Inforce Data Compression Methods for Actuarial Modeling

Presented to 46th Actuarial Research Conference
University of Connecticut

August 13, 2011

Matthew Wininger, FSA, MAAA
Deloitte Consulting LLP

Question for consideration

Suppose you are projecting the number of future deaths for a set of fixed deferred annuities.
Your projection model has a group of 10,000,000 lives and a projection step of monthly for 50 years.
The model input data file is too large to run individually and you decide to combine your policyholder data by policyholder date of birth.

What is the optimal level of granularity to categorize DOB to balance runtime and accuracy?

Potential solutions

- Group all policyholders together in the same year of birth -> 1 category per birth year
- Group all policyholders together by quarter of birth -> 4 categories per birth year
- Group all policyholders together by month of birth -> 12 categories per birth year
- Group all policyholders together by week of birth -> 52 categories per birth year
- Group all policyholders together by day of birth -> 365 categories per birth year

How can we quantitatively evaluate the level of granularity if a seriatim run is not possible?

Why is understanding compression methods important?

Improve model runtime

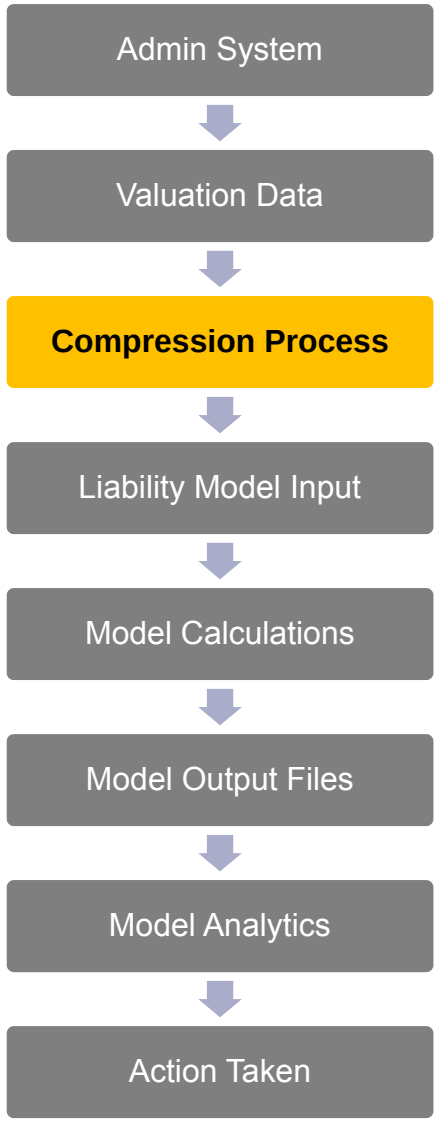
The compression process can be a source of error and/or efficiency in a model. If a user increases their compression ratio from 10x to 20x, they cut model runtime in half. When you have only four days to close your books, every hour counts.

Understand model attribution analysis

Changes in compression should be separately attributed when changing, refining, or updating models. Are they? How does a user attribute changes in the compression? Do users test appropriate alternatives?

Evaluate compression bias

It's helpful to be aware of the consolidation process to understand how it works to understand how the actuarial liabilities are reported. Have users recently evaluated the impact of compression on modeled results?



Cell compression terminology

Cell - An inforce model data point.

Seriatim - A set of cells without grouping, categorization, or remapping. One cell = one policy.

Grouping - A set of inforce data aggregated across certain elements defined by an algorithm. One cell has ≥1 policies.

Categorization - A process by which data elements are systematically and deliberately summarized to prepare for compression. Ex: Summarizing Issue Month into Issue Quarter.

Remapping - A data summarization technique whereby data elements are possibly altered. Ex: Products {A, B, C} are remapped to {A, C, C}.

Compression - Grouping process by which policies with similar characteristics are aggregated together, generally for actuarial modeling. Compression involves grouping, categorization, and/or remapping. A compression is done to reduce model runtime by reducing model points via similar groupings. A compression is defined by rules, formal or not.

Compression Bias - Model error due to inappropriate or excessive categorization or remapping. Ex: creates an unintentional *benefit of aggregation* which reduces model accuracy. Compression bias could overstate or understate results and may be nonlinear.

