A.D. 1308
unipg
DIPARTIMENTO
DI ECONOMIA

Pricing and Risk Management of Long Dated Financial Contingent Claims with Applications to Life Insurance

Doctoral Dissertation of:
Filippo Granocchia

Supervisor: Prof. Gianna Figà-Talamanca
Year 2023/2024 - XXXVII Cycle

Abstract

This PhD Thesis is focused on pricing and risk management techniques for long dated contingent claims where no complete market is available for hedging purposes. The problem is approached from a potential life-insurer standpoint, with the purpose of selling long dated guarantees on a certain investment fund portfolio. With this in mind, we aim to investigate and define some best practices to design and build a strong risk management framework. The First Chapter of the Thesis sets the ground with the definition and review of Variable Annuities, while also tackling practical topics such as risk management practices (hedging) and performance measurement. The Second Chapter deep dives into more practical issues, such as the definition of possible models to use to model the underlying risk factors, as well as the implementation of a suitable hedging strategy based on *greeks* for such long dated contingent claims. Finally, the Third Chapter introduces a more innovative approach for pricing and risk management, by combining risk minimization hedging with Neural Networks.

Acknowledgements

This Thesis marks the end of a challenging but extremely rewarding and interesting multi-years journey. When in 2021 I enrolled to such PhD Program, I had very high expectations on what this could have meant for my personal and professional development and, as of now, I feel quite comfortable to say that these expectations have been far exceeded. I would really like to thank the Economics Department of the University of Perugia for this opportunity, as well as all the Professors who shared their knowledge and experience on the extremely insightful PhD courses which I had the pleasure to attend.
A great thank you goes also to my Supervisor, Prof. Gianna Figà-Talamanca, for all the coordination work, valuable ideas and contributions she gave during these years.
Finally I would like to extremely thank all my Family for always believing in me, and especially my beloved wife, Chiara Maria, who really made this journey possible by helping, supporting and motivating me, as well as my kids: Caterina Maria, who spent quite some time "studying and researching" with daddy and Leone, who is yet to be born, but I believe already gave me some precious inspiration.

Contents

List of Figures

List of Tables

List of Algorithms

Nomenclature

AI	Artificial Intelligence
ATM	At the Money
AV	Account Value
$CVaR$	Conditional Value at Risk
ES	Expected Shortfall
$GMAB$	Guaranteed Minimum Accumulation Benefit
$GMDB$	Guaranteed Minimum Death Benefit
$GMIB$	Guaranteed Minimum Income Benefit
$GMWB$	Guaranteed Minimum Withdrawal Benefit
ITM	In the Money
KPI	Key Performance Indicator
LLP	Last Liquid Point
LPM	Lower Partial Moment
$LSTM$	Long Short-Term Memory
NN	Neural Network
OTM	Out of the Money
SLV	Stochastic Local Volatility
VaR	Value at Risk
VAs	Variable Annuities

1 | Variable Annuities' Risk Management Overview

The first chapter of this PhD work is devoted to analyze the world of Variable Annuities products, with specific focus on their features and on the common pricing and risk management techniques used by practitioners.

1.1. Introduction

Over the last decades, life insurers started to develop more and more complex products with the aim to both, offer attractive and (potentially) remunerative packages to increasingly sophisticated clients, and not to fall behind a generally complex financial environment, where structuring innovations to get the best (and cheapest) risk-return solutions are continuously developed and pushed for sale by market participants.
Variable Annuities products (VAs) fall within this scope, and were primarily designed to offer, to clients looking to secure an attractive retirement package, the benefit (in terms of expected returns) of investing into global financial markets (both equities and bonds), while keeping risks limited thanks to the various (market-linked) guarantees provided by the insurance undertakings.
Similarly to unit-linked life insurance products, premiums related to VAs are invested into a contractually specified allocation among one or more investment funds proposed to clients by the insurance company. For VAs, on top of the investment into the units, clients are also charged for the additional guarantees related to the unit-linked investment. The price charged for this latter component is clearly dependent on the type and level of these guarantees and insurance companies have wide structuring possibilities. Usually, the offered guarantees show a higher degree of complexity and longer maturities than liquid financial options quoted by investment banks. This element, besides significantly affecting VAs costs for policyholders, has to be carefully assessed by the insurance company during the product design phase as it may subsequently result in hedging difficulties.
As also reported in [1], some basic VAs structures can be identified as follows:

- **GMWB** (Guaranteed Minimum Withdrawal Benefit): deferred or immediate, it entitles the policyholder to make periodic guaranteed withdrawals (plus potential upside returns), even if the account value has fallen to zero;
- **GMAB** (Guaranteed Minimum Accumulation Benefit): at maturity, the policyholder is entitled to get the maximum between the account value and a pre-defined guaranteed level;
- **GMIB** (Guaranteed Minimum Income Benefit): minimum guaranteed lifetime or term annuity starting at a predefined age on a defined benefit base;
- **GMDB** (Guaranteed Minimum Death Benefit): minimum guaranteed benefit in case of policyholder's death.

