4,025	4,052
Winding up		1,780	1,888	3,248	3,129	3,072	3,109
	Public	0	0	0	0	0	0
	Private	1,780	1,888	3,150	3,129	3,072	3,109
Total (open, closed, frozen and winding up)		54,110	57,328	55,320	54,438	49,723	48,889
	Public	310	321	330	327	307	314
	Private	53,800	57,007	54,990	54,111	49,417	48,575
Out of scope		23,593	20,375	20,350	21,232	12,107	12,941
	Public	29	18	12	15	29	22
	Private	23,564	20,357	20,338	21,217	12,077	12,919
TPR Register total		77,703	77,703	75,670	75,670	61,830	61,830
	Public	339	339	342	342	336	336
	Private	77,364	77,364	75,328	75,328	61,494	61,494

Source: Occupational Pension Schemes Survey

Estimating standard errors

As set out in Annex 2, it has for the first time been possible to produce some standard errors for the estimates of scheme numbers from the survey. As the Annex shows, those estimates can only be approximate, but they suggest for example that the 95% confidence interval for the number of open schemes for 2009 is broadly plus or minus 4,000 around the figures shown.

The differences between the existing and new estimates shown above are well within this interval.

Conclusion

This investigation has updated the methodology and improved the robustness of the scheme counts estimates; however from the table above it is clear the figures are not hugely different. Further analysis will be carried out when the 2010 data becomes available with the intention of reintroducing scheme numbers into the OPSS annual report.

A further improvement for the estimation of scheme counts would be to increase the sample size for the smaller schemes. However due to the constraints on the total sample size this would mean reducing the number of forms sent to larger schemes. As scheme counts are only one of the many variables that the Occupational Pension Schemes survey collects, it is inevitable that a compromise is involved. The sample is designed to maximise the reliability of the membership totals as the primary focus of the survey is on membership.

This is a topic currently being discussed in the Future for Occupational Pension Schemes Survey consultation which is available here: <http://www.ons.gov.uk/about/consultations/open-consultations/the-future-of-the-occupational-pension-schemes-survey/index.html>. Readers of this article are invited to respond to the consultation. The closing date for responses is the 24th June 2011. A summary of responses will be published on the ONS website approximately 2 months after the consultation closes.

Annex 1 – Scheme count calculations

Current weighting methodology

Design weight
$$a_h = \frac{N_h}{n_h}$$

Final weight
$$w_h = a_h \frac{\sum_{i \in s_h} x_i}{\sum_{i \in s_{h,r}} x_i}$$

Where:

N_h is the number of schemes in the population of stratum h

n_h is the number of sampled schemes in stratum h

w_h is the weight for stratum h

s_h is the full sample in stratum h

$s_{h,r}$ is the responding sample in that stratum

x_i is the auxiliary variable of unit i (in this case register whole scheme membership for unit i)

Proposed new methodology

Stage 1
$$\hat{N}_{h,i1} = N_h \times \frac{n_{h,i1}}{n_h}$$

Stage 2 and 3
$$\hat{N}_{h,i} = \hat{N}_{h,i1} \times \frac{n_{h,ir}}{n_{h,r}}$$

Final weight
$$w_{h,i} = \frac{\hat{N}_{h,i}}{n_{h,ir}}$$

Where:

$\hat{N}_{h,i1}$ is the estimate of the total remaining in the population of stratum h after the first stage of out-of-scope identification

N_h is the number of schemes in the population of stratum h

$n_{h,i1}$ is the number of units in the sample of stratum h that is in-scope at the first stage of out-of-scope identification.

n_h is the number of sampled schemes in stratum h

$\hat{N}_{h,i}$ is the estimator of the number of in-scope units in stratum h ,

$n_{h,ir}$ is the number of in-scope units in the responding sample in stratum h ,

$n_{h,r}$ is the total responding sample size in stratum h .

$w_{h,i}$ is the final weight for in scope units under the new methodology.

Annex 2 – Estimating standard errors for counts of schemes in the Occupational Pension Scheme Survey

Standard Errors

Since the weighting process is in three main steps, ideally this should be incorporated into the estimation of standard errors.

Several methods for doing this are presented by Valliant (2004)². In his paper, only the group jackknife V_J incorporates every part of the weighting process. The various design and estimation steps for the variance estimators are listed in his paper, and will not be discussed here.

The Generalized Estimation System (GES) of Statistics Canada³ is the ONS standard estimation and weighting tool and was used to perform the estimation. The simplest method to achieve this is to treat the out-of-scope identification, which essentially occurs in two steps, as one step. This introduces a small bias in the standard errors relative to the method being used to produce the estimates.

A 'naïve' estimator was applied, which is described as follows: non responding units were excluded from the sample and the design weight was calculated using the full population counts; all out-of-scope schemes were treated as one group; estimation was then carried out using an expansion estimator. The out-of-scope units (i.e. known ineligible) were treated as a domain in the estimation. Coefficients of variation (CVs) from this 'naïve' approach were then calculated.

In practice, the estimates obtained using GES and using the formal weighting scheme are very similar to one another. Therefore, it would appear safe to assume that it is reasonable to use the CVs obtained from GES to approximate the true CVs.

Results

Table 1 shows the point estimates obtained under the new methodology together with the approximate CVs. The counts are dominated by the private sector because of the number of small private sector schemes. The CVs of the TPR register totals for public sector, private sector and total have a CV of zero because these counts are obtained directly from the TPR register and no estimation is involved. In contrast, some of the estimates have relatively large CVs. However, some of these are associated with very small estimates.

² Valliant, R., (2004) 'The Effect of Multiple Weighting Steps on Variance Estimation', *Journal of Official Statistics*, Volume 20, No.1, pp. 1-18.

³ Estevao, V., Hidirolou, M.A., and Särndal, C.-E., (1995) 'Methodological Principles for a Generalized Estimation System at Statistics Canada', *Journal of Official Statistics*, Volume 11, No. 2, pp. 181-204.

Table 1: Estimated Population Counts and Coefficients of Variation for OPSS from 2007 to 2009

Sector	Status	2007		2008		2009	
		Estimate	CV	Estimate	CV	Estimate	CV
Private	Open	30456	8.4%	34027	6.8%	24930	7.7%
Private	Closed	20109	11.2%	10084	14.0%	16484	10.3%
Private	Frozen	4554	23.7%	6871	17.4%	4052	18.2%
Private	Winding up	1888	32.8%	3129	24.6%	3109	20.2%
Private	Out of scope	20357	11.5%	21217	9.9%	12919	11.9%
Private	TPR total	77364	0.0%	75328	0.0%	61494	0.0%
Public	Open	139	3.9%	236	2.8%	231	2.6%
Public	Closed	121	5.6%	20	25.8%	25	19.0%
Public	Frozen	60	10.6%	72	9.5%	57	11.5%
Public	Winding up	0	0.0%	0	0.0%	0	0.0%
Public	Out of scope	18	20.6%	15	27.3%	22	19.8%
Public	TPR total	339	0.0%	342	0.0%	336	0.0%
All	Open	30595	8.4%	34262	6.7%	25161	7.6%
All	Closed	20230	11.2%	10104	14.0%	16509	10.2%
All	Frozen	4614	23.3%	6943	17.2%	4109	17.9%
All	Winding up	1888	32.8%	3129	24.6%	3109	20.2%
All	Out of scope	20375	11.5%	21232	9.9%	12941	11.9%
All	TPR total	77703	0.0%	75670	0.0%	61830	0.0%