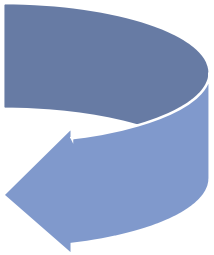
Compression Ratio - Average number of policies found in a cell. Higher compression ratio leads to model efficiency, at the possible cost of introducing compression bias. Ex: Depending on purpose a VA model could have a compression ratio between 10:1 and 2000:1.

Multiplier Effect - For each additional grouping selection utilized, this multiples the cell count by the number of elements in the group. Ex: if a model compresses policy to nearest issue year, and it is now desired to compress to nearest issue month, there will be 12 times as many cells. (This example assumes independence of variables.)

Cell compression example

Seriatim Data

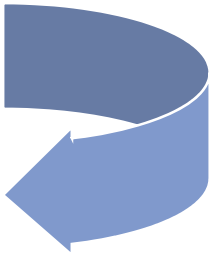
Policy Number	Product Type	Issue Month	Issue Year	NAR Ratio	Account Value
10000001	Victory	4	2005	113%	100,000
10000002	Pinnacle	5	2005	108%	50,000
10000003	Victory	6	2005	98%	75,000



- *Categorize Issue Quarter and NAR Band*

Categorized Inforce Data

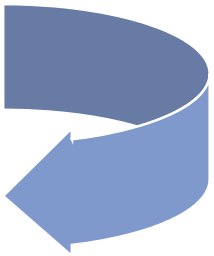
Policy Number	Product Type	Issue Quarter	Issue Year	NAR Band	Account Value
10000001	Victory	2	2005	1.05-1.15	100,000
10000002	Pinnacle	2	2005	1.05-1.15	50,000
10000003	Victory	2	2005	0.95-1.05	75,000



- *Remap Product Group*

Categorized and Remapped Inforce Data

Policy Number	Product Group	Issue Quarter	Issue Year	NAR Band	Account Value
10000001	Victory	2	2005	1.05-1.15	100,000
10000002	Victory	2	2005	1.05-1.15	50,000
10000003	Victory	2	2005	0.95-1.05	75,000



- *Compress by consolidating similar cells with matching grouping elements*

Compressed Inforce Data

Policy Count	Product Group	Issue Quarter	Issue Year	NAR Band	Sum of AV
2	Victory	2	2005	1.05-1.15	150,000
1	Victory	2	2005	0.95-1.05	75,000

Basic compression features

How are the compression calculations typically done?

- Excel via pivot tables, or
- In admin system directly via a subroutine, or
- In an Access or Oracle database

Simple variable annuity compression example

- SELECT FROM Current Month Valuation Data
- GROUP BY Issue Year, Net Amount at Risk (NAR) Band, Benefit Type, Attained Age Group
- SUM Policy Count, Policy AV, Gross Remaining Benefit (GRB), NAR$
- AVERAGE Attained Age Weighted by AV

Grouping vs. Calculation Elements

- Grouping. In this example they are Issue Year, NAR Band, Benefit Type, Attained Age Group
- Calculation. In this example they are Policy Count, Policy AV, GRB, NAR$ and Attained Age

Two ways to reduce model points

- First, use a simple “Group By” function. This reduces seriatim to a compression level with very little compression bias
- Second, introduce categorization and/or remapping. This changes the values of the grouping elements, and *begins to introduce compression bias.*

Basic compression features, continued

Is every policy uniquely assigned to a single cell?

- In simple compressions, yes
- Policy division may be required or desired
    - Depends on modeling purpose
    - Depends on product features
- Ex: fund regression calculations

Incremental evolution vs. generational

There may not be a formal process to adjust the compression. It could be done ad hoc, in reaction to a new product or modeling feature. It may be done only after a serious model error occurs.