Few additional elements may be also highlighted in connection to VAs' guarantees:

- Guarantees are provided by the insurance company and not by the fund itself, *i.e.* the product category is different than protected or guaranteed funds;
- Guarantees are individual as they are linked to each single policyholder. This means that the single contract position is guaranteed rather than the value of the fund unit itself, and that the guaranteed level of the underlying funds may differ for each policyholder depending on the funds NAV (Net Asset Value) levels at the time of investment;
- Usually, guarantees are explicitly and separately charged, and the payoff at maturity of the contract is clearly defined, so the policyholder can transparently track the evolution of the investment.

VAs have been widespread instruments in US and, especially after the 2008 Financial Crisis, gained considerable traction also in Europe due to the protection features they offer to clients in case of markets downturns. In the recent years, VAs, Including Registered index-linked annuity (RILA), have remained a popular investment solution in US with 98.8Bn $ sold in 2023[1], while the EU market for such products remained quite thin and was certainly not helped by a financial environment where interest rates are negative or very low across the whole yield curve (considering EUR swap rates).

[1]Fixed Annuities sold for 286.6Bn $ in 2023. Source: LIMRA U.S. Individual Annuities Sales Survey

1.2. Simple Prototype of a Variable Annuity Product

A simple prototype of VA product may be represented through the structuring of an insurance policy combining a GMAB (Guaranteed Minimum Accumulation Benefit) with a GMDB (Guaranteed Minimum Death Benefit) rider. For such case, the final payoff for the policyholder would be defined by the GMAB component in case of survival until maturity date and it would be otherwise described by the GMDB component in case of death before maturity date.

Let us assume for simplicity that the product allows for a single premium only, which is paid at time $t = 0$ by the policyholder and that is invested until the end of the accumulation phase T (or T^D, $T^D \leq T$ in case of premature death of the insured), which we also assume to be the final maturity date of the contract. It is also assumed a deterministic behavior of the policyholder with no total or partial lapses.

The premium net of loadings P is invested into the investment fund(s) chosen by the policyholder at subscription of the contract. In $t = 0$ we have that P is equal to the account value of the contract A_t. For $t > 0$, A_t is defined by the evolution of the reference fund(s). It is important to remark that these funds may be invested into a wide range of assets and that the chosen allocation will have a considerable impact on the pricing and hedging of the products.

Assuming that the death benefit is not triggered during the lifetime of the product, at maturity T, the payoff of the GMAB component b_T^A is as follows [2]:

$$b_T^A = \max\{A_T, G_T^A\}, \tag{1.1}$$

where A_T is the account value at maturity (*i.e.* the pure investment funds return) and G_T^A is the contractually defined guarantee.

More specifically, the insurance company has considerable flexibility in the structuring of the guarantee G_T^A. Few examples are illustrated here below:

- G_T^A is defined through a static coefficient on the initially invested premium P:

$$G_T^A = \alpha P, \tag{1.2}$$

  where an $\alpha = 1$ would imply a 100% guarantee on the initial investment, an $\alpha < 1$ would imply a partial guarantee and an $\alpha > 1$ would imply a minimum guaranteed return.

- The minimum guarantee level is defined through the roll-up of the premium P at a pre-defined rate. For example, under continuous compounding the guarantee will evolve as follows:

$$G_T^A = P\mathrm{e}^{\delta T}, \tag{1.3}$$

where δ represents the guaranteed instantaneous rate which can be fixed for the entire life of the product or can change according to some contractual specifications (*e.g.* in a time dependent way or with an inflation-linked structure). The roll-up compounding mechanism is contractually defined and may also reflect different schemes than the continuous one, e.g. simple compounding (as shown in Figure 1.1).

- The G_T^A guarantee is structured through a ratchet mechanism, where the account value level is observed with a certain time schedule and used as a parameter to set the guarantee itself:

$$G_T^A = \max_{t_i < T} A_{t_i}, \tag{1.4}$$

with $i = 0, 1, \dots, n$ and $0 = t_0 < t_1 < \cdots < t_n < T$. Such mechanism allows to lock-in the investment fund(s) returns and it can be a useful tool in terms of product design so to reduce the policyholder lapse risk in case the evolution of the fund(s) would considerably outperform the level of the guarantee, therefore resulting in no incentive for the client to keep paying high guarantee fees for an out of the money guarantee.

The above reported guarantee structures may be also suitable to be combined together.

