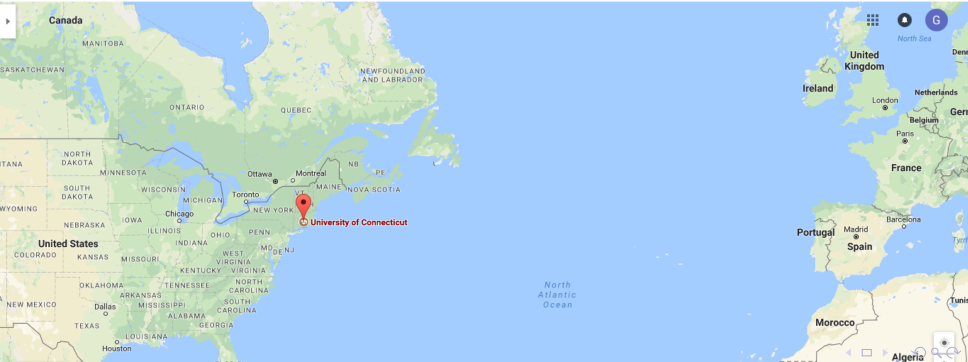
Valuation of Large Variable Annuity Portfolios: Challenges and Potential Solutions

Guojun Gan, PhD, ASA

Department of Mathematics, University of Connecticut, Storrs, CT, USA

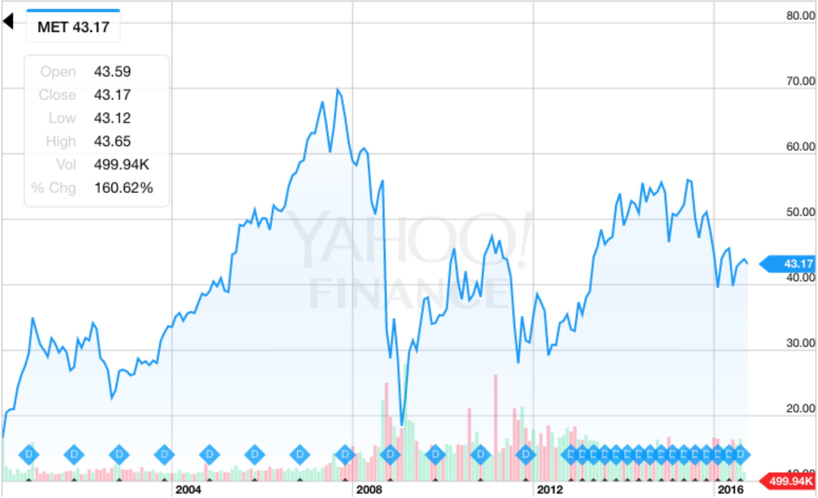
Workshop on Data Sciences applied to Insurance and Finance
Université catholique de Louvain, Belgium
September 15, 2017



Outline

- Computational problems from variable annuities (VA)
- Metamodeling approaches
- Some numerical examples

MetLife



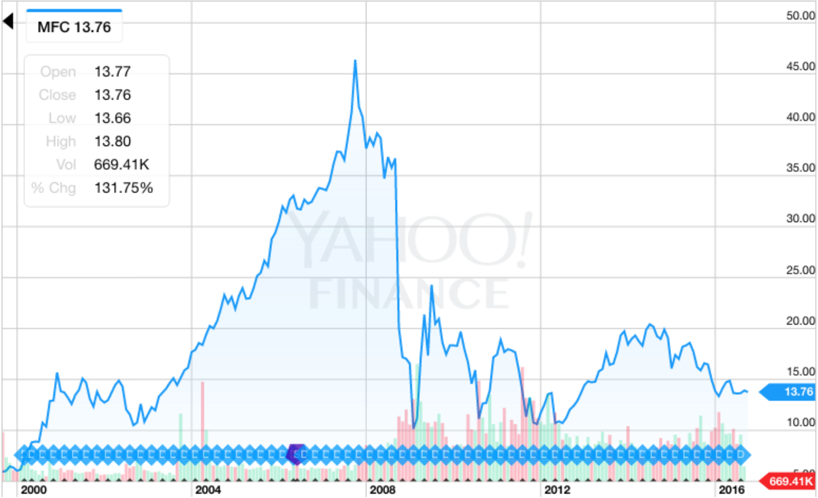
MET 43.17
Open 43.59
Close 43.17
Low 43.12
High 43.65
Vol 499.94K
% Chg 160.62%
80.00
70.00
60.00
50.00
43.17
40.00
30.00
20.00
499.94K
2004
2008
2012
2016

