2020 VIRTUAL ANNUAL MEETING & EXHIBIT

OCTOBER 26–29, 2020
Hedging Equity-Based Insurance Guarantees using Predictive Equity Analytics: --The Latest Innovation

Mark Zanecki  ASA, MAAA
President, IHA Consultants
2020 SOA Annual Meeting (virtual)
October 27, 2020  Session 5E (3:00 pm – 4:00 pm)
SOA Antitrust Compliance Guidelines

Active participation in the Society of Actuaries is an important aspect of membership. While the positive contributions of professional societies and associations are well-recognized and encouraged, association activities are vulnerable to close antitrust scrutiny. By their very nature, associations bring together industry competitors and other market participants.

The United States antitrust laws aim to protect consumers by preserving the free economy and prohibiting anti-competitive business practices; they promote competition. There are both state and federal antitrust laws, although state antitrust laws closely follow federal law. The Sherman Act, is the primary U.S. antitrust law pertaining to association activities. The Sherman Act prohibits every contract, combination or conspiracy that places an unreasonable restraint on trade. There are, however, some activities that are illegal under all circumstances, such as price fixing, market allocation and collusive bidding.

There is no safe harbor under the antitrust law for professional association activities. Therefore, association meeting participants should refrain from discussing any activity that could potentially be construed as having an anti-competitive effect. Discussions relating to product or service pricing, market allocations, membership restrictions, product standardization or other conditions on trade could arguably be perceived as a restraint on trade and may expose the SOA and its members to antitrust enforcement procedures.

While participating in all SOA in person meetings, webinars, teleconferences or side discussions, you should avoid discussing competitively sensitive information with competitors and follow these guidelines:

- Do not discuss prices for services or products or anything else that might affect prices
- Do not discuss what you or other entities plan to do in a particular geographic or product markets or with particular customers.
- Do not speak on behalf of the SOA or any of its committees unless specifically authorized to do so.
- Do leave a meeting where any anticompetitive pricing or market allocation discussion occurs.
- Do alert SOA staff and/or legal counsel to any concerning discussions
- Do consult with legal counsel before raising any matter or making a statement that may involve competitively sensitive information.

Adherence to these guidelines involves not only avoidance of antitrust violations, but avoidance of behavior which might be so construed. These guidelines only provide an overview of prohibited activities. SOA legal counsel reviews meeting agenda and materials as deemed appropriate and any discussion that departs from the formal agenda should be scrutinized carefully. Antitrust compliance is everyone’s responsibility; however, please seek legal counsel if you have any questions or concerns.
Presentation Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the participants individually and, unless expressly stated to the contrary, are not the opinion or position of the Society of Actuaries, its cosponsors or its committees. The Society of Actuaries does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented. Attendees should note that the sessions are audio-recorded and may be published in various media, including print, audio and video formats without further notice.
Agenda

1. Executive Summary

2. Overview of current stochastic PDE modeling of $S_t$ including sensitivities (“the Greeks”)

3. Insurer’s guarantee minimum benefits hedges are they reliable?

4. Two statistical models for $S_t$ behavior under filters $\{P, N\}$

5. Overview Predictive Equity Analytics: TRI-SIGNAL framework

6. Hedging segregated account guaranteed minimum benefits using TRI-SIGNAL equity prediction

7. Discussion and Conclusion

8. Questions

Executive Summary – Predictive Equity Analytics: Motivation

WARNING: There are two common reactions to the cutting-edge material you are about to see and hear:

1) (open-minded) a sense of surprise coupled with euphoria with some degree of trepidation. [It took me 2 months to accept and self-modify my graduate finance training.]

2) (closed minded) “it can’t be done”, no matter the supporting theory, facts and evidence as the new findings are not yet incorporated into textbooks or exam material. (White paper is available.)

The basic idea is simple and immensely powerful. If an investor had next trade-day equity price direction prediction capability that was sufficiently reliable, how would s/he invest?