Further to the GMAB component, a prototype product can also offer a GMDB rider to enhance the perimeter of the policy coverage (*i.e.* retirement objective achieved through the GMAB and death coverage achieved through the GMDB).
For this latter, the guaranteed payoff is uncertain and potentially triggered by the policyholder's death during the accumulation phase (*i.e.* only if the death happens for $T^D \leq T$). In terms of potential benefits to offer, the flexibility degrees are quite wide also for the GMDB case and actually all the structures already mentioned for GMAB can be recalibrated for the GMDB, with the only difference that the guaranteed payoff will be only paid in case of death within T. Here the benefit b_t^D is formalized as follows:

$$b_{T^D}^D = \max\{A_{T^D}, G_{T^D}^D\}, \quad T^D \leq T, \tag{1.5}$$

where G_t^D represents some contractually specified guaranteed amount.
On top of the guarantee structures already mentioned for the GMAB case, for the GMDB also a reset feature can be considered. In this case, the G_t^D guarantee will be reset at

certain (contractually defined) observation dates. If the client dies before T, the account value recorded at the last observation date will be taken into account for the guarantee purpose:

$$G^D_{T^D} = A_{\max\{t_i : t_i < T^D\}}, \tag{1.6}$$

where times t_i, $i = 1, 2, \dots, n$ define the reset dates.
It is important to remark that, unlike the ratchet case, here the guaranteed minimum amount can decrease over time.

The guarantees embedded in such products are similar to financial options and, as remarked in [3], the total contract value at time $t < T$ can be decomposed in the non-guaranteed component plus the guarantee by using a put-decomposition approach. For VAs the first term is equal to the account value invested into the fund(s), while the second one is effectively the present value of the guaranteed benefits specified by the policy termsheet, which can be seen as a put option on the fund(s) investment.
For the illustrated example (GMAB + GMDB), the present value of the insurance product V_t would then be:

$$V_t = A_t + O^A_t + O^D_t, \tag{1.7}$$

where A_t can be observed from the fund(s) NAV and O^A_t and O^D_t represent respectively the present value of the (put option-like) accumulation and death benefits. The pricing of these two latter components for $t < T$ represents an important challenge for the insurance company as several variables need to be studied and modeled. More specifically, a pricing approach based on the common market practice used for financial derivatives is applied in such context. This involves discounting, through risk-free discount factors, the future expected payoff under the risk-neutral probability measure. Due to the complexity of the products to be modeled and priced, numerical methods such as Monte Carlo simulation are commonly considered as the best practice to use for such purpose. However, also approximations based on Black-Scholes like closed formulas have been tested in the past years, especially for small insurance portfolios [4].

1.2.1. Visualization of Possible GMAB Structures

It may be useful, at this point, to visualize the life-cycle of some possible GMAB structures. We consider a product with 20 years to maturity and an initial investment on the account value of 100 monetary amount.

As shown in Section 1.2, we may want to consider three main types of guarantee: fixed, roll-up and ratchet. The first two types may be relatively easy to understand as they can be structured in a "deterministic" way, in the sense that they are not driven by any stochastic factor. The ratchet becomes a bit more challenging due to its path-dependency to the account value.

Figure 1.1 shows how this 20 years GMAB could look like over time under the three structures discussed above. The account value evolution is simulated and shows a sharp increase in the first 10 years, followed by a huge drop between 10 and 15 years, with a final recovery, however remaining below the initial level.

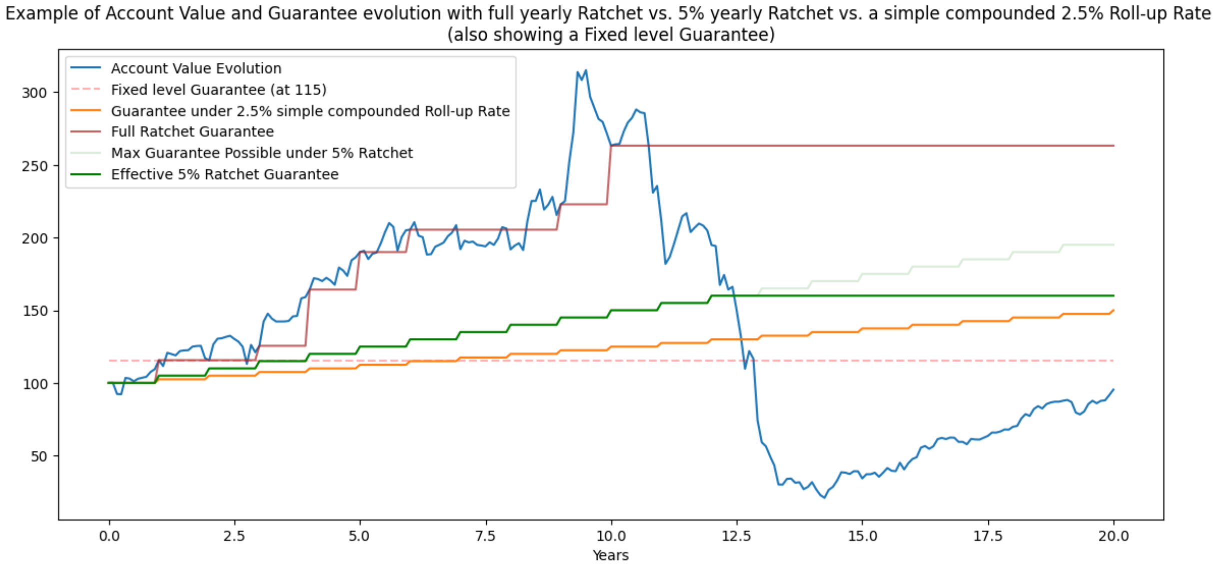


Figure 1.1: Example of a possible GMAB prototype with Account Value evolution and different Guarantee structures (Full Ratchet, Capped Ratchet, Roll-up and Fixed level).