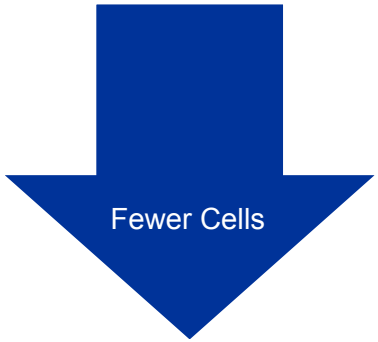
Compression Validations

- At minimum confirm the control totals for key calculation fields match before and after the compression process
- May indicate incorrect valuation data or erroneous calculations
- Possibly add filtering elements, ex: select only policies with AV > 0
- We'll discuss this in more depth later in the presentation

Top Level Adjustment

- Occasionally implemented as a way to overcome previously identified and quantified compression bias
- May be a linear adjustment to fix a non-linear issue
- Need to make sure the top-level adjustments are validated, documented, and refreshed appropriately

Compression tradeoffs and externalities

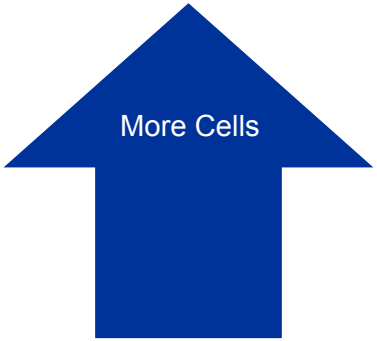


Reasons for More Compression

- Reduces model runtime; allows for more scenarios or faster results
- Control over infrastructure costs: hardware vs. software investment tradeoff
- May be required by model software or hardware constraints

Reasons for Less Compression

- Appropriate for high policyholder optionality
- Increased model accuracy in key scenarios
- Trace model results to policyholder cell drivers



Compression Externalities

Incorrect valuation data

Model calculation bias

Scenario selection bias

Analysis bias

Failure to understand or take appropriate action based on model results

Illustrative effect of compression on model results

Situation

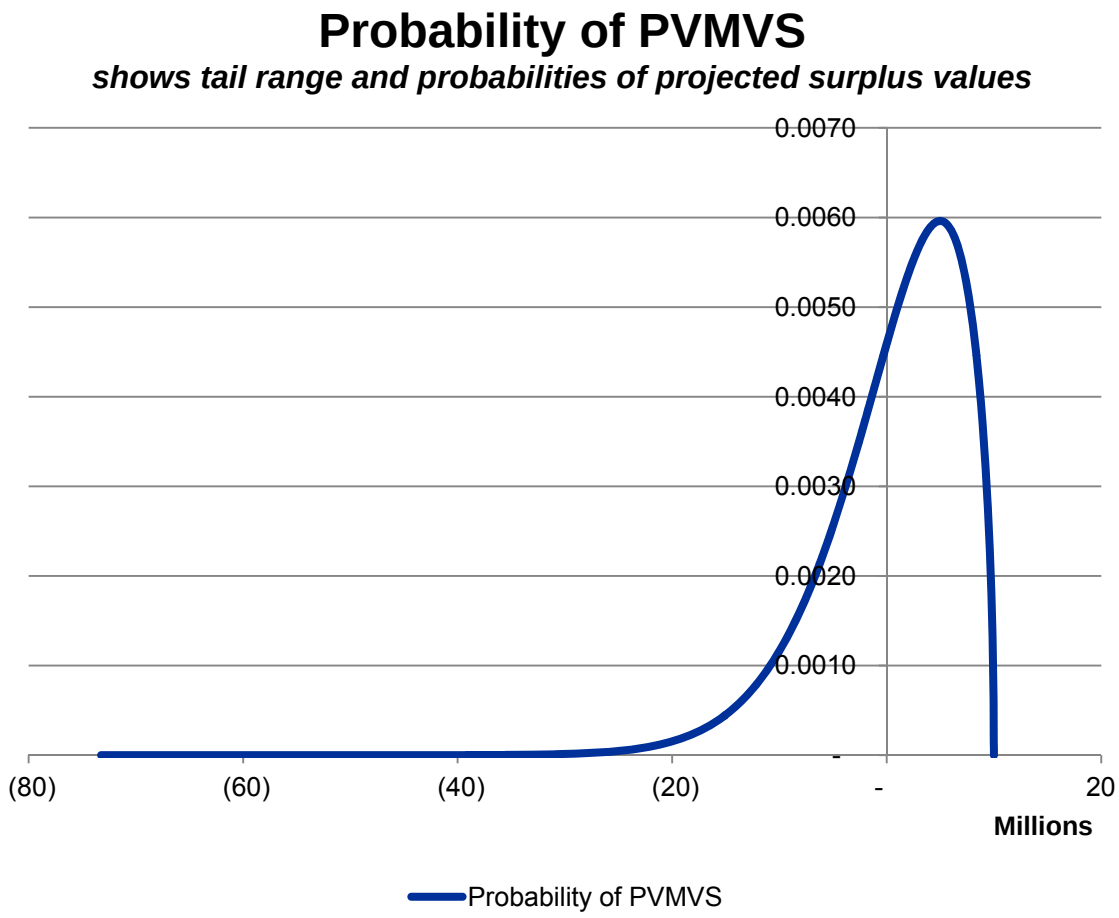
You have a generic asset adequacy analysis model, designed to calculate the present value market value of surplus (PVMVS).

