Government
Actuary's
Department

Data for social security valuations

Author: Martin Lunnon FIA
E-mail: martin.lunnon@gad.gov.uk

Mailing address: Government Actuary's Department, Finlaison House, 15-17 Furnival Street, London EC4A 1AB United Kingdom
www.gov.uk\gad

IAA PBSS colloquium, Oslo

Monday 8th June 2015

Data for social security valuations

1. Data sufficiency

- What data should be requested to project finances of a social security arrangement (part of a review)

2. Data reliability

- Are the data received correct
- Are they complete

In both cases –remember materiality

Professional background

- ISAP2 (and ISAP1)
- TAS(D) (and other TASs)
- North American standard ASoP23
- ILO internal guidelines/
  ISSA-ILO Actuarial Guidelines

Not all directly applicable to any given piece of work

Professional background – ISAP2 (and ISAP1)

ISAP2

- Section 2.2 – types, sources of data that may be used
- Section 2.3.4 – use of data from sources other than the program being reviewed for new arrangements/benefits
- Section 3.1.3.c – include info on data in report
- Section 3.2.1.a – actuary's opinion that data are sufficient and reliable
- Appendix – more details of data that may be described in report

ISAP1 (covers all actuarial work)

- Section 2.5 on data – sufficiency and reliability; data validation; entity-specific data; data deficiencies

Professional background – TAS D (and other TASs)

TAS D

- Purpose – data subject to sufficient scrutiny so users rely on results; actions on data deficiencies increase reliability
- Definition of data – wide
- Assess need for data for each exercise
- Definition of data
- Documentation
- Validation
- Incomplete or inaccurate data

TAS M – documentation of data for each "realisation"; suitable for model; grouping – need to rationalise and explain; cleaning – explain

TAS R – describe data, state sources; describe uncertainties, fixes

Professional background

North American standard

- Actuarial Standard of Practice 32 "Social Insurance"
- Developed by Academy of Actuaries, adopted by Actuarial Standards Board
- 1998, updated 2011
- No section on data
- "[R]eport should identify the information, documents, and data used, including their sources, and whether the actuary undertook any independent verification" (paragraph 4.1.3)

Professional background

ILO internal guidelines

- "Internal guidelines for actuarial analysis of a national social security pension scheme"
- 1998
- Section 2.2.4.2 on "Data base and assumptions"
- Scheme-specific and wider (demographic, economic) data
- Report should
  - summarise content, quality, origin of statistical data
  - statistical annex on main elements of data base
  - include comments on quality and reliability of data
- Template for data request (very high level) – detail in appendix

Other issues and discussion

1. Why talk about data

- Work on data often major part of work in reviewing a social security arrangement

2. Variability of social security arrangements

- Can't have "one-size-fits-all" approach to data specification
- But can use check lists

3. In-house vs consultant actuaries (and ad hoc requests)

- How proactive can actuaries be with regards to data systems – liaison costs time and effort

Data sufficiency – what data do we need?

1. Starting point of valuation

- Data on liabilities and potential liabilities at valuation date
- Also assets and finances – but not main focus of this talk

2. Setting the assumptions

- Exposed to risk
- Data on movements between statuses (claims for benefit, ends of claims, end of liability for other reasons)

3. Validation of methodology

- Related to 1? Related to analysis of surplus

Data sufficiency – what data do we need?

Starting point of valuation – date issues

- As at valuation date (generally)
- But can this be produced later?
- Allowance for late postings/recording of items
- Adjustments?
- Age definition – as at date of valuation or date of extract
- Use year of birth instead

Data for social security valuations

Data sufficiency – what data do we need?

Starting point of valuation – grouped vs individual

- Social security arrangements are big (generally)
- Full data on individuals
  - capacity to transmit securely, analyse
  - is it complete?
- Grouped data
  - homogenous groups
  - TAS M
  - non-linearities
- Sample
  - unbiased?
  - sufficiently big?
  - rating-up factor?

Data sufficiency – what data do we need?