From the picture, it is possible to understand how the four types of guarantee work. The fixed one stays flat over the entire 20 years, while the roll-up one, constantly increases over time at the 2.5% simple compounding rate. The 5% ratchet kicks in until the 12th year thanks to the outstanding performance of the account value and then flattens out afterwards, due to its drop. Similarly, the full ratchet is triggered until the 10th policy anniversary, locking in the almost entire performance of the AV. At the contract expiration, all the four guarantees end up to be in the money, with the full ratchet bringing the highest return to the policyholder. In this specific example, we only analyzed the possible payoff of each type of guarantee without considering the corresponding price to be paid. Clearly the choice of the appropriate solution would need to both assess the desired payoff structure, as well as the associated price.
The picture can also be used for some product design considerations. As it will be remarked in the following Section 1.3, not all the risks related to VAs can be hedged away. Among the non-hedgeable ones, policyholder behavior can certainly play a major role.

Depending on the product structuring, this risk may be quite relevant, especially in the case where the guarantee fees are not collected entirely upfront, but they are spread over the lifetime of the insurance contract and therefore partially lost by the insurance company in case of lapse. In the example reported, the account value massively increases in the first 10 years, getting to levels much higher than all the three "capped" types of guarantee. Assuming that the guarantee fees are paid every year until maturity, a rationale policyholder may in this case consider the possibility to surrender the policy, getting back the account value and maybe purchasing a new contract with a new (much higher) guarantee level. In such scenario, the insurance company would then need to fully unwind the hedge portfolio related to that policy, *de facto* having to absorb, with its own balance sheet, the remaining present value of the uncollected guarantee fees. A surrender charge to the policyholder may be used to at least partially mitigate such risk. Also the full ratchet mechanism could be used as a lapse risk mitigation tool. In fact, the level of the guarantee would closely follow the peaks in the AV, always remaining at least close to the moneyness. In case the policyholder would then surrender, it could be very likely that the guarantee will be in the money, so the insurance company would at least compensate the missed guarantee fees PV with the hedge portfolio gains (as the policyholder could only claim the guarantee at product maturity).

1.3. Risks for VAs Insurers

As a matter of fact, life insurance companies active in the VAs business have to deal with usually complex products that generate heterogeneous exposures and, therefore, risks. More specifically, as remarked in [5], the main aggregated risk drivers are here below explored.

1.3.1. Market Risk

By definition, market risk represents the risk of an adverse price movement on a certain asset or portfolio (of liabilities in this case). Due to the nature of VAs, this is normally the main risk that insurance undertakings have to face when dealing with such products, as the guarantee value (*i.e* what the company owes to the policyholder) is strictly related to the evolution of the underlying funds which is invested into financial instruments (mostly equities and bonds).

Depending on the way guarantees are structured (*e.g.* maturity, guarantee level, strategic asset allocation of the chosen fund(s), optional guaranteed life or term annuity after maturity), the portfolio can be exposed to different market risk factors:

- Underlying Fund Risk (Delta): changes in the NAV of investment fund(s) picked by the policyholder have a clear effect on the liabilities' value. The NAV change is mostly driven by market movements, but is also impacted by the overall costs charged by the fund(s), which have to be considered and modeled during the product design phase, and monitored on an ongoing basis. Due to the put-option like structure of VAs, a decrease in the NAV would bring to an increase of the guarantees value. It is very important to note that the fund portfolios can be populated with several kinds of assets classes (e.g. equities, bonds, FX) and therefore the ultimate risks for the insurance company arise from the effective instruments held by these funds.

- Interest Rate Risk (Rho): due to the usually long dated nature of VAs liabilities, changes in interest rates have a considerable effect on the present value of the long-term income provided by the contract guarantees. The sensitivity to such risk factor may increase considerably on products like GMIBs where the policyholder can exercise the option to get a term or lifetime annuity after the accumulation phase.

- Volatility Risk (Vega): being an option-like structure, VA guarantees are sensitive to the changes in the future expected volatility of the financial assets where the underlying funds are invested. Market changes on the quoted implied volatilities have a direct impact on the liabilities.

- Interest Rate Volatility Risk (IR Vega): this risk is more pronounced for products having a swaption like feature, for example GMIBs, where the client can choose to annuitize at a certain rate the guaranteed amount after the accumulation phase. Here, an increase of IR volatility would increase the probability of low rates at the annuitisation time with a consequent liabilities' PV increase.

- Second order (Gamma) and cross effects: even if not having a direct impact on the liabilities PV, a movement in the market risk factors can result in a change on first order sensitivities (even the ones related to a different risk factor) which can, on one hand, reduce the efficiency of the hedging program and, on the other hand, force the company to frequent restructurings of the hedge portfolio, bearing relevant transaction costs.

- Liquidity Risk: to set up a proper hedging, the insurance company would need to find liquid instruments on the financial markets that strongly approximate the behavior of the guarantees in terms of PV. Due to the long dated nature of VAs, this task is not always easy to accomplish. It is for example unlikely to be able to trade equity options with tenors beyond the 10Y expiry.