Prudential

PRU 77.78
Open 78.40
Close 77.78
Low 77.69
High 78.42
Vol 206.36K
% Chg 134.35%
110.00
100.00
90.00
80.00
77.78
70.00
60.00
50.00
40.00
30.00
20.00
206.36K
2004
2007
2010
2013
2016

The Hartford

Manulife Financial



Financial risks associated with variable annuities

- Traditional actuarial approaches cannot address these financial risks adequately (Hardy, 2000)
- Dynamic hedging is a popular approach to mitigate these financial risks in practice

Dynamic hedging requires calculating the Greeks of the portfolio

- Due to the complexity of the guarantees, there are in general no closed-form formulas for calculating the Greeks
- Insurance companies resort to Monte Carlo simulation

Monte Carlo simulation is computationally demanding for large portfolios of variable annuities

- Each policy needs to be projected with its own characteristics
- Long-term projection (e.g., 30 years) is involved
- Many economic scenarios (e.g., 1000 scenarios) are required.

(Dardis, 2016)

An example

- 100,000 policies
- 1,000 risk-neutral scenarios
- 360 time steps (monthly steps for 30 years)

The total number of cash flow projections for this portfolio is:

$$1,000 \times 12 \times 30 \times 100,000 = 3.6 \times 10^{10}.$$

Suppose that a computer can process 200,000 cash flow projections per second. Then it would take this computer

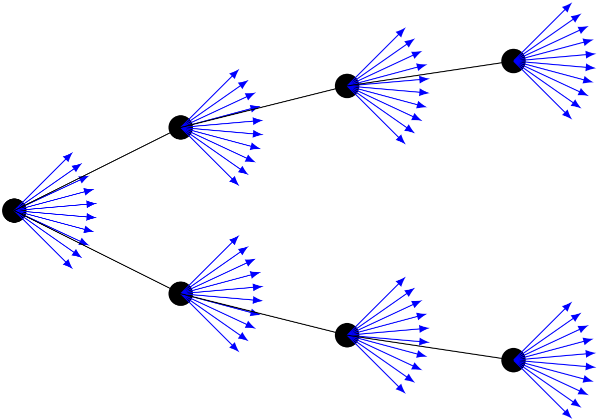
$$\frac{3.6 \times 10^{10} \text{ projections}}{200,000 \text{ projections/second}} = 50 \text{ hours}$$

to process all the cash flows of the portfolio.

Two computational problems related to dynamic hedging

- Calculate partial dollar deltas, partial dollar Rho, and dollar vegas for daily hedging purpose
  - Need to know the Greeks of the liability portfolio within a short time interval (e.g., a second) in order to rebalance the hedge position in a timely manner
- Reflect the dynamic hedging program in quarterly financial reporting
  - A stochastic-on-stochastic framework is usually used to simulate the dynamic hedging program

Nested simulation



An example of nested simulation

- 100,000 policies
- 1,000 inner loop risk neutral scenarios
- 1,000 outer loop real world scenarios
- 360 monthly time steps

$$\boxed{100,000 \times 1,000 \times 1,000 \times 360 \times 360/2 = 6.48 \times 10^{15}!}$$

$$\boxed{\frac{6.48 \times 10^{15} \text{ projections}}{200,000 \text{ projections/second}} \approx 1027 \text{ years!}}$$

(Reynolds and Man, 2008)

Metamodeling Approaches

A metamodel is a model of the Monte Carlo simulation model (Friedman, 2013) that can be used to replace the Monte Carlo simulation model to value the VA contracts in a large portfolio. A metamodeling approach involves four major steps (Barton, 2015):

1. select a small number of representative VA contracts from the portfolio;
2. run the Monte Carlo simulation model to calculate the fair market values of the selected VA contracts;
3. build a metamodel based on the selected VA contracts and the corresponding fair market values; and
4. use the metamodel to estimate the fair market values of all VA contracts in the portfolio.

