Backtesting of dynamic hedging of registered index-linked annuities

Comparing static vs. dynamic hedging with a Monte Carlo perspective for decision making

Hervé Andrès
Alexandre Boumezoued
Ken Qian
Katherine Wang

The standard approach for hedging registered index-linked annuities (RILAs) in the industry is to statically hedge by buying back the set of options sold to the policyholder from the market. Static hedging can be viewed as a less risky hedging strategy because options purchased are guaranteed to replicate the liability payoff when the RILA matures. However, some insurers may also periodically dynamically hedge in order to take advantage of synergies with other parts of the business. For instance, some insurance companies with large existing variable annuity blocks may choose to statically hedge the call spread portion of the RILA, and then dynamically hedge the net delta exposure from the RILA out-of-the-money (OTM) put option and an existing variable annuity block to take advantage of the offsetting delta exposure between the two products. The reduced net delta exposure will decrease the notional required for hedging, which can reduce transaction costs and the size of potential collateral payments. Insurers can also be motivated to dynamically hedge to avoid paying high bid-ask spreads charged by the dealers on longer-tenor options.

The mark-to-market accounting volatility associated with dynamic hedging of RILA products is a concern for some companies, and is driven by realized equity market volatility and changes in implied volatility. Strategies that involve an insurer opportunistically deciding when to hedge statically or dynamically (for example, based on the current market environment) have been relatively under-explored. This research paper intends to explore such a strategy by attempting to forecast the dynamic hedge P&L by simulating stochastic scenarios of the market implied volatility surfaces and the underlying asset price. We then model dynamic hedging along the scenarios to understand the projected P&L distribution, and compare this with the actual hedge P&L that was observed historically. To the extent that this simulated dynamic hedge P&L is similar to the actual hedge P&L, the company could then use the simulated dynamic hedge P&L values to make decisions on whether or not to dynamically or statically hedge.

For this application, we make use of a new stochastic model that considers the path-dependent feature of implied volatility as a response to the underlying asset. The backtest results show that there are valuation dates where the expected dynamic hedge P&L from the simulation analysis is similar to the actual dynamic hedge P&L observed in the following hedge period. For dates where they are different, the primary source of deviation between the simulated expected P&L and the actual realized dynamic hedge P&L is realized equity market volatility, such as periods of extremely high “tail” realized market volatility from the 2020 COVID period. However, under less extreme market conditions, using the scenarios from the model can potentially help with decision making on expected performance of hedging strategies, as well as the range of potential dynamic hedge P&Ls, which can help the hedge manager make a decision that is reflective of the insurer’s risk tolerance.

Section 1 of this paper will provide more background information on RILAs, Section 2 will describe the asset and implied volatility model, Section 3 will provide more detail on the dynamic hedging setup, and Section 4 will summarize the comparison of the dynamic hedging results from the simulated scenarios and the historical data.

1. RILA introduction

RILA sales have seen significant growth over the past few years, as RILAs offer policyholders the opportunity to participate in the upside potential of a market index while providing downside protection at the same time. The returns on RILAs are linked to the performance of a market index such as the S&P 500 and are subject to a cap rate, the maximum return policyholders can earn. Policyholders can also specify a level of protection against significant losses, typically via buffer or floor. A buffer protects policyholders from the first layer of losses from the underlying index at a pre-determined buffer level while a floor protects policyholders from further losses beyond the pre-determined floor level. For example, for a RILA product with 10% buffer, policyholders are protected against the initial 10% decline in the index before incurring losses if the index continues to decline beyond 10%; for a RILA with 10% floor, policyholders would incur losses from the initial negative return and insurance companies are then liable for any additional losses beyond that point.

On top of the level of protection and the underlying index chosen by policyholders at the time of issuance, the policy term can be specified too. The most common RILA terms in the market are one-year and six-year terms. For a one-year RILA product with a 20% cap rate and 10% buffer, the typical static hedge option's portfolio would consist of:

- Long position in a one-year ATM call
- Short position in a one-year 10% OTM put
- Short position in a one-year 20% OTM call

Figure 1 shows the option payoff for a RILA with a 20% cap and 10% buffer, and Figure 2**Error! Reference source not found.** shows the corresponding payoff for a 20% cap and 10% floor.