Answer: purchase stocks predicted to increase, sell stocks predicted to decrease if owned and short stocks predicted to decrease if not owned. Diversification would be re-defined as “dynamic diversification” weighting allocations to the set of stocks expected to increase in price. Up-front costs of implementing pre-defined portfolios to capture gains and diversify risk are re-scoped to only those stock expected to increase in price with full investment or partial investment and the remainder in a side interest bearing account.
Executive Summary – Predictive Equity Analytics: *Left Tail Risk* and *Veracity/Reliability*

We propose a **new solution**, **predictive equity analytics**, that is able to avoid drawdowns and capture gains (long run and short run) thus satisfying investors' desire to achieve high returns with limited drawdown risk i.e. **“natural hedging”** – the account value is generally non-decreasing over time held primarily in a side account as cash deposit.

Predictive equity capability implies that **left-tail return distribution risk** has been eliminated, censored in some way or significantly reduced.

**Veracity/reliability of predictive capability accuracy can be measured in three ways:**
1. back testing over some historical period (**currently 85%+**),
2. going forward series of next trade day predictive analysis reports and
3. overview of underlying financial, economic, behavioral and statistical theory. (See whitepaper.)
Example: YTD Segregated Account Gross Return vs SP500 thru Aug. 10, 2020

Comparison of 2020 YTD Gross Returns for TRI-SIGNAL Predictive SP500 vs SP500 ending Aug. 10

Account is generally non decreasing step function over time e.g. “SELF HEDGING.”

At 9.5 elapsed months: 19%+

12 month forecast: [18 – 30%+]

2020 SP500 Predictive Trend ending Aug. 10, 153 Trade days, YTD= 19.3%, Mean: 1.01, Stddev: 0.020, Sharpe: 50.8, Skew: 1.58, Kurtosis: 1.81

2020 SP500 Daily Trend ending Aug. 10, YTD= 2.3%, Mean: 1.0005, Stddev: 0.026, Sharpe: 37.8, Skew: -0.44, Kurtosis: 5.27
Hedging Intuition - Segregated Account Value Desired Outcome

Segregated account value return implication from predictive capability is avoidance or mitigation of downside loss and capture of upside gain with the potential for compounding of investment returns.

- “The account value(t) >= account value(t=0) in probability almost surely.”
- Account value is generally a non-decreasing step function over time and in probability.

- The need and cost for conventional physical hedging is significantly reduced (not eliminated.)
- Liquidity of segregated accounts and general account is high for equity portion.
- Current policyholders are willing to pay “extra fees” for access to investment return beyond guarantee minimum benefit.
- Operationally use “pooled equity account for trading” with gain/loss allocation back to individual policyholder accounts is straight forward and minimizes transaction costs. (Little to no out-sourcing to financial intermediary, services will be internal to insurer.)
Overview of current stochastic PDE modeling of $S_t$ including sensitivities ("the Greeks")
Segregated account performance depends on policyholder elected mix of equity and fixed income investments which can change over time.

We will focus on equity component for this hedging presentation using SP500 index since fixed index annuity sales are stronger than variable annuity sales.
Conventional Hedging is based on following Taylor approximation:

\[
\Delta f(S, t, r, \sigma) = \frac{\partial f}{\partial S} \Delta S + \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial r} \Delta r + \frac{\partial f}{\partial \sigma} \Delta \sigma + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (\Delta S)^2 + \ldots
\]

Depending on which terms are being hedged, some of the different approaches are as follows:

<table>
<thead>
<tr>
<th>Hedge</th>
<th>Hedge Cost</th>
<th>Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delta hedge</td>
<td>$$</td>
<td>nominal</td>
</tr>
<tr>
<td>Delta, gamma hedge</td>
<td>$$$</td>
<td>advanced</td>
</tr>
<tr>
<td>Delta, gamma, rho hedge</td>
<td>$$$$</td>
<td>high</td>
</tr>
<tr>
<td>Delta, gamma, rho and vega hedge</td>
<td>$$$$$</td>
<td>higher yet</td>
</tr>
</tbody>
</table>

Taylor expansion is local approximation and requires normal market function. When markets fail: 2008, 2018, 2010 Taylor expansion fails as local approximation is no longer valid. “All hedge programs are good up until they fail.”