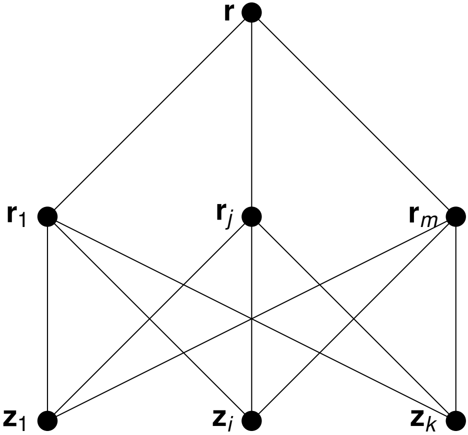
Experimental Design Methods

- Random sampling
- Data clustering
- Latin hypercube sampling
- Conditional Latin hypercube sampling

Metamodel Forms

- Kriging (Gan, 2013; Gan and Lin, 2015; Gan, 2015; Gan and Lin, 2016)
- GB2 regression models (Gan and Valdez, 2017)
- Neural networks (Hejazi and Jackson, 2016)

A two-level metamodeling approach was proposed to address the first computational problem



(Gan and Lin, 2016)

The level-one metamodel I

- Conditional Latin hypercube sampling
- Universal kriging

Under the universal kriging method, the partial dollar Delta of policy $\mathbf{x}_i$ on the hth tradable index when the market levle on the next day is $\mathbf{r}_l$ can be expressed as:

$$\hat{f}(\mathbf{x}_i, \mathbf{r}_l, h) = \boldsymbol{\lambda}_i^T f(Z, \mathbf{r}_l, h) = \sum_{j=1}^{k} \lambda_{ij} f(\mathbf{z}_j, \mathbf{r}_l, h), \quad (1)$$

where $\boldsymbol{\lambda}_i = (\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{ik})^T$ is a vector of kriging weights,

$$\begin{pmatrix} \mathbf{A}(Z) & \mathbf{B}(Z) \\ \mathbf{B}^T(Z) & \mathbf{0} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{\lambda}_i \\ \mathbf{v}_i \end{pmatrix} = \begin{pmatrix} \mathbf{A}(Z, \mathbf{x}_i) \\ \mathbf{B}(\mathbf{x}_i)^T \end{pmatrix}, \quad (2)$$

The level-one metamodel II

In Equation (2), $\mathbf{A}(Z)$ is a $k \times k$ matrix defined as

$$\mathbf{A}(Z) = \begin{pmatrix} \gamma(\mathbf{z}_1, \mathbf{z}_1) & \gamma(\mathbf{z}_1, \mathbf{z}_2) & \cdots & \gamma(\mathbf{z}_1, \mathbf{z}_k) \\ \gamma(\mathbf{z}_2, \mathbf{z}_1) & \gamma(\mathbf{z}_2, \mathbf{z}_2) & \cdots & \gamma(\mathbf{z}_2, \mathbf{z}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(\mathbf{z}_k, \mathbf{z}_1) & \gamma(\mathbf{z}_k, \mathbf{z}_2) & \cdots & \gamma(\mathbf{z}_k, \mathbf{z}_k) \end{pmatrix}, \tag{3}$$

where $\gamma(\cdot, \cdot)$ is the exponential semivariogram function defined as

$$\gamma(\mathbf{x}, \mathbf{y}) = 1 - \exp\left(-3\frac{\|\mathbf{x} - \mathbf{y}\|}{\beta}\right).$$

The matrix $\mathbf{B}(Z)$ is a $k \times (d+1)$ matrix defined as

$$\mathbf{B}(Z) = \begin{pmatrix} 1 & z_{11} & \cdots & z_{1d} \\ 1 & z_{21} & \cdots & z_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & z_{k1} & \cdots & z_{kd} \end{pmatrix}. \tag{4}$$

The level-one metamodel III

The column vector $\mathbf{A}(Z, \mathbf{x}_i)$ and the row vector $\mathbf{B}(\mathbf{x}_i)$ are defined as

$$\mathbf{A}(Z, \mathbf{x}_i) = \begin{pmatrix} \gamma(\mathbf{z}_1, \mathbf{x}_i) \\ \gamma(\mathbf{z}_2, \mathbf{x}_i) \\ \vdots \\ \gamma(\mathbf{z}_k, \mathbf{x}_i) \end{pmatrix} \tag{5}$$

and

$$\mathbf{B}(\mathbf{x}_i) = \begin{pmatrix} 1 & x_{i1} & \cdots & x_{id} \end{pmatrix}, \tag{6}$$

respectively.