Product and Risks

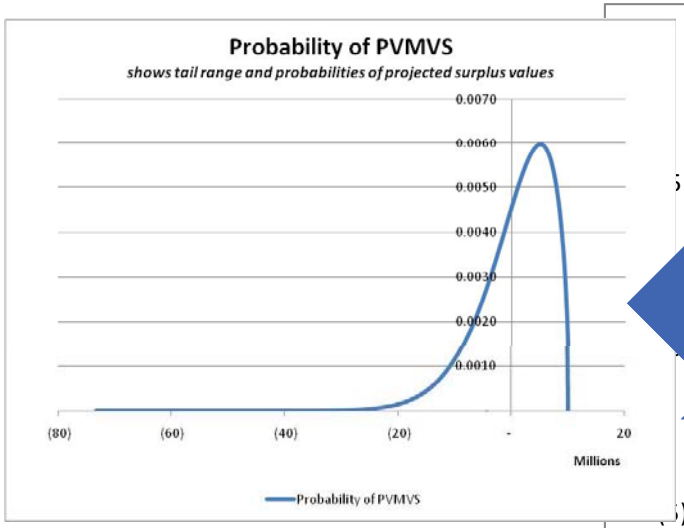
For illustrative purposes, the product and risks are not very important, just important that there is a distribution.