1.3.2. Longevity/Mortality Risk

Variable annuities are not just financial instruments, but rather a complex mix which gathers together financial components with actuarial ones. Often, mortality rates have an influence on the pricing, reserving and notional amount of the guarantees to be hedged. Depending on the product's features, the insurer may be exposed to either mortality (e.g. GMDB structure) or longevity risk. This latter would be faced when, for example, products are based on a living benefit for the policyholder (i.e. the more clients live, the higher the cost for the insurance company). Like more traditional life insurance products, longevity risk can play therefore a considerable role on VAs, as it can result in an increased duration of the payments associated with the living benefits which will have to be finally borne by the insurance company.

1.3.3. Policyholder Behavior

Similarly to the mortality effect, policyholders behavior has a strong influence on the notional amount of the guarantee to be hedged as, in case of lapse before maturity date, clients will not be allowed to benefit from the guarantee and will only get back the account value (but will also avoid to pay future expected fees). Failing to properly assess lapse behavior may result in serious P&L volatility for the insurance company as it may lead to considerable under/over hedging. A common approach to reduce such effect is to define the lapse behavior as a function of the in-the-moneyness of the guarantee and time to maturity.

As a matter of fact, modeling such component is particularly hard, as policyholder behavior may not be always completely rationale, and therefore guided by the effective moneyness of the guarantee, but rather driven by personal choices or exogenous factors hard to capture with a model.

Lapse risk is also hard to hedge (it would entail trading some tailor made derivatives and a counterparty willing to trade them), so the best solution would be to limit it through the product design phase (*e.g.* by selling single premium products only or by adding a ratchet mechanism so to avoid deep out of the money guarantees)

1.3.4. Hedging Risk

As already highlighted throughout this Thesis, hedging VAs is key to maintain a decent Balance-Sheet stability. However, the risk that the implemented hedging program would

not perform as expected is a material one, and this may potentially result in an actually even bigger damage for the company than to not to hedge at all. It may happen for example that instruments chosen for hedging purposes would not properly fit the liabilities structure, creating huge P&L volatility, or that losses would arise due to the default of a counterparty. It is for such reasons that proper risk margins, as well as future expected trading costs, have to be considered and accounted for in the product design phase, and that a proper hedge portfolio monitoring on an ongoing basis and operational and counterparty risk mitigation techniques have to be put in place.
Further risks may also arise from questionable or wrong choices in terms of models to use for the main financial drivers. Particularly for long dated policies, this may have a direct and considerable impact on both pricing and risk metrics calculations (and therefore on the hedge portfolio structuring), potentially leading to undesired P&L effects.

1.3.5. Regulatory Considerations on VAs

Due to the very delicate purpose and high complexity of VAs, any insurance undertaking interested to deal with such products should thoughtfully assess, in addition to the pure challenges arising by the VA products, the regulatory requirements applying to them.
For example, under risk-based regulatory regimes such as Solvency II in Europe, the financial guarantees embedded into VA products can result in considerably high capital requirements.
As shown in [6], a key element to ensure the sustainability of the business is to mitigate non-financial and certain financial risks through product design and to set in place an accurate and effective hedging program in order to offset as much as possible the PV changes, driven by financial factors, of the liabilities with those of the hedge portfolio. However, despite the hedge accuracy, the insurance company needs to also be aware of the current and, especially, future capital requirements led by changes in some un-hedged exogenous parameter, which may result into economic risk and reduced profitability of the business.

1.4. Risk Management of Variable Annuities

As outlined in Section 1.3, dealing with Variable Annuities exposes Life Insurance Companies to a wide range of Financial and Non-Financial risks. While Non-Financial risks need to be generally mitigated through product design as they are not directly hedgeable since there is no liquid nor complete market to be used, a careful VA provider would try to immunize as much as possible its net portfolio from Financial (Market) risks.

Although this objective should be theoretically possible under the Market completeness (and therefore frictionless) assumption, this would be practically hard to obtain in practice, especially for long dated (10+ years) liabilities with underlying equity indexes.

There can be countless ways to approach the hedging problem. In general every financial operator would need to design the appropriate hedging policy based on several key elements such as, but not limited to, the nature and volume of the business to hedge, its Risk Appetite Framework, the availability of proper hedging instruments on the market.

Chapter 2 and 3 will deep dive into some practical hedging implementations to be used when dealing with long dated equity derivatives. Here we would like to first focus on the basic hedging theory with the following approaches:

- **Static Hedging**: A strategy built to avoid rebalancings over a certain time horizon.
- **Dynamic Hedging**: A continuous rebalancing approach, based on the delta of the derivative.
- **Delta-Gamma Hedging**: An extension of delta hedging that accounts for gamma, the second derivative with respect to the underlying asset.
- **Super-Hedging**: A robust method that covers the derivative's payoff in extreme scenarios.