For elevated volatility levels some type of delta, gamma hedge is superior to delta only hedge.

Monitoring delta and hedging rho and vega are critical for hedge risk management. Rebalancing always causes gamma loss.

**Hedging using derivatives, options, swaps, options on variance and futures is complex, requires experience and absence of market-wide disturbances.**

There is no way to ensure “a fair price.” At best, you have a negotiated deal.

Mr. Yu Feng FSA from AegonUSA will discuss optimal hedging using neural nets.
How reliable are an insurer’s guarantee minimum benefits?
Insurer’s guarantee minimum benefits hedges are they reliable? GMxB Financial Estimates are Uncertain

Our situation is somewhat unsatisfactory from both actuarial and financial viewpoints. Important questions seemingly do not have sufficiently reliable answers despite our best efforts:

- Are the guarantee minimum benefits correctly bundled? Ans: Yes
- Are the guarantee minimum benefits correctly priced? Ans: ???
- Are the guarantee minimum benefits correctly reserved? Ans: ???
- Are the guarantee minimum benefits correctly hedged? Ans: ???

In order to answer the last three guarantee minimum benefit questions, we are going to need a point of view based on a fundamentally different modeling approach.

**Digression – thought experiment.**

For a set of stocks and some future time period:

- is it possible to identify the subset of stocks expected to increase?
- is it possible to identify the subset of stocks expected to decrease?
- is it possible to identify the subset of stocks expected to neither increase nor decrease?

The optimal trading/hedging strategy would be to buy the stocks expected to increase in value and sell the stocks expected to decrease in value. If no buy/sell opportunities are identified, then invest in interest-bearing account.
Intuition - Segregated Account Value Desired Outcome

Segregated account value return implication from predictive capability is avoidance or mitigation of downside loss and capture of upside gain with the potential for compounding of investment returns.

- “The account value(t) >= account value(t=0) in probability almost surely.”
- Account value is generally a non-decreasing step function over time and in probability.

- The need and cost for conventional physical hedging is significantly reduced (not eliminated.)
- Liquidity of segregated accounts and general account is high for equity portion.
- Current policyholders are willing to pay “extra fees” for access to investment return beyond guarantee minimum benefit.
- Operationally use “pooled equity account for trading” with gain/loss allocation back to individual policyholder accounts is straight forward and minimizes transaction costs.
Two statistical models for $S_t$ behavior under filters $\{P,N\}$
Two statistical models for $S_t$ behavior under filters $\{P, N\}$

We break from standard theory at this point and hypothesize that rather than one statistical model, we have two statistical models for $S_t$ behavior:

- Model 1 behavior specifies positive trend filtering, $P$.
- Model 2 behavior specifies negative trend filtering, $N$.

**Model 1- Overall Positive Trend:**

$$
\frac{dS_t^P}{S_t} = (\mu_t^P - q)dt + \sigma_t^P dW_t^P \geq 0 \text{ almost surely in mean-variance for any combination of } \{ (\mu_t^P - q), \sigma_t^P dW_t^P \} \text{ such that } \frac{dS_t^P}{S_t} \geq 0.
$$

*where more narrow assumptions are:*

- $dW_t^P \geq 0$ is a censored normal random variable at 0 and $\mu_t^P = E_t[dS_t^P] \geq \max(0, q)$,
- $\sigma_t^P \geq 0, q \geq 0$.