FIGURE 1: OPTION PAYOFF FOR A RILA WITH 20% CAP AND 10% BUFFER

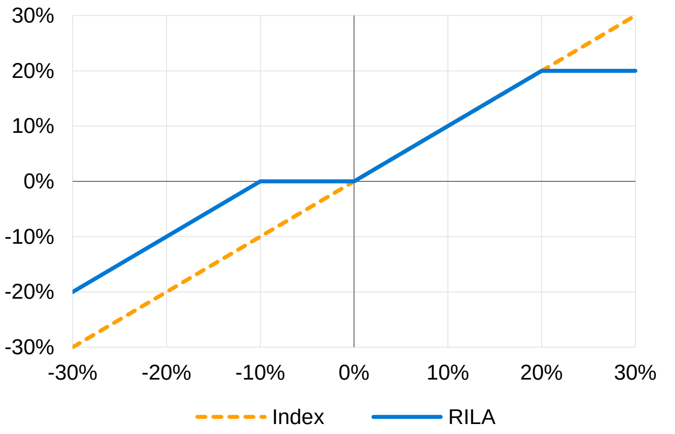
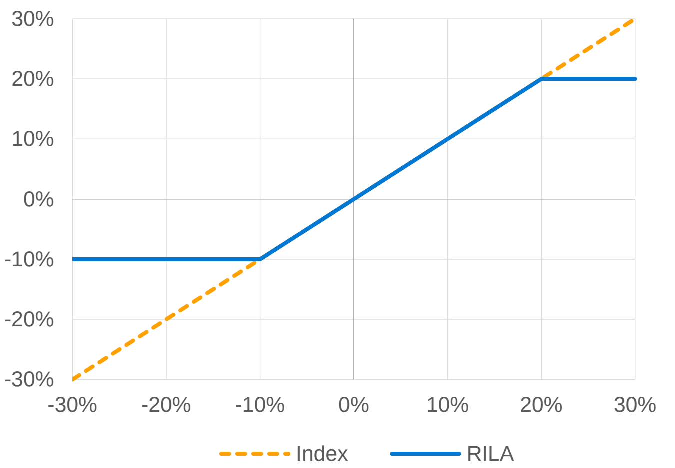


FIGURE 2: OPTION PAYOFF FOR A RILA WITH 20% CAP AND 10% FLOOR



2. Path-dependent implied volatility model

In order to estimate the probability distribution of the P&L of a dynamic hedging strategy, we utilized a new model[1], developed by two authors of the present paper along with Benjamin Jourdain, that jointly simulates the evolution of the underlying asset price (i.e., S&P 500) and the corresponding implied volatility surface (IVS). Note that other models could be considered for this task as well, such as the Black-Scholes model or the model from Cont & Vuletić (2023).[2] The purpose of this section is to provide a broad overview of the assumptions underlying the selected model.

There are many models that simulate the joint dynamics of an asset price and the corresponding IVS. Within these models, the dependence structure between the asset price and the IVS is captured through simple assumptions such as a Gaussian copula, common noise terms, or using the short-term implied volatility as a term in the asset stochastic volatility dynamics. In the paper "Implied volatility (also) is path-dependent", we have explored this dependence structure on historical data. More precisely, we conducted an empirical study of the historical joint behavior of the IVS and the underlying asset price based on a model structure first introduced in Guyon and Lekeufack (2023),[3] and we discovered a more complex dependence structure. More precisely, we showed that a large part of the movements of the at-the-money (ATM) implied volatilities for tenors of up to two years can be predicted using only the past returns and past squared returns of the underlying asset price. The model resulting from our empirical study describing the behavior of the ATM implied volatility over time can be written as follows:

$$IV_t^{ATM} = \beta_0 + \beta_1 R_{1,t} + \beta_2 \sqrt{R_{2,t}}.$$

The definition and interpretation of each term is provided below.

- IV_t^{ATM} is the ATM implied volatility at time t for a given tenor.
- R_1 is a trend feature defined by:

$$R_{1,t} = \sum_{t_i \le t} K_1(t - t_i)\, r_{t_i}$$

  where r_{t_i} is the daily return between day t_{i-1} and day t_i of the underlying asset price and K_1 is a decreasing kernel weighting the past returns.

1. Andrès, H., Boumezoued, A., & Jourdain, B. (2023). Implied volatility (also) is path-dependent. arXiv preprint arXiv:2312.15950.
2. Cont, R., & Vuletić, M. (2023). Simulation of arbitrage-free implied volatility surfaces. Applied Mathematical Finance, 30(2), 94-121.
3. Guyon, J., & Lekeufack, J. (2023). Volatility is (mostly) path-dependent. Quantitative Finance, 23(9), 1221-1258.