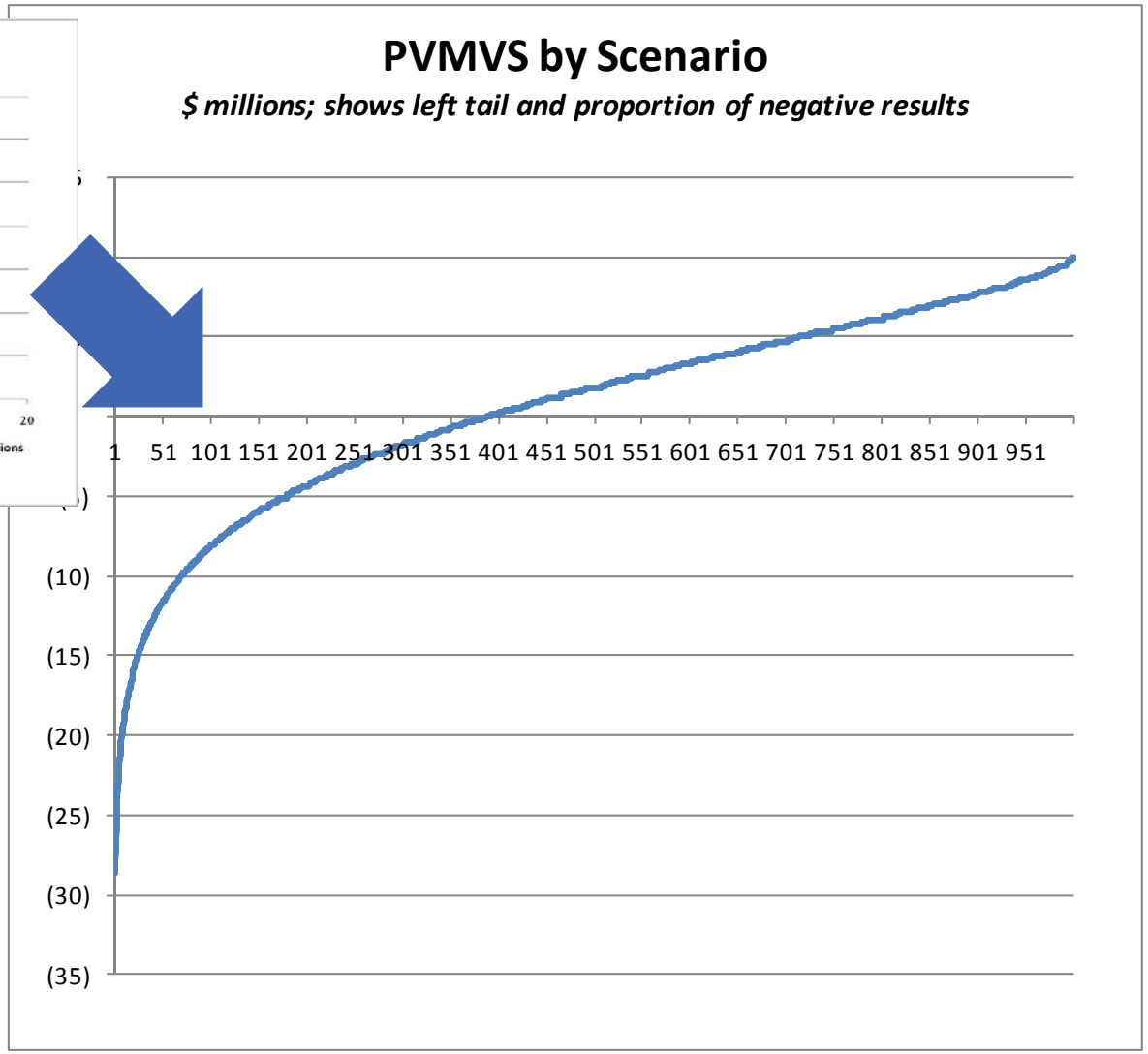
There is a positive expected value, an upper limit limited by premium collected; and a long left tail due to insured risks.

This illustrates the seriatim run across 1000 economic scenarios.