1.4.1. Static Hedging

Ideally and conceptually, static hedging would represent the best solution for an operator seeking for perfect portfolio immunization. Let's still take into account the case of an Insurance Company selling guarantees on investment products. It would be perfect if, as soon as the product is sold and the margin is locked in, the Company would be able to hedge away all the financial risks in one go and have perfect offsetting of the liabilities book with the hedge portfolio until the insurance product expires, or at least until a certain future rebalancing date. This objective may be achieved in case the product sold would be basically replicable with financial instruments commonly traded on the market, but that would be rarely the case for complex and long dated payoff structures.

The main breakthrough on the static hedging field is represented by the methodology proposed by Peter Carr and Liuren Wu in [7]. The Authors assume frictionless markets and no arbitrage, with the underlying's spot price S_t always satisfying the condition $S_t \geq 0$ and constant continuously compounded risk-free rate r and dividend yield q, Furthermore, the analysis is restricted to the class of models where the risk neutral evolution of the stock price is Markov in the stock price S and in the calendar time t.

Consider the time-t price of a European call with strike K and maturity T, denoted as $C_t(K,T)$. It is assumed that the state is fully described by the stock price and time implies that there exists a call pricing function $C(S,t;K,T;\mathbb{Q})$ such that:

$$C_t(K,T) = C(S_t,t;K,T;\mathbb{Q}), \quad t \in [0,T],\ K \geq 0,\ T \in [t,T].$$

Thus, the call pricing function relates the call price at t to the state variables (S_t,t), the contractual parameters (K,T), and to a vector of deterministic model parameters $\mathbb{Q}$.

The paper then introduces a theorem formalizing a new spanning relation between the value of a European option at one maturity and the value of a continuum of European options at some nearer maturity. The practical implication of this theorem is that one can span the risk of a given option by taking a static position in a continuum of shorter-term, usually more liquid, options.

Theorem 1.1 *Under no-arbitrage and Markovian assumption, the time-t value of a European call option maturing at a fixed time $T \geq t$ is related to the time-t value of a continuum of European call options at a shorter maturity $u \in [t,T]$ by*

$$C(S,t;K,T;\mathbb{Q}) = \int_0^\infty w(K)C(S,t;K,u;\mathbb{Q})\,dK, \quad u \in [t,T],$$

for all possible nonnegative values of S and at all times $t \leq u$. The weighting function $w(K)$ is given by

$$w(K) = \frac{\partial^2}{\partial K^2}C(K,u;K,T;\mathbb{Q}).$$

Interestingly, the option weights $w(K)$ are independent of S and t. This ensures that the hedging allocation would be independent from the point in time $t \leq u$ and the level of S, meaning that that no rebalancing will be needed before the expiration date of the spanning portfolio.

Although providing an elegant and insightful hedging approach, the practical implementation looks particularly complicated because of the fact that markets are not effectively frictionless and transaction costs may therefore become relevant, especially to trade many options far from the moneyness and, most important, this approach does not really perform well under stochastic volatility. The main issue the hedger will encounter would be that when the shorter dated hedging instruments will get to expiration, then he will need to roll over his "static" hedging strategy with new options traded at a new, and most likely different, volatility level than the initial one (or in any case the one which was expected on a forward basis for the rebalancing date), potentially nullifying the hedging appropriateness. Rollover risk will therefore play a key role in this case.

An interesting extension for this hedging approach with the aim to extend the spanning relationship to a multi-period option expiry scheme is provided in [8], with the following main result:

Theorem 1.2 *Under no-arbitrage and Markovian assumption, the time-t value of a European call option maturing at a fixed time $T > t$ relates to the time-t value of a continuum of European call options having shorter maturities $0 < u_2 < u_1 \leq t$ by:*

$$C(S,t,K,T) = \int_{K_{11}}^{K_{12}} w(K_1) C(S,t,K_1,u_1)\, dK_1 + \int_0^{\infty} \tilde{w}_2(K_2) C(S,t,K_2,u_2)\, dK_2,$$

with weights:

$$w(K_1) = \frac{\partial^2 C}{\partial K_1^2}(K_1, u_1, K, T),$$

and

$$\tilde{w}_2(K_2) = \int_0^{K_{11}} w(K_1) w_2(K_2, K_1)\, dK_1 + \int_{K_{12}}^{\infty} w(K_1) w_2(K_2, K_1)\, dK_1,$$

where:

$$w_2(K_2, K_1) = \frac{\partial^2 C}{\partial K_2^2}(K_2, u_2, K_1, u_1),$$

and $0 < K_{11} < K < K_{12} < \infty$ denotes the range of liquid strikes available at initial time t_0, corresponding to the options with maturity u_1.

The Authors present the application of their approach on both Black-Scholes and Merton Jump Diffusion frameworks. Although this approach may provide more flexibility in the choice of data points to use for hedging purposes, it would still expose the hedger to volatility rollover risk.