The level-two metamodel I

- Latin hypercube sampling
- Ordinary kriging

Under the ordinary kriging method, the partial dollar Delta of the portfolio on the hth tradable index when the market level is $\mathbf{r}$ can be expressed as

$$\hat{g}(\mathbf{r}, h) = \sum_{l=1}^{m} w_l \cdot \hat{f}(X, \mathbf{r}_l, h), \tag{7}$$

where $\hat{f}(X, \mathbf{r}_l, h)$ is the partial dollar Delta of the portfolio on the hth tradable index when the market level is $\mathbf{r}_l$ and $w_1, w_2, \ldots, w_m$ are the kriging weights.

$$\begin{pmatrix} V_{11} & \cdots & V_{1m} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ V_{m1} & \cdots & V_{mm} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_m \\ \theta \end{pmatrix} = \begin{pmatrix} D_1 \\ \vdots \\ D_m \\ 1 \end{pmatrix}, \tag{8}$$

The level-two metamodel II

where θ is the Lagrange multiplier to ensure the sum of the kriging weights equal to one,

$$V_{ls} = 1 - \exp\left(-\frac{3}{\beta}\|\mathbf{r}_l - \mathbf{r}_s\|\right), \quad l, s = 1, 2, \ldots, m,$$

and

$$D_l = 1 - \exp\left(-\frac{3}{\beta}\|\mathbf{r} - \mathbf{r}_l\|\right), \quad j = 1, 2, \ldots, m.$$

A portfolio of synthetic variable annuity policies

Attribute	Values
Guarantee type	{DBRP, DBRU, WB, WBSU, MB}
Gender	{Male, Female}
Birth date range	[1/1/1950, 1/1/1980]
Issue date range	1/1/2000 1/1/2014]
Valuation date	1/1/2014
Maturity range	[15, 30]
Account value range	[50000, 500000]
Maturity	{10, 11, 12, . . ., 25}
Number of investment funds	10

All investment funds are mapped to five indices

Fund	US Large	US Small	Intl Equity	Fixed Income	Money Market
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	0	0	0	1	0
5	0	0	0	0	1
6	0.6	0.4	0	0	0
7	0.5	0	0.5	0	0
8	0.5	0	0	0.5	0
9	0	0.3	0.7	0	0
10	0.2	0.2	0.2	0.2	0.2

Some validation measures

$$\text{RMSE}(h) = \sqrt{\frac{1}{J}\sum_{l=1}^{J}(\hat{y}_{lh} - y_{lh})^2}, \tag{9}$$

$$\text{RAAE}(h) = \frac{\sum_{l=1}^{J}|\hat{y}_{lh} - y_{lh}|}{J \times \sigma_h}, \tag{10}$$

$$R_h^2 = 1 - \frac{\sum_{l=1}^{J}(\hat{y}_{lh} - y_{lh})^2}{\sum_{l=1}^{J}(\mu_h - y_{lh})^2} \tag{11}$$

$$\text{RMAE}(h) = \frac{\max_{1\le l\le J}|\hat{y}_{lh} - y_{lh}|}{\sigma_h}, \tag{12}$$

$$\text{APE}(h) = \frac{1}{J}\sum_{l=1}^{J}\frac{\hat{y}_{lh} - y_{lh}}{y_{lh}}, \tag{13}$$

$$\text{AAPE}(h) = \frac{1}{J}\sum_{l=1}^{J}\left|\frac{\hat{y}_{lh} - y_{lh}}{y_{lh}}\right|, \tag{14}$$

The metamodels in different levels are validated separately and jointly

To validate the level-one metamodel, we set

$$\hat{y}_{lh} = \hat{f}(X, \mathbf{r}_l, h), \quad y_{lh} = f(X, \mathbf{r}_l, h), \quad J = k,$$