Translating the probability distribution to scenario results





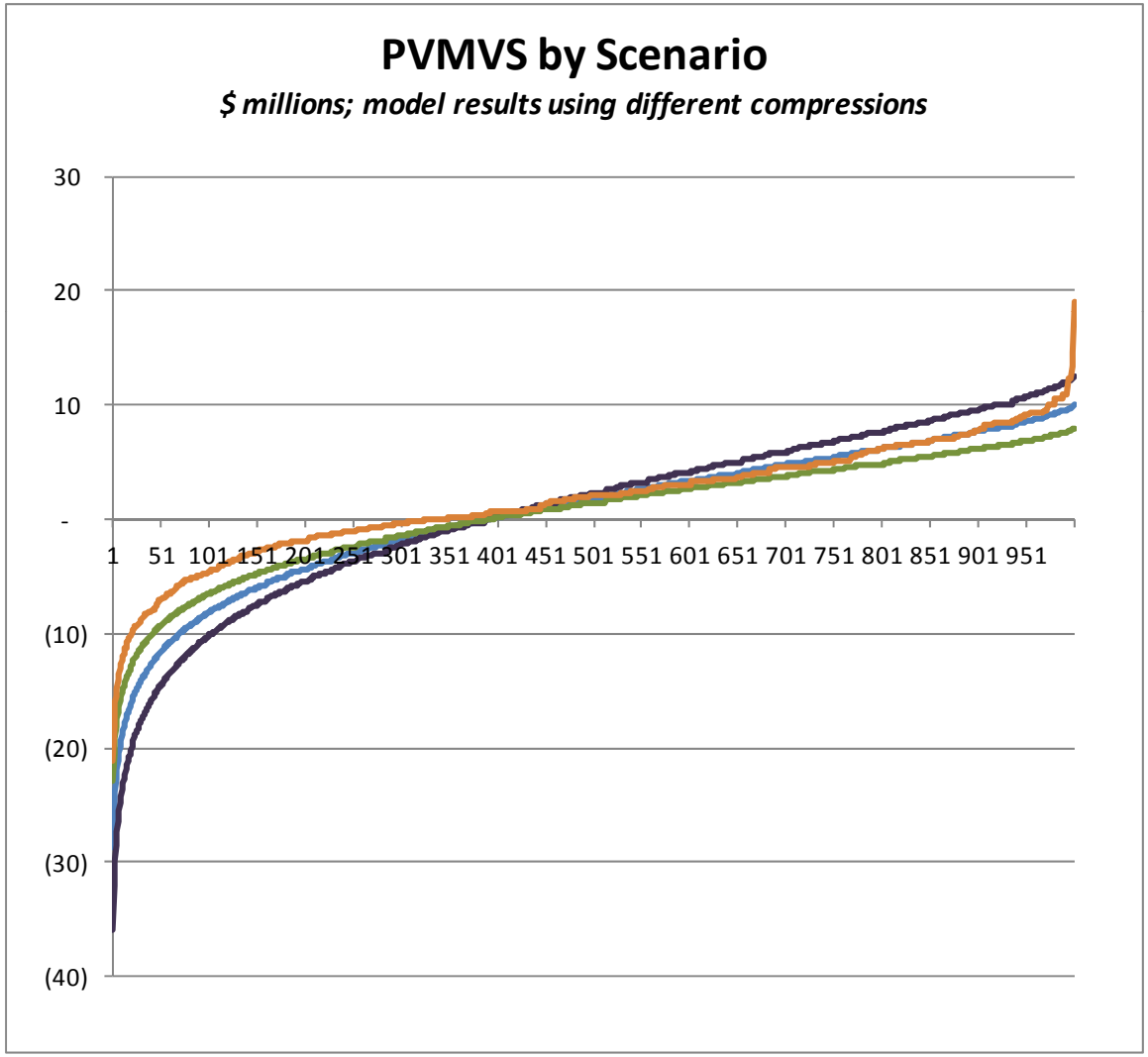
Scenario Results

The scenario results are ranked and displayed from smallest to largest PVMVS.

How do you know when a compression is good? Or good enough?

Which of these compressions is the best one, *if best is defined as least biased, or least biased given the runtime required to calculate it?*

It's not clear which compression is best by simply looking at the model output. A question often overlooked is: is *any* compression good enough?



Evaluating a compression quantitatively

This illustrates a typical compression test. Note the increasing pattern of compression bias with more compression or higher CTE value, compared to the baseline seriatim run.

(in $ millions)

CTE Value	Baseline	Compression A	Compression B	Compression C
50	(4.0)	(3.9)	(3.8)	(3.6)
65	(6.0)	(5.9)	(5.7)	(5.4)
70	(7.0)	(6.9)	(6.7)	(6.3)
80	(9.0)	(8.8)	(8.6)	(8.1)
90	(13.0)	(12.7)	(12.4)	(11.7)
Cell Count	**15,000**	**8,000**	**4,000**	**1,000**

Measuring compression bias (as a percent of the baseline):

CTE Value	Baseline	Compression A	Compression B	Compression C
50		-1.0%	-3.2%	-6.2%
65		-1.3%	-3.9%	-7.4%
70		-1.8%	-4.2%	-8.7%
80		-2.0%	-4.8%	-9.5%
90		-2.1%	-5.1%	-10.3%
Cell Count	**15,000**	**8,000**	**4,000**	**1,000**

This example illustrates there is no clearly optimal choice. In practice you may not have the information conveniently available to make this tradeoff decision.

Compression requirements and recommended practice

C3 Phase II Practice Note – 9/2006

Q4.2 What granularity of models is usually appropriate?
A: For large blocks of business, the actuary may choose to employ grouping methods to in-force seriatim data in order to improve model run times. The actuary normally uses enough model points that the VA RBC result would not materially change with additional model points (model cells). Grouping methods usually retain the characteristics required to model all material risks and options embedded in the liabilities. The actuary may wish to consider describing the degree of granularity chosen in the supporting memorandum.