1.4.2. Dynamic Hedging

Dynamic hedging, introduced in [9], requires to continuously adjust the position in the underlying asset to maintain a risk-neutral portfolio. The key to dynamic hedging is the *delta* (Δ), defined as the sensitivity of the option's price to changes in the underlying asset price:

$$\Delta = \frac{\partial V}{\partial S},$$

where V is the option value and S is the underlying asset price. In practice, a delta-neutral portfolio is achieved by holding Δ units of the underlying asset against the option position.

Dynamic hedging relies on the assumption of continuous trading and normally distributed asset returns. However, in reality, trading frictions such as transaction costs and liquidity constraints make continuous hedging challenging. When asset prices exhibit jumps, the strategy's effectiveness is further compromised. To account for these issues, practitioners often adjust the frequency of rebalancing, balancing between hedging accuracy and cost.

1.4.2.1. Gamma and Theta Risks

Dynamic hedging alone may not fully address the risks inherent in derivative positions. Two additional sensitivities, *gamma* (Γ) and *theta* (Θ), are important to consider:

$$\Gamma = \frac{\partial^2 V}{\partial S^2}, \quad \Theta = \frac{\partial V}{\partial t}.$$

Gamma measures the rate of change of delta with respect to the underlying price, while theta represents the sensitivity of the option price to time decay. A high gamma means the option's delta will change quickly with price movements, complicating the hedging process.

Also additional risk factors such as volatility and interest rate risk may need to be factored into the hedging strategy, with a even higher degree of complexity in the hedging process. Further insights on this approach will be provided in the following Chapter, especially in 2.6.1.

1.4.3. Delta-Gamma Hedging

A popular hedging approach often used by practitioners is the Delta-gamma hedging. This aims to mitigate both delta and gamma risks by constructing a portfolio that is neutral to small and moderate changes in the underlying asset. This approach uses a combination of the underlying asset and other derivatives to neutralize the effects of changes in both the level and curvature of the option's price relative to the underlying [10].

For a portfolio to be delta-gamma neutral, it must satisfy:

$$\Delta = 0, \quad \Gamma = 0.$$

Achieving gamma neutrality often involves using additional options with different strikes or maturities. While this can reduce exposure to large price movements, the complexity and cost of managing a delta-gamma neutral portfolio increase, especially in volatile markets where gamma can vary significantly.

In practice, delta-gamma hedging is implemented by constructing a portfolio with offsetting positions in options that reduce the net gamma exposure. For example, if an option position has a high gamma, additional options with an opposing gamma effect are included in the portfolio to stabilize its response to large movements.

1.4.4. Super-Hedging

Super-hedging is a conservative approach that seeks to create a portfolio ensuring the payoff covers the derivative's obligations in all scenarios, including extreme market conditions. This technique is particularly relevant for risk-averse investors or in markets with high volatility [11].

In super-hedging, the goal is to construct a portfolio P such that:

$$P \geq V,$$

across all possible states, guaranteeing that the payoff of the derivative is fully covered. While super-hedging provides a robust solution, it often requires a larger initial investment due to the additional protective measures taken, particularly when there is significant uncertainty about future market conditions.

1.5. Measuring the Hedge Performance

As briefly discussed in the previous sections and as it will be extensively shown in the rest of this work, there is no universal approach to use to hedge a portfolio of derivatives. Many ways can be assessed, and every operator is likely required to decide which one to pick depending on the defined Risk Appetite Framework and many other aspects relating to market liquidity, strategy complexity and so on.

After a hedging strategy has been decided and implemented, it becomes crucial to assess its effectiveness. Here again, many metrics can be built for this purpose. Cotter and Hanly discuss the following ones in [12].

1.5.1. Hedging Effectiveness Metric 1 - The Variance

The variance metric (HE_1) measures the percentage reduction in the variance of a hedged portfolio compared to the variance of an unhedged portfolio. The performance metric is defined as:

$$HE_1 = 1 - \frac{\text{VARIANCE}_{\text{Hedged Portfolio}}}{\text{VARIANCE}_{\text{Unhedged Portfolio}}}. \tag{1.8}$$

This gives the percentage reduction in the variance of the hedged portfolio compared to the unhedged portfolio. A value of $HE_1 = 1$ indicates a 100% reduction in variance, while $HE_1 = 0$ indicates no risk reduction. A larger HE_1 indicates better hedging performance. Technically, the metric could even become negative in the unwelcome case where the hedged portfolio results more volatile than the unhedged one. Obviously in this case a review of the hedging program may be advisable.

While easy to calculate and interpret, this measure has limitations. Variance is a two-sided measure that does not differentiate between positive and negative returns, which may not align with hedgers' focus on downside risk. Moreover, for non-normal return distributions, variance alone may not fully characterize risk, necessitating alternative metrics.

1.5.2. Hedging Effectiveness Metric 2 - The Semi-Variance

The semi-variance measures variability below the mean and is defined as:

$$\text{Semi-Variance} = \mathbb{E}\left[\max(\tau - R, 0)^2\right] = \int_{-\infty}^{\tau} (\tau - R)^2 dF(R), \tag{1.9}$$

where τ is the target return and it is set to the expected return, R is the net portfolio return, and $F(R)$ is the cumulative distribution function of R. When τ is set to the expected return, the semi-variance becomes half the variance for symmetric distributions. The performance metric is defined as:

$$HE_2 = 1 - \frac{\text{Semi-Variance}_{\text{Hedged Portfolio}}}{\text{Semi-Variance}_{\text{Unhedged Portfolio}}}. \tag{1.10}$$

The semi-variance focuses on downside risk, addressing a key limitation of variance. However, it applies a single weighting to deviations from the mean and may not account for investors' differing risk preferences. As per the interpretation, it is similar to the first metric presented.