Indicator positive trend model is commencing or is active:

$I_t^P$ with probability $0 \leq p(I_t^P) \leq 1$

$q$ is continuous margin offset.
Model 2- Overall Negative Trend:

\[
\frac{dS_t^N}{S_t} = (\mu_t^N - q)dt + \sigma_t^N dW_t^N \leq 0 \text{ almost surely in mean-variance for any combination of } \{(\mu_t^N - q), \sigma_t^N dW_t^N\} \text{ such that } \frac{dS_t^N}{S_t} \leq 0.
\]

where more narrow assumptions are:
\[dW_t^N \leq 0 \text{ is a censored normal random variable at } 0 \text{ and }\]
\[
\mu_t^N = E_t[dS_t^N] \leq \max(0, q),
\]
\[
\sigma_t^N \geq 0, q \geq 0
\]

Indicator negative trend model is commencing or is active:
\[I_t^N \text{ with probability } 0 \leq p(I_t^N) \leq 1\]

Given filters \{P, N\}, mean-variance dominance of returns follows.
Overall horizontal \(S_t\) behavior is the compliment to \{P, N\} with probability \(1 - p(I_t^P) - p(I_t^N)\). For horizontal behavior, an investor is better off investing in an interest-bearing account.

Existence and Sufficiency of \{P,N\} Filtering of \(S_t\)
Construction of \{P,N\} filters is out of scope. Existence and sufficiency of \{P,N\} are satisfied by providing historical TRI-SIGNAL results and measuring predictive accuracy. Alternatively, a forward-looking accuracy verification is also available. It would be “impossible” to accurately predict the behavior of 4,500+ stock prices over 500 trade-days without an accurate and reliable modeling framework which proves existence and sufficiency of \{P,N\}. TRI-SIGNAL predictive accuracy is 85%+.
Demonstration - Segregated Account Value Linked to SP500

Sample of 332 stocks from SP500 as demonstration: view of first 250, alphabetical order with 121 trade-day period ending July 6, 2020.

Rows contain individual stock series. Columns contain market close prices.

How you would invest? Ans: with no predictive capability use standard diversified portfolio in form of ETF for lowest cost, mutual fund, SPDR or possibly some other active strategy using momentum, factors, sustainable(ESG) etc.
**TRI-SIGNAL Predictive View** of first 250, SP500 stocks with 121 trade-day period ending July 6, 2020 with in form of heat map highlighting.

This snapshot is first 250 stocks arranged by symbol alphabetically. View by sector and by symbol within sector is also available.

**Cell Highlight Color coding:**
1. Short run buy: light green
2. Long run and short run simultaneous buy: dark green
3. Short run sell: red
4. Monitor / No action: yellow

An all-encompassing description of the figure would be characterized as “random” in agreement with traditional financial theory. For any individual stock, “random or nearly random.” This is similar to internet https encryption of messages; messages appear to be random but in reality, they are not. The true message is recoverable with the proper “prime number key” and algorithm.

SP500 index drawdown 2020 warnings:
2/24, 6/9, 6/11.

No matter the view, visually you can see patterns (local non-randomness) across rows and down columns.

Cause: algorithmic trading.

Our hedging / trading strategy is self-evident:
- Purchase stock indications: columns with “more green” or rows with “more green.”
- Sell stock indications: columns with “more red” or rows with “more red.”
Overview TRI-SIGNAL Framework and Robustness
TRI-SIGNAL – any stock price series can be decomposed into three distinct signals that are predictive.

The level and trend of Series_direction_signal_3 predicts trend increase and decrease.

Series_direction_signal_1 predicts indication of buy or sell opportunity.

Whenever series_direction_signal_3 is at a “high level relative to average” or when trend in series_direction_signal_3 is positive, the realized stock price trend is generally positive.

Whenever series_direction_signal_1 is 0.6 or higher an investment opportunity analysis event has occurred.

Prediction is accomplished by using combinations of series_direction_signal_1 and series_direction_signal_3. TRI-SIGNAL predictive accuracy is 85%+.

Breaking news events are not factored into the signals if they occur after market close and prior to market open.
Example of Equity Price Direction Prediction – Rule Set

Visual Signal Pattern for Next Trade-day Stock Price Direction**:
Signal 1 > = 0.6 is predictive for local peaks and inflection points.
Signal 3 = 0.97 or trending up indicates increase.
Signal 3 < 0.97 or trending down indicates decrease.
Signal 1 > = 0.6 and (Signal 3 > = 0.97 or Signal 3 trending up from 0.95) is predictive for price increase with accuracy 85%+.
Signal 1 > = 0.6 and (Signal 3 < 0.97 or Signal 3 trending down from 0.97) is predictive for price decrease with accuracy 85%+.