- R_2 is an activity or volatility feature defined by:

$$R_{2,t} = \sum_{t_i \le t} K_2(t - t_i) r_{t_i}^2$$

where K_2 is a decreasing kernel weighting the past squared returns.

We considered time-shifted power-law kernels for K_1 and K_2, which allows capturing both short and long memory:

$$K_i(\tau) = \frac{Z_{\alpha_i,\delta_i}}{(\tau + \delta_i)^{\alpha_i}}$$

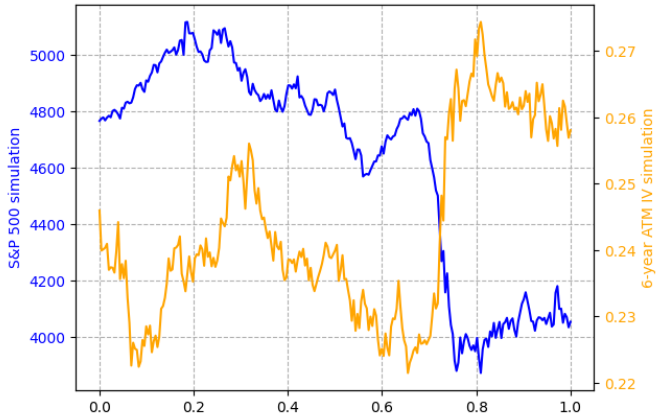
where Z_{α_i,δ_i} is a normalization constant and α_i, δ_i positive parameters. We calibrated this model on historical implied volatilities of the S&P 500 and the Euro Stoxx 50 and measured its performance using the R^2 score, which measures the proportion of the variance of the data that can be explained by the model. For the S&P 500, we obtained R^2 scores between 85% and 93% on the training set and between 62% and 77% on the test set. For the Euro Stoxx 50, we obtained R^2 scores between 85% and 90% on the training set, between 70% and 81% for the 15 first maturities on the test set and between 50% and 70% for the last maturities. These high R^2 scores show that implied volatility can be well predicted using the past path of the underlying asset price. Moreover, we obtained $\beta_1 \le 0$, which means that a drop in the underlying asset price leads to an increase in the implied volatility, and $\beta_2 \ge 0$, which means that an increase in the variability of the underlying asset price also leads to an increase in the implied volatility. Both effects make sense from an intuitive point of view: if the stock market drops and becomes more volatile, options investors will update their anticipations and consider that the stock is riskier, which translates into higher implied volatilities.

Based on this empirical study, we also propose a joint model for the dynamics of the full implied volatility surface using a parsimonious version of the Surface Stochastic Volatility Inspired (SSVI) parameterization (first introduced by Gatheral and Jacquier, 2014[4]), whose parameters are assumed to be stochastic processes. Mathematically, this means that the implied volatility at time t of an option of log-forward moneyness k and tenor T is given by:

$$IV_t(k,T) = \frac{\theta_{t,T}}{2}\left(1 + \rho_t \varphi_t(\theta_{t,T})k + \sqrt{\left(\varphi_t(\theta_{t,T})k + \rho_t\right)^2 + (1 - \rho_t^2)}\right)$$

where $\theta_{t,T} = a_t T^{p_t}$, $\varphi_t(\theta) = \frac{\eta_t}{\sqrt{\theta(1+\theta)}}$ and a, p, ρ, η are stochastic processes, whose detailed dynamics can be found in the original paper. The most important element in these dynamics is the fact that the processes a and p, which govern the ATM level of the implied volatility surface, depend on the past returns and the past squared returns of the underlying asset price, which allows reproducing the influence of the underlying asset price onto the implied volatility that one can observe on historical data. To showcase this influence, we calibrate the model on daily implied volatility surfaces of the S&P 500 provided by bank dealer quotes from 2007 to 2021. The calibrated model will then simulate paths of the S&P 500 and its implied volatility surface over one year with a daily time step. In Figure 3, we present one of these paths, where the effect of a sudden drop in the S&P 500 triggers a sudden increase in the implied volatility.