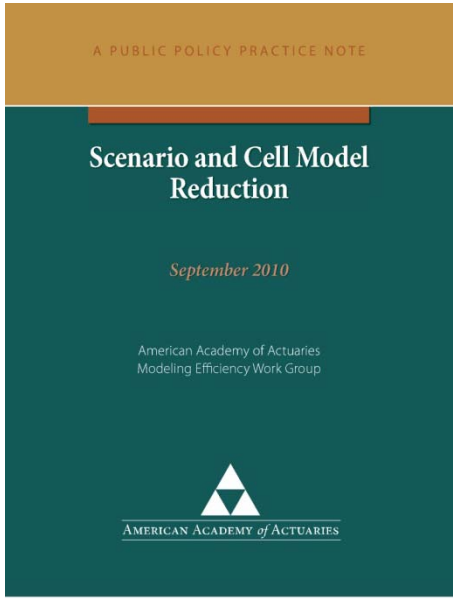
VACARVM Practice Note – 7/2009

Q4.2 What granularity of models is usually appropriate?
A: For large blocks of business, the actuary may choose to employ grouping methods to in-force seriatim data in order to improve model run times. The actuary should normally use enough model points such that results would not materially change with additional model points (model cells). Grouping methods usually retain the characteristics required to model all material risks and options embedded in the liabilities. AG 43 Section IV) D states that the Conditional Tail Expectation Amount at the option of the company may be determined by applying the methodology to subgroupings of contracts, Appendix 8 of AG 43 and Appendix 11 of C-3 Phase II both specify that the supporting memorandum should specify the grouping of contracts. The actuary may wish to consider describing in the supporting memorandum any testing performed to support the degree of granularity that has been used in the modeling of results.

Results

Compression is very much a judgment call. Disclosure of the high level method is required, but disclosure of the testing approach is not required.

9/2010 Modeling Efficiency Working Group Practice Note



This practice note is intended to provide information on common practices and approaches related to the use of reduced scenarios or reduced cell models for purposes of principle-based approaches to reserves and capital. Some of the concepts are **covered in this presentation**, and other concepts are not discussed in depth.

(1) Introduction

(2) Reduction Techniques
- Using a Reduced Scenario Set
- **Using a Reduced Cell Model**
- Using a Proxy for a Model of the Business
- Using a Reduced Scenario Set and a Reduced Cell Model
- Using a Reduced Scenario Set and a Reduced Cell Model, with Adjustment for Estimated Error

(3) Validating Results
- **Static and Dynamic Validation**
- **Validating with Reduced Cell or Reduced Scenario Sets**
- Validating as of an earlier projection date

- Ideally, a model is run using a "full set" of scenarios (a number of scenarios such that adding further scenarios would be very unlikely to materially affect results, "convergence")
- Because of practical constraints, an actuary may have to use a reduced scenario set intended to approximate the "full set."

- Sometimes a collection of non-insurance instruments may be used as a proxy for the cash flows from the liability or asset model as a whole.
- These instruments often feature cash flows or market values that can be determined for any economic scenario using closed form solutions (reduces model run-time).

- An actuary can take more time to validate a reduced run at an earlier valuation date.
- This method helps to alleviate time and resource constraints.
- Any differences between the results at the "test date" and the valuation date may be difficult to attribute if there were significant changes to either liability composition or market conditions.

Compression testing and other validation methods

Full seriatim categorization test

The most comprehensive method is to run each cell through the model individually. This method is the AG43 Standard Scenario test. This involves categorization, but not necessarily grouping or remapping. If this is possible, it's generally the best validation method.

Often this test is impossible, impractical, or undesirable:

- Many projection models have an effective upper limit on number of model cells.
- The calculation could take too long or generate too much output to store.
- Aggregate or dynamic modeling features may not work correctly; ex: reinsurance treaty modeling.

Should additionally test impact of grouping, then of remapping.

Point validations

A good substitute for a full seriatim categorization test is to chose a subset of cells or scenarios.

- Can run single cells as a categorization, or choose cells with one policy.
- Desirable to run several calibration scenarios of same cell.
- Desirable to run several cells through same scenario.
- Develop a fixed set of "test cells" which test common and extreme values.