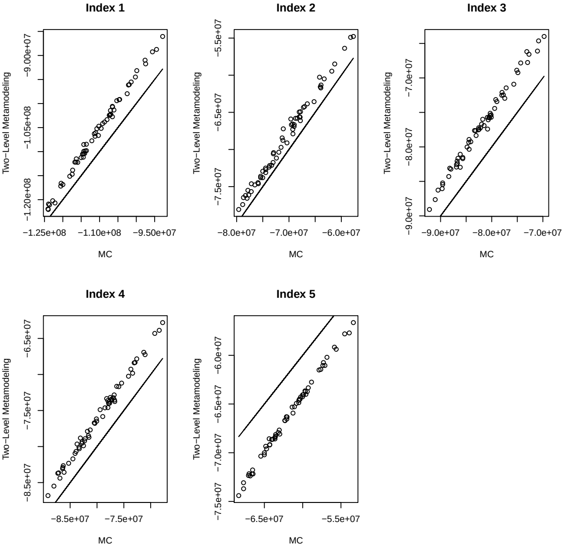
To validate the level-two metamodel alone, we set

$$\hat{y}_{lh} = g(\mathbf{s}_l, h), \quad y_{lh} = f(X, \mathbf{s}_l, h), \quad J = m,$$

To validate the two-level metamodeling approach with the level-one and level-two metamodels together, we set

$$\hat{y}_{lh} = \hat{g}(\mathbf{s}_l, h), \quad y_{lh} = f(X, \mathbf{s}_l, h), \quad J = m,$$

There are some biases when $k = 220$



Accuracy of the two-level metamodeling approach with $k = 220$ and $m = 50$.

	RMSE	RAAE	R-Squared	RMAE	APE	AAPE
Index 1	4,698,664	0.5811	0.6316	0.8675	-0.0419	0.0419
Index 2	2,302,026	0.4021	0.8188	0.7146	-0.0317	0.0317
Index 3	4,988,233	0.9079	0.1395	1.2809	-0.0609	0.0609
Index 4	3,746,782	0.7114	0.4688	0.9593	-0.047	0.047
Index 5	4,496,789	1.2115	-0.5247	1.6475	0.0717	0.0717

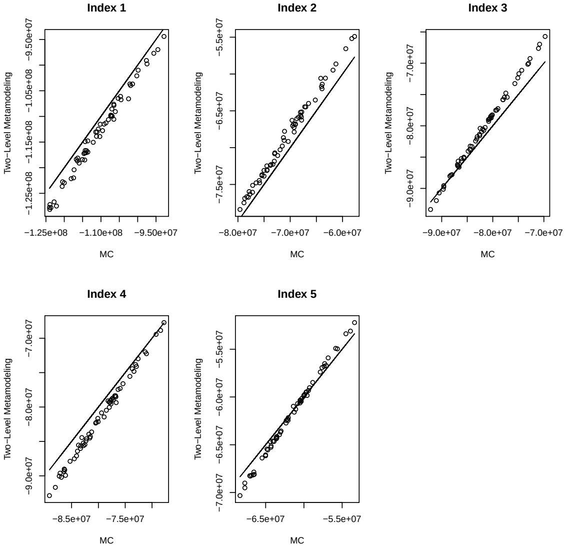
Accuracy of the level-one metamodel with $k = 220$ and $m = 50$.

	RMSE	RAAE	R-Squared	RMAE	APE	AAPE
Index 1	4,901,431	0.3274	0.8703	0.6525	-0.043	0.0431
Index 2	3,085,133	0.2763	0.8884	0.6799	-0.0335	0.0378
Index 3	5,076,421	0.4822	0.7307	0.9234	-0.0602	0.061
Index 4	3,814,150	0.3984	0.8175	0.6768	-0.0466	0.0466
Index 5	4,579,153	0.7101	0.4679	1.0208	0.0729	0.0729

Accuracy of the level-two metamodel with $k = 220$ and $m = 50$.

	RMSE	RAAE	R-Squared	RMAE	APE	AAPE
Index 1	632,799	0.066	0.9933	0.1923	-0.0012	0.0047
Index 2	428,126	0.0655	0.9937	0.1836	-0.0011	0.0051
Index 3	364,905	0.0515	0.9954	0.1664	-0.0015	0.0035
Index 4	390,177	0.06	0.9942	0.1712	-0.0014	0.004
Index 5	311,483	0.0694	0.9927	0.2025	-9.00E-04	0.0042

Biases reduced when $k = 440$



The accuracy of the two-level metamodeling approach when $k = 440$ and $m = 50$.