1.5.3. Hedging Effectiveness Metric 3 - The Lower Partial Moment (LPM)

The LPM generalizes downside risk measures. It is defined as:

$$\text{LPM}_n(\tau; R) = \mathbb{E}\left[\max(\tau - R, 0)^n\right] = \int_{-\infty}^{\tau} (\tau - R)^n dF(R), \tag{1.11}$$

where n determines the weighting of deviations from τ. For $n = 2$ and τ set to the mean, the LPM is equivalent to the semi-variance. Risk-averse investors often use $n > 1$. As per τ, it can be seen again as the investor's minimum acceptable level of return. Popular values are zero or the risk free rate.

The performance metric is:

$$HE_3 = 1 - \frac{\text{LPM}^n_{\text{Hedged Portfolio}}}{\text{LPM}^n_{\text{Unhedged Portfolio}}}. \tag{1.12}$$

The LPM is robust to non-normal distributions and provides insights into asymmetry in the joint distribution of spot and futures returns.

1.5.4. Hedging Effectiveness Metric 4 - Value at Risk (VaR)

Value at Risk (VaR) estimates the maximum potential loss over a given time horizon with a specified confidence level. It is defined as the $(100-x)^{\text{th}}$ percentile of the portfolio return distribution over the next N days. The VaR can be calculated from the LPM (Equation 1.11) with $n = 0$:

$$\text{VaR} = F^{-1}(1 - \text{LPM}_0), \tag{1.13}$$

where F^{-1} is the inverse cumulative distribution function. The performance can be then evaluated as:

$$HE_4 = 1 - \frac{\text{VaR}_{\text{Hedged Portfolio}}}{\text{VaR}_{\text{Unhedged Portfolio}}}. \tag{1.14}$$

1.5.5. Hedging Effectiveness Metric 5 - Conditional Value at Risk (CVaR)

CVaR, or Expected Shortfall, measures the average loss beyond the VaR threshold. It is defined as:

$$\text{CVaR} = \int_{-\infty}^{\tau} (\tau - R) dF(R), \tag{1.15}$$

where τ is set to the VaR level. CVaR accounts for losses beyond the $(100-x)^{\text{th}}$ percentile. The performance metric can be then computed as:

$$HE_5 = 1 - \frac{\text{CVaR}_{\text{Hedged Portfolio}}}{\text{CVaR}_{\text{Unhedged Portfolio}}}. \tag{1.16}$$

1.6. Modeling Considerations for VAs

Many aspects related to the VAs' life-cycle have been tackled so far, ranging from possible structures and product design to the ongoing risk analysis/management and hedging, also including some regulatory considerations. Another key element for a successful and sustainable VA business, which is also a necessary condition for the above aspects to be properly handled, is to use effective models/techniques/assumptions for the risk drivers impacting the evolution of the VA products. This is particularly important to produce both reliable pricing and risk sensitivities, needed to set up an efficient hedge portfolio.
The list of the main risk drivers to consider has been already debated in Section 1.3. Once identified, these risks need to be properly modeled and quantified in order to come up with a product pricing which would be sufficient to hedge or to cover them under certain assumptions, plus clearly adding a profit margin to make the business attractive.
Many contributions can be found in the literature using several model and techniques. In this thesis, we mostly focus on equity and equity volatility risk arising from the account value, assuming that the fund is fully invested into a unique equity index with a liquid enough options market. For this purpose, as it will be outlined on Chapter 2, we focus on a specific variation of a stochastic volatility model with a displacement factor. This model extends both the standard Black & Scholes [9] as well as the Heston stochastic volatility one [13].
The reasons behind the choice of such model lie on the fact that it allows a very good fit on the market implied volatilities while ensuring a stochastic behavior of the instantaneous variance. At the same time, the introduction of a displacement factor keeps ensuring the model tractability, while also guaranteeing that the variance does not fall to 0, therefore making it easy to be interpreted as a sort of minimum volatility floor and without the need to impose constraints on the parameters.
For completeness, further alternatives to be considered could have also been:

- **Local Volatility Model**: The Dupire Local Volatility Model [14] extends the Black-Scholes framework by allowing the volatility to depend deterministically on both the asset price S and time t.

The asset price under the risk-neutral measure follows:

$$dS_t = (r - q)S_t dt + \sigma_{\text{LV}}(S_t, t)S_t dW_t,$$

where:

- r is the risk-free rate,
- q is the continuous dividend yield,
- $\sigma_{\text{LV}}(S_t, t)$ is the local volatility,
- W_t is a standard Brownian motion.

Dupire derived an explicit expression to extract