An alternative approach is to run all cells through a subset of scenarios

- This subset should adequately model the tail and also the shape of the entire distribution

Static & dynamic validation

- This should be designed to reveal model biases, independent of the compression used.

Compression testing and validation methods, continued

Improve Until Good Enough

Test different compressions until the refinements don't result in any material output changes.

- Depends on definition of materiality.
- Must be sure to test "non local" solutions.

Remember, you may observe model biases independent of the compression.

When the behavior regime changes, do your bands?

Suppose in 2006 a company banded NAR ratio by the following groups:
{0-0.5, 0.5 – 0.8, 0.8 – 0.9, 0.9 – 1, 1-1.15, 1.15 – 1.25, 1.25+}

Then after the financial crisis the policyholders average NAR ratio increases to 1.2. The model must have redefined bands to account for new expectations of tail behavior.

Modeling the tail

The tail can refer to the model output tail – the worst scenarios by the key measures – or those cells which result in the worst model output. Reviewing tail values is important to understand what compression results trigger extreme behavior; then can calibrate your bands.

Scoring methods

May have a predefined evaluative criteria to select among different compressions.
Evaluations should be independent of model results.
May consider cell count; some sort of intraband measures.

Compression best practices checklist

- ❑ Compression algorithm is clearly documented and change history is maintained.
- ❑ Compression is validated by:
  - ❑ Control totals
  - ❑ Distribution checks for grouped elements
  - ❑ Distribution checks for remapped elements
- ❑ Changes to compression are appropriately tested using one or more of the following methods:
  - ❑ Seriatim categorization
  - ❑ Test cells
  - ❑ Test scenarios
  - ❑ Attribution tested on model results
- ❑ Static and dynamic validations are performed
- ❑ Tail scenarios are reviewed to understand sources/drivers
- ❑ Sources of compression bias on model results are understood, monitored, and adjusted if appropriate
- ❑ The degree of granularity and choices for grouping are supported by appropriateness testing, refreshed periodically.

Advanced compression features

Version control features
- Compression owner will track changes to the compression calculations
- Adds capability to reproduce prior compression results
- Log all the elements used, the qualitative method (sum, WA on AV, etc) and remapping rules

Default categorization feature
- Runs each policy into one cell without grouping but with remapping
- Facilitates compression validation and single-cell testing

Cell IDs with traceable inputs
- In simple compressions, it may be difficult or impossible to tell exactly which policies compose a cell
- This becomes more difficult if policies are subdivided across several cells
- An advanced compression will ‘tag’ each policy with an compression cell ID

Nonlinear banding / clustering
- Example of a linear banding: issue quarter
- What happens if 75% of your business was sold in 2Q and you require monthly projections?
- Might make sense to redefine issue date bands as: {1Q, April, May, June, 3Q, 4Q}

- Greatest granularity for bands with highest risk or modeling interest
- Helps better identify and model policyholder behavior in the tail

Advanced compression features, continued

Multi-stage compressions
- May apply different grouping and calculation rules sequentially
- Goal is to reduce the number of cells with few cell points
- *Generally model runtime is a function of cell count, not compression ratio*

Behavior review / prediction analysis
- Advanced compression technique where prior policyholder behavior is used to categorize
- Ex: Has the policyholder taken irregular partial withdrawals in past few years?
- Ex: Is this policy a lapse risk by some predefined criteria?

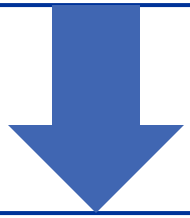
Asset compression methods
- Not widely used, yet.
- Asset call and prepayment schedules are generally unique and significantly influence market values.
- Asset diversity is generally greater than liability diversity for a given block.
- Simplistic asset compression may be appropriate if low invested asset balances, such as term life.
- Would not be appropriate for spread based insurance products.

Sampling Methods and Advanced Modeling Techniques
- An emerging actuarial practice

Question for open discussion

I'm not aware of any statistical tools which quantify compression bias over multiple output parameters.

What statistical tools can optimize the design of inforce data compression for a multi-scenario econometric projection?



This presentation used PVMVS as the single output variable by which compression bias was measured.

What statistical tools can optimize the compression design *when the econometric model has several output variables which are unevenly biased by compression*?