	RMSE	RAAE	R-Squared	RMAE	APE	AAPE
Index 1	2,781,176	0.3327	0.8709	0.6853	0.0233	0.0233
Index 2	2,124,997	0.3718	0.8456	0.6566	-0.0293	0.0293
Index 3	1,808,466	0.2776	0.8869	0.7427	-0.0182	0.0193
Index 4	2,142,709	0.3766	0.8263	0.7322	0.024	0.024
Index 5	803,153	0.1776	0.9514	0.5359	0.004	0.0104

Accuracy of the level-one metamodel when $k = 440$ and $m = 50$.

	RMSE	RAAE	R-Squared	RMAE	APE	AAPE
Index 1	3,733,350	0.226	0.9247	0.6062	0.0246	0.0283
Index 2	2,530,037	0.2294	0.925	0.5583	-0.0302	0.0316
Index 3	1,953,643	0.164	0.9601	0.3728	-0.0195	0.0218
Index 4	2,479,114	0.231	0.9229	0.6286	0.0246	0.0255
Index 5	803,141	0.1044	0.9836	0.3303	0.0036	0.0105

Accuracy of the level-two metamodel when $k = 440$ and $m = 50$.

	RMSE	RAAE	R-Squared	RMAE	APE	AAPE
Index 1	632,799	0.066	0.9933	0.1923	-0.0012	0.0047
Index 2	428,126	0.0655	0.9937	0.1836	-0.0011	0.0051
Index 3	364,905	0.0515	0.9954	0.1664	-0.0015	0.0035
Index 4	390,177	0.06	0.9942	0.1712	-0.0014	0.004
Index 5	311,483	0.0694	0.9927	0.2025	-9.00E-04	0.0042

Runtime of the two-level metamodeling approach

	$k = 220$, $m = 50$	$k = 440$, $m = 50$	Full, $m = 60$
`clhs`	138.75	169.44	NA
Monte Carlo	639.55	1,279.10	42,090.08
Level-One Metamodel	8.12	16.66	NA
`lhs`	0.02	0.02	NA
Level-Two Metamodel	0.19	0.33	NA
Total (seconds)	786.63	1,465.55	42,090.08
Total (hours)	0.22	0.41	11.69

Future work

- How to determine the sample size?
- Error estimation.
- Efficient experimental design methods.
- Many partial dollar Deltas are zero.
- Other applications (e.g., economic capital)

Please get in touch

Guojun Gan, PhD, ASA

Department of Mathematics

University of Connecticut

Storrs, CT

guojun.gan@uconn.edu

References I

Barton, R. R. (2015). Tutorial: Simulation metamodeling. In *Proceedings of the 2015 Winter Simulation Conference*, pages 1765–1779.

Dardis, T. (2016). Model efficiency in the U.S. life insurance industry. *The Modeling Platform*, (3):9–16.

Friedman, L. W. (2013). Simulation metamodeling. In Gass, S. and Fu, M., editors, *Encyclopedia of Operations Research and Management Science*, pages 1404–1410. Springer US.

Gan, G. (2013). Application of data clustering and machine learning in variable annuity valuation. *Insurance: Mathematics and Economics*, 53(3):795–801.

Gan, G. (2015). Application of metamodeling to the valuation of large variable annuity portfolios. In *Proceedings of the Winter Simulation Conference*, pages 1103–1114.

Gan, G. and Lin, S. (2015). Valuation of large portfolios of variable annuities under nested simulations: A functional data analysis approach. *Insurance: Mathematics and Economics*.

Gan, G. and Lin, X. S. (2016). Efficient greek calculation of variable annuity portfolios for dynamic hedging: A two-level metamodeling approach. *North American Actuarial Journal*, In Press.

Gan, G. and Valdez, E. A. (2017). Regression modeling for the valuation of large variable annuity portfolios. *North American Actuarial Journal*. In Press.

Hardy, M. (2000). Hedging and reserving for single premium segregated fund contracts. *North American Actuarial Journal*, 4(2):63–74.

References II

Hejazi, S. A. and Jackson, K. R. (2016). A neural network approach to efficient valuation of large portfolios of variable annuities. *Insurance: Mathematics and Economics*, 70:169 – 181.

Reynolds, C. and Man, S. (2008). Nested stochastic pricing: The time has come. *Product Matters! - Society of Actuaries*, 71:16–20.