FIGURE 3: SIMULATION OF THE S&P 500 AND THE CORRESPONDING SIX-YEAR ATM IMPLIED VOLATILITY OVER ONE YEAR WITH A DAILY TIME STEP USING THE PATH-DEPENDENT SSVI MODEL OF ANDRÈS ET AL. (2023)



3. RILA hedging setup

The dynamic hedging analysis in this paper focuses on a six-year RILA on the S&P 500. We had particular interest in whether an insurer could benefit from focusing on dynamic hedging when market implied volatilities are elevated to avoid purchasing options at a premium. However, the net vega of the RILA (i.e., its sensitivity to the change in the implied volatility) is reduced when the vegas of the short OTM legs of the RILA are combined with the long position in the ATM call. As a result, our backtest assumed static hedging of the short OTM option positions and dynamic hedging of the long ATM call. The dynamic hedge P&L calculated in this analysis is then equivalent to the dynamic hedge P&L of a long position in a six-year ATM call option.

The dynamic hedge itself consists of using SPX futures to hedge equity delta, and a zero-coupon swap to hedge the rho of the option. Hedges were rebalanced on a weekly basis. Higher-order equity and interest rate impacts, as well as vega impacts from the movements in implied volatility that could not be hedged with equity futures, were left unhedged and would flow through into the dynamic hedge P&L.

4. Gatheral, J., & Jacquier, A. (2014). Arbitrage-free SVI volatility surfaces. Quantitative Finance, 14(1), 59-71.

For a given valuation date, we focused on the dynamic hedge P&L over a one-year hedging period in order to have more historical sampling periods for which to analyze the dynamic hedge P&L. Under this setup, the insurer is dynamically hedging the six-year ATM call option for one year, and then statically hedging by purchasing the necessary option back (which would be a five-year option). The dynamic hedge calculation was repeated monthly from 2016 to 2021 to cover a range of market environments.

The goal of this analysis was to establish whether the dynamic hedge P&L calculated along the stochastic scenarios is informative of the actual dynamic hedge P&L that was observed in the following period. The two main drivers of the dynamic hedge P&L over the one-year hedging period will be: 1) the change in implied volatility levels over the year and 2) the realized volatility over the year. To the extent that a scenario generator can reasonably forecast implied volatilities and the resulting market dynamics, then the dynamic hedge P&Ls calculated along those scenarios could provide a range of potential P&Ls that the company could expect to have if it decided to dynamically hedge. This information can help the insurer decide whether to engage in dynamic hedging at the outset.

4. Dynamic hedge P&L results

The dynamic hedge P&L for the six-year-long ATM call option was calculated monthly from January 1, 2016, to December 31, 2021, according to the following steps:

1. For each valuation date t, we calculated the dynamic hedge P&L over one year using the historical evolution of the S&P 500 and its implied volatility over $[t, t+1]$.

2. For each valuation date t, we calibrated the model described in Section 2 using daily implied volatility surfaces from 2007 to t, and we generated 1,000 scenarios from t to $t+1$.

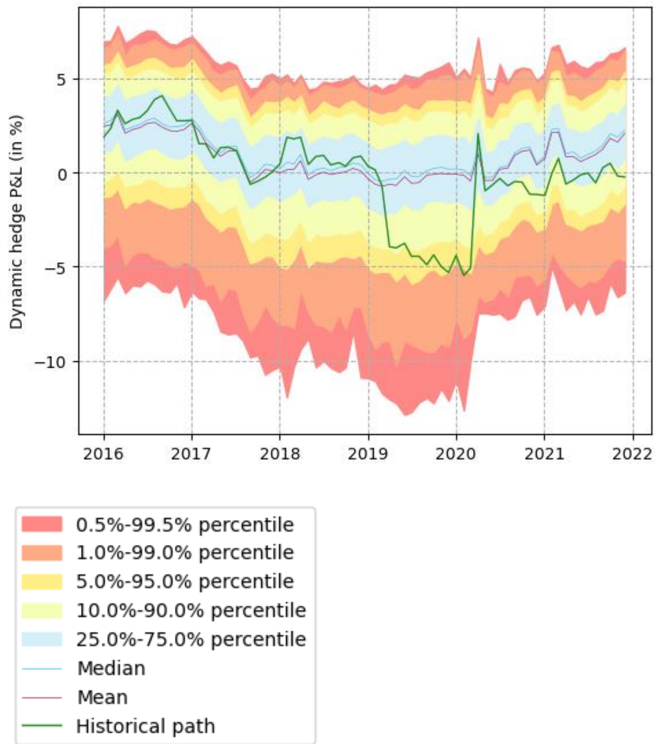
After implementing these two steps, we obtained the historical dynamic hedge P&L as well as 1,000 Monte Carlo dynamic hedge P&Ls at each valuation date. In Figure 4: QUANTILE ENVELOPES OF THE DYNAMIC HEDGE P&L FOR A SIX-YEAR ATM CALL OPTION UNDER A ONE-YEAR HEDGING PERIOD FROM JANUARY 2016 TO DECEMBER 2021Figure 4, we show the evolution of several quantiles and the mean of the dynamic hedge P&L distribution obtained from the simulated scenarios, as well as the corresponding historical dynamic hedge P&L. The P&L is expressed as a percentage of notional.