Note: 2020 Spring and 2018 Fall market downturns are highlighted in light red.
**Some stocks require signal 3 threshold adjustment for accuracy 85%+, e.g. Nvidia (NVDA) threshold is 0.95 rather than 0.97.

Stock List: AIGLENT TECHNOLOGIES (A)  APPLE (AAPL)  JPMORGAN (JPM)  NVIDIA (NVDA)  UNION PACIFIC RAILROAD (UNP)
Hedging Segregated Account Guaranteed Minimum Benefits using TRI-SIGNAL Equity Prediction

- Illustration: Segregated Account Investment Return for Policy starting January 2, 2020 thru July 2, 2020 using SP500
YTD Segregated Account Gross Return vs SP500 thru Aug. 10, 2020

Comparison of 2020 YTD Gross Returns for TRI-SIGNAL Predictive SP500 vs SP500 ending Aug. 10

At 9.5 elapsed months: 19%+. 12 month forecast: [18 – 30%+]

Account is generally non decreasing step function over time.

2020 SP500 Predictive Trend ending Aug. 10, 15 Trade days, YTD= 19.3%, Mean: 1.01, Stddev: 0.020, Sharpe: 50.8, Skew: 1.58, Kurtosis: 1.81

2020 SP500 Daily Trend ending Aug. 10, 153 Trade days, YTD= 2.3%, Mean: 1.0005, Stddev: 0.026, Sharpe: 37.8, Skew: -0.44, Kurtosis: 5.27

Left-tail filtering.
Policyholder Segregated Account Value Progression

Policyholder Sub-Account Value at end of each sub-period is given by:

$$F_k = F_0 \frac{S_k}{S_0} \left(1 - \frac{m}{n}\right)^k \quad k = 1, 2, \ldots, nT$$

with $n$ evaluations per year, $m$ annual charge amount.

The account value goes down with deterministic fee deduction over time and fluctuates with underlying reference portfolio.

Using TRI-SIGNAL applied to SP500 to avoid downturns and capture short-run/long-run gains:

$$F_k = F_0 \times \frac{Cum_{TRI\_SIGNAL\_SP500\_Return}}{n} \left(1 - \frac{m}{n}\right)^k \quad k = 1, 2, \ldots, nT$$

If TRI-SIGNAL return $\geq$ minimum guarantee, reserve liability only for future performance is required at much lower net risk level than standard model.
**Hedging Guarantee Minimum Benefit Modeling Probabilities**

Guaranteed minimum death benefit
Guaranteed minimum withdrawal benefit
Guaranteed minimum income benefit
Guaranteed minimum accumulation benefit

Highest probability

Lowest probability

**Highest hedging challenge**

**TRI-SIGNAL prediction has significantly removed SP500 downside volatility effect and captured gain opportunities – avoided downturn from mid February to end of April.**

**Hedging consisted of physical assets within SP500 at opportunistic moments which resulted in substantially lower hedging costs compared to any other hedging strategy.**

Return of premium guarantee was maintained and surpassed thru July 2 with extra margin for a roll-up as well as insurer profit.

**TRI-SIGNAL demonstration over 6 months implies that similar performance would be expected for hedging guaranteed minimum accumulation benefits over durations of any length.**
Discussion and Conclusion
Discussion

- By using both TRI-SIGNAL and TRI-SIGNAL DEEP ANALYTIC report sets, it is possible to boost accuracy greater than 90% with proper training and experience interpreting graphs combined with human inference.

- TRI-SIGNAL heatmap gives short run and long run quantification to determine if a specific investment strategy is working.

- TRI-SIGNAL combined with research and an investment strategy is the optimal trade monitoring and investment analytic solution.

- Pros:
  - Segregated account volatility is replaced with generally non-decreasing step function which reduces need for standard hedging and has high liquidity.
  - \( \{P, N\} \) filtering of \( S_t \) is easy to understand and implement in automated fashion.
  - Hedging costs are significant reduced versus standard hedging techniques.
  - Dynamic diversification - purchasing subset of stocks expected to increase is less costly and has higher expected return than static diversification with weighting.
  - Consistently surpasses conventional benchmark returns.

- Cons:
  - Requires daily monitoring and trade execution orders. “It is not set it and forget it.”
Conclusion

• Machine learning enables upstream filtering of $S_t$ to create TRI-SIGNAL physical hedge which consists of $\{S_t^P\}$ and an interest bearing account. The need and cost for standard hedging are both reduced.

• $S_t$ behavior under $\{P,N\}$ filtering significantly removes downside volatility enabling generally non deceasing step function for segregated account value over any horizon.

• For GMxB liabilities that can be viewed as some type of look-back put option, TRI-SIGNALs generally non decreasing segregated account value mitigates the need for hedging.

• Segregated accounts total assets under management behavior fundamentally changes which results in greater policyholder retention and expense recoupment.

• Financial reporting using prescribed standard methods to determine hedging levels exceed TRI-SIGNAL physical hedging results which dynamically remove / mitigate downside volatility.

• TRI-SIGNAL prediction combined with equity research is the optimal hedging framework.
Thank you for your time and interest.

Questions?
Appendix: TRI-SIGNAL Modeling Framework Robustness

TRI-SIGNAL Modeling Framework Diagnostics: DEEP ANALYTIC Reports

TRI-SIGNAL modeling performance is reported by 4 independent diagnostics using graphs in conjunction with raw accuracy measures from 121 trade-day heat map:

1. **Linear regression**: concurrence of predictive power, correlation > 0.80
2. **Time series**: plot of realized daily trend versus proxy long run mean track nicely.
3. **Mean-variance**: two-dimensional movement in expected return vs standard deviation vs realized return demonstrates concurrence or coming correction.
4. **Distributional Evidence**: distribution of proxy of long run mean is “similar” to distribution of realized daily trend.

See Figures below.
Agilent Technologies Inc. (A) - 121 Trade-day Regression: Realized Daily Trend (Blue dots) vs Proxy Long-run Mean Daily Trend (Orange dots) Prior to Market Open 7/6/2020 with Correlation: 0.82
Agilent Technologies Inc. (A) - 121 Trade-day Time Series:
Realized Daily Trend vs Proxy True Mean Trend Prior to Market
Open 7/6/2020
Agilent Technologies Inc. (A) - 42 Trade-day Histogram: Proxy True Mean Trend (Mean: 1.003545, StdDev: 1.543499E-02, Kurtosis: 1.387734, Skewness: -0.1753182) Prior to Market Open 7/6/2020

Agilent Technologies Inc. (A) - 42 Trade-day Histogram: Realized Daily Trend (Mean: 1.004276, StdDev: 1.678785E-02, Kurtosis: 3.60784, Skewness: -0.6675908) Prior to Market Open 7/6/2020
Disclaimer

The information provided is for educational purposes only and should not be construed as tax, legal or financial advice or guidance.

This presentation, data, information and all other contents or materials contained herein or attached hereto (collectively, the "Materials") were intended and prepared for general informational purposes only and should not be construed as advice or opinions on any specific facts or circumstances.

These Materials are made available on an "as is" basis without any warranty of any kind (express or implied), including without limitation respect to the accuracy, completeness, timeliness, or sufficiency of these Materials.

IHA Consultants Inc. directors, officers, employees, and representatives (collectively, “Representatives”) disclaim any liability to any person or organization for loss or damage caused by or resulting from any actions based on or reliance placed on these Materials.

In any event, IHA Consultants is to be held harmless for any special, indirect, consequential, or punitive loss or damage of any kind arising from or relating to these Materials.

Viewers of these Materials are advised to undertake an independent review of any legal, accounting, investment advice, taxation, regulatory, and actuarial implications of any transaction described herein.
Hedging with Neural Network

Yu Feng, FSA, CFA
Transamerica Life Insurance Co.

10/27/2000
Agenda

1. A brief introduction to neural network
2. Hedging to fair value with neural network
3. Hedging to stat with neural network
A brief introduction to neural network
History of neural network

- Neural Network research goes way back
- Only became particularly useful during last decade
  - Image recognition
  - Natural language processing
- Why now
  - Faster hardware: GPU, TPU, Neural Engine
  - Better software: improved network architecture, new activation functions, robust optimizer, modern software framework (tensorflow, pytorch etc.)
  - Bigger data to work with
  - Active community support, critical mass of interest

<table>
<thead>
<tr>
<th>50’</th>
<th>perceptron</th>
</tr>
</thead>
<tbody>
<tr>
<td>mid 70’</td>
<td>backpropagation</td>
</tr>
<tr>
<td>mid 90’</td>
<td>convolutional neural network</td>
</tr>
<tr>
<td>2009</td>
<td>ImageNet</td>
</tr>
<tr>
<td>2015</td>
<td>ResNet</td>
</tr>
<tr>
<td>2016</td>
<td>AlphaGo</td>
</tr>
</tbody>
</table>
How to train neural network

- Large labeled dataset
- Keep adjusting weights until desired outputs
- Gradient based weight adjustments

Forward pass - calculate loss

Backpropagation - adjust weights

Sigmoid activation function
Hedging to fair value with neural network
Future greeks

• Stochastic modeling is commonly used by actuaries
• Traditionally, nested stochastics is needed when calculating value/greeks for future node
• Nested stochastics with outer loop/inner loop setup is extremely computational intensive
• Least Square Monte Carlo. Reduce inner loops size by curve fitting.
• Some attempt to use neural network for the fitting
Train NN to produce Black-Scholes

- An algorithm to train neural networks to discover the future greeks of financial options
- The inputs of the process are:
  1. One set of economic scenarios. Could be real world. Shocked scenarios are NOT needed
  2. Option cash flow associated with each scenario
  3. That’s it. We do NOT need any prior knowledge of Black Scholes formula.

- The output of the process is a trained neural network, with times and index levels as inputs, delta as output.
- The training setup is original. We do not have a target output for neural network itself
- The training target is at batch level, where delta neural network is applied multiple times
  - Similar after hedge g/l among all scenarios (highest hedging effectiveness)!
Training setup, an example

- One year (252 days) at the money European call option
- 2% interest rate
- 4096 scenarios, daily time step, 0% drift and 16% volatility
  - The drift rate is different from risk free rate
- The delta network has
  - two inputs, time and index level
  - two hidden layers (16 nodes and 8 nodes, tanh activation)
  - One output, sigmoid activation, which will be trained as delta
- The training target is set up at batch level
  - Deltas are calculated with delta network applied at each time step of each scenario
  - G/L of delta hedge is calculated
  - Then calculated the after hedge cost of the options = delta hedge G/L + payout
  - The loss function is the variance of after hedge cost
  - The average after hedge cost is the risk neutral price at time zero

Training result (1)

- The program is developed with pytorch framework, using AdamW and LBFGS optimizer
- Only takes seconds to train on google Colab
- The neural network delta matches Black Scholes formula extremely well.
- The neural network delta also extrapolates well
- The neural network independently rediscovered Black-Scholes formula!
### Training result (2)

<table>
<thead>
<tr>
<th></th>
<th>Before Hedge</th>
<th>Hedged with NN</th>
<th>Hedged with BS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Std</td>
<td>Mean</td>
</tr>
<tr>
<td>Training scenarios</td>
<td>6.258%</td>
<td>10.053%</td>
<td>7.353%</td>
</tr>
<tr>
<td>Validation scenarios</td>
<td>6.237%</td>
<td>10.014%</td>
<td>7.353%</td>
</tr>
<tr>
<td>risk neutral scenarios</td>
<td>7.329%</td>
<td>10.900%</td>
<td>7.350%</td>
</tr>
</tbody>
</table>

- Hedge effectiveness with neural network is actually slightly better compared to Black-Scholes, even for out of sample scenarios.
- The after hedge mean is very close to the Black-Scholes formula price of 7.352%.
- We can now calculate future delta for a real world scenario set.
Hedging to stat with neural network
Liability setup

- 25 year GMWB like option
- 100% fund invested in S&P 500, mapped to US Diversified Equity for stat calculations
- 2.5% risk free rate
- 5 year accumulation phase. Withdrawal base rolls up at 3.5% per year, with monthly ratchet
- 20 year withdrawal phase. Monthly withdrawal at 4.5% of WB per year.
- 1.5% annual fee on account value, charged monthly

- We will explore several risk management methods
  - Unhedged
  - Hedge to CTE98 directly
  - Explicit hedge with neural network
Test Scenarios and unhedged result

- S&P 500 total return from June 1995 to June 2020
- Others are picked from AAA scenario, based on $\text{AVG}(\text{US yr5-25})/\text{HWM}(\text{US yr0-5})$
- Total Asset Requirement=$\max(\text{CTE98}, 0)$
TAR(CTE98) hedge

- Calculate sensitivity of max(CTE98,0) to index and delta hedge monthly
- Accumulated distributable earning profile is mostly smooth over time. Some occasional drawdown
- Still a lot variation among scenarios in final outcome
- Delta curve has a big range and jumps a lot
Hedge with neural network

- 8K independently generated training scenarios based on AAA logic
- Only trained once at time 0
- Same training target: minimize stdev of pv(after hedge cashflow)
- Neural network now has 3 inputs: AV, WB and time to maturity
- Again AdamW and LBFGS optimizer
- Training took 15 seconds on Colab Nvidia P4 GPU

- Extremely computational involved to generate the after hedge distributable earning profile with traditional explicit method. Need to run stochastic on stochastic for each point on the chart.
- Only took 14 seconds with the neural network on Colab Nvidia P4 GPU
Result with neural network

**Fee hedge**

- smaller delta range
- smoother line

**Smaller NB strain**

**very tight group**

**Not going to -1 fast enough, retrain needed**

**Negative cte98**
A new training target

What we like about neural network hedge:
• More predictable eventual outcome
• Smoother delta curve, less rebalance
• Smaller initial new business strain

What we don’t like neural network hedge
• The big drawdown due to cte98 flooring
• All up side equity risk hedged away
• Those are typical issue for fair value hedge

A new blended training target to fix it:
• Minimize Std (old target)
• Minimize cte98 if positive
• Minimize the hedge size

Retrain every 3 month.
New Target Result

- Smaller delta initially
- Stopped hedging fees when certain no claim
- 100% delta when certain ITM
- Downside protected
- More profit retained
- Smaller drawdown
New target discussion

- About 15 minute to retrain on Colab T4 GPU
- Still smooth delta curve
- Slighter smaller hedge in the beginning
- The model reduces hedge size much faster after the downside risk is sufficiently protected
- Drawdown due to cte98 flooring is much smaller
  - The biggest drawdown is compared to NB strain without neural network hedge,
  - potentially could be smoothed out with voluntary reserve
- Even smaller initial new business strain
  - With current extreme low interest rate, this benefit might evaporate
Discussion and Conclusion
Discussion

- Why explicit hedge with neural network instead of implicit fair value hedge?
  - Potentially smaller total asset requirement since implied volatility typically is higher compared to AAA scenario volatility
  - We can do a halfway between fair value hedge and real-world hedge and retain features from both sides.
  - CTE98 floor issue. With neural network hedge, we can reduce the hedge position gradually when the liability gets out of money.
Conclusion

• It’s now possible to model explicit hedge for stat capital calculation

• Given a lognormal scenario set and stdev training target, neural network hedge is fair value hedge
  • For put or call like cashflow, the algorithm can match Black-Scholes delta

• The neural network can calculate the fair value hedge position/price with real world scenarios

• With a blended training target, the neural network model can work around the cte98 floor issue of fair value hedge and retains more upside equity risk
SOA
2020 VIRTUAL ANNUAL MEETING & EXHIBIT