Starting point of valuation – other issues

- Options and guarantees (esp. if using grouped data)
- Accrued contributions – are all contributions equal (high-earnings sort service contributors vs low-earnings long servers)
- Need to perform operations on scheme data akin to benefit calculations
- Materiality

Data sufficiency – what data do we need?

Setting the decrement assumptions (movements between states)

- Number of movements
- Exposed-to-risk
- Credibility
- Multiple decrements
- Consistency of analysis of movements to set assumptions with projection methodology
- Materiality

Data sufficiency – what data do we need?

Setting the assumptions – exposed to risk

How measure

- Exactly? Is this possible from database?
- Average of "census data"?
  - similarity to data for starting point of valuation?
  - but discontinuities within year(s)?
- Include those who have subsequently exited/changed status

How much analysis of experience?

- Each year in inter-valuation period? Or longer?
  - Establish trends

Data for social security valuations

Data sufficiency – what data do we need?

Setting the assumptions – movements

Consistency with exposed to risk

- Including those who've subsequently changed status again in programme

Movements without financial effects – captured accurately?

Multiple sequential movements

Data sufficiency – what data do we need?

Setting the assumptions – other issues

- Credibility
  - May well affect even material aspects of arrangement
  - Grouping across data "cells"
  - Comparison with other rates
- Amounts vs lives
  - Hard to get data for "amounts" investigation?
  - Stratify lives by amount instead to capture same effect?
- Age definition (again)

Data sufficiency – what data do we need?

Summary

- Understand the arrangement
  - Cause of liabilities
  - How it can be modelled – "states" and "movements"
  - Comparison with other rates
- Materiality
  - Discuss with client
  - Try to anticipate elements of arrangement that can be omitted from modelling as non-material
- Engage client early on data issues (ideally < valuation date)
  - Try to understand administrative system and thus constraints
- Decide on grouped vs individual vs sample data (with limitations)
- Assess data credibility, decide on suitable experience analysis

Data for social security valuations

Data reliability – do data reflect arrangement's liabilities?

- Completeness
- Consistency with other data sources
  - Audited accounts
  - Last time's data
  - Other external data
- Internal consistency
- Common sense checks

Even if data were fine last time…

Data reliability – do data reflect arrangement's liabilities?

Data checking

Like checking anything…

- What do you (reasonably) expect
- Do data accord with this?
- If not, can differences be explained?

If can codify "what you expect"…

- Push data checking onto client?
- But still have to check client's checked data?

Data reliability – do data reflect arrangement's liabilities?

Completeness

- Hard (impossible?) in abstract
- See consistency with other data sources

Data reliability – do data reflect arrangement's liabilities?

Consistency with other data sources

- Audited accounts
  - receipts (in accounts) vs accruals (in database)
  - definitions and breakdowns
  - late postings and other date issues
  - easier for some membership or transaction types than others (where there's a financial effect)
- Last time's data (where available)
  - straight comparison, understanding of changes
  - analysis of surplus
- Other external data – national statistics, international data

Resultant rating up/down of results?

Data reliability – do data reflect arrangement's liabilities?

Internal consistency

- Internal consistency of each table of data/database
- Consistency between tables of data/database
  - For example – are movements from state A to state B in table of data/database for state A and table of data/database for state B
  - Amounts data and lives data consistent
  - Summary statistics

Data reliability – do data reflect arrangement's liabilities?

Common sense checks

- Misclassification of ages/dates of birth
- Extreme values (large/small)
- …
- Use of graphs to identify outliers?
- …
- You tell me

Data reliability – do data reflect arrangement's liabilities?

Summary

- Build on data specification
- Consider likely sources of error or problems
  - Discuss with client
- Codify what data being "right" (or at least "acceptable") means
- If can codify tightly, push data checking onto client
  - But still have to check that checks have been applied
- Check vs accounts
- Check vs last time's data (if available)
- Assess data credibility for experience analysis
- Disclose problems identified to client, in report

Data for social security valuations

What have I missed?