Figure 4 shows that the average P&L calculated along the scenarios tracks quite closely with the historical hedge P&L for the first few years of the backtest, but there are larger differences through most of 2019 and 2021. These differences were primarily due to higher realized volatility in the historical equity returns that the simulated scenarios were not able to capture. For instance, large losses from dynamic hedging were expected when hedging an option sold in 2019 because the large volatile returns in March 2020 from the COVID pandemic would have caused large gamma losses from dynamic hedging. Since such an exogenous event is difficult to predict, it was unsurprising that the average P&L obtained using the scenarios would be higher than the historical one.

We performed a more detailed attribution of the P&L for different valuation dates to provide a better understanding of the dynamic hedge P&L. The P&L over a week can be decomposed into the following components:

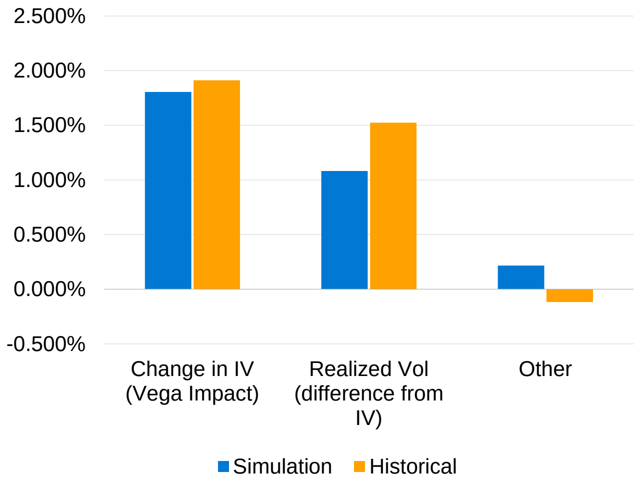
- **Impact of change in implied vol** (calculated on a weekly basis as: $Vega \times (EndingIV - StartingIV)$): This captures the impact of unanticipated changes in the implied volatility (outside of volatility skew) on the option value.
- **Realized vol impact:** This is calculated as: $Gamma \times ((SPReturn^2 - ImpliedVol^2)/2)$. To the extent that realized volatility is less than the implied volatility throughout the hedging period, this will be a gain; otherwise, it will be a loss.
- **Other:** This includes unhedged cross effects and higher order equity return impacts.

FIGURE 4: QUANTILE ENVELOPES OF THE DYNAMIC HEDGE P&L FOR A SIX-YEAR ATM CALL OPTION UNDER A ONE-YEAR HEDGING PERIOD FROM JANUARY 2016 TO DECEMBER 2021



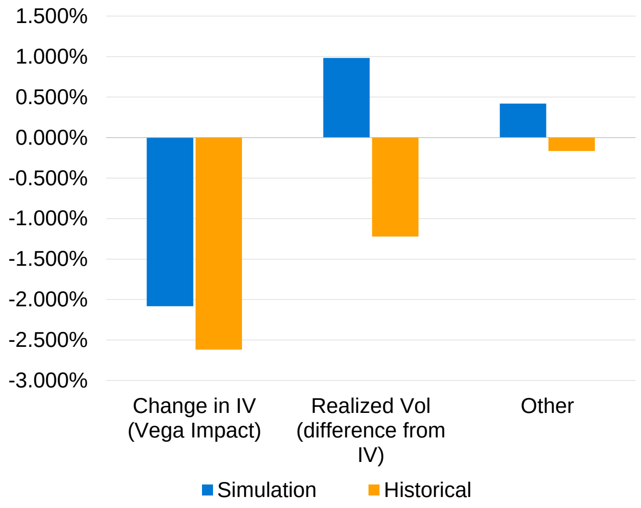
In the dynamic hedge P&L attribution for March 1, 2016, shown in Figure 5, the average P&L impact from implied volatility movements and realized volatility was quite similar between the model simulations and the historical data. Most of the other valuation dates in the backtest where the average hedge P&L from the model lined up closely with the historical scenario had a similar P&L attribution For these dates, the simulated scenario's realized volatility in the equity returns and forecast of implied volatility was fairly representative of what happened the following year in the historical scenario.

FIGURE 5: DYNAMIC HEDGE ATTRIBUTION ON MARCH 1, 2016



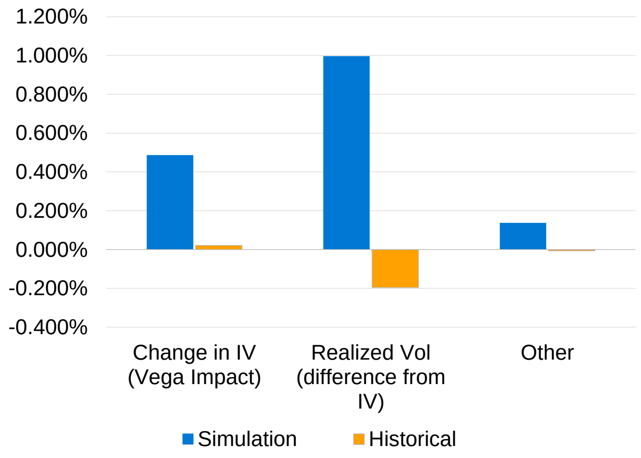
For the April 29, 2019, valuation date, the dynamic hedge P&L was significantly lower in the historical scenario compared with what was calculated along the simulated scenarios. The attribution in Figure 6 shows that this P&L discrepancy is mainly driven by higher realized volatility from the ensuing year in the historical data compared with the model. Along the stochastic scenarios, realized volatility was on average lower than the implied volatility at the time of the valuation date, thus resulting in gains on average (around 100 bp) from the realized volatility component. However, under the historical scenario, realized volatility was much higher because the one-year hedging window included the volatile period of February and March 2020 during the start of the COVID pandemic. This valuation date is representative of other valuation dates where the historical dynamic hedge P&L was worse than what was calculated along the simulated scenarios due to higher realized volatility in the historical scenario.

FIGURE 6: DYNAMIC HEDGE ATTRIBUTION ON APRIL 29, 2019



There were a few valuation dates in the historical backtest where differences in implied volatility forecasts also contributed to hedge P&L differences between the simulated paths and the historical scenario. The hedge P&L attribution chart in Figure 7 for November 1, 2021, shows that the main driver of the hedge P&L discrepancy is still the realized volatility component (the model was forecasting a large gain of around 100 bp, while there was a 20 bp loss along the historical scenario). However, differences in implied volatility forecasts also contributed to the total hedge P&L difference. On average, the model forecasted slight decreases in implied volatilities, contributing to an approximately 50 bp gain from volatility changes (net of anticipated impacts from skew); in contrast, the implied volatility changes were negligible along the historical scenario for the ensuing year.

FIGURE 7: DYNAMIC HEDGE ATTRIBUTION ON NOVEMBER 1, 2021



Conclusion and future considerations

The dynamic hedging backtest in this paper shows that for certain valuation dates, modeling dynamic hedging along the simulated scenarios can be a reasonable indicator of what the resulting P&L would be for the ensuing hedge period. The historical backtest shows that dynamically hedging had noticeable gains from 2016 to 2019, while it did not perform as well in the following years. The hedge P&L calculated along the simulated scenarios is consistent with this trend in 2016 to 2019, although there are larger discrepancies after that due to differences in realized volatility between the scenarios and the historical data. A company could then use these hedging simulation results to decide to construct a strategy of when to dynamically or statically hedge. For example, one possible rule could be to dynamically hedge whenever the average P&L calculated along the simulated scenarios is greater than some threshold level, such as 100 bp.

A future area to explore would be to compare the dynamic hedge P&L predictions from the simulation model used in this paper with alternative models, or from just using a simple heuristic. For example, a simple heuristic could be to simply not dynamically hedge whenever the current six-year ATM implied volatility is 100 bp higher than some long-term estimate of the six-year ATM implied volatility. Performing a comparison to this heuristic was outside the scope of this paper, but future research is encouraged.

CONTACT

Hervé Andrès
herve.andres@milliman.com

Alexandre Boumezoued
alexandre.boumezoued@milliman.com

Ken Qian
ken.qian@milliman.com

Katherine Wang
katherine.wang@milliman.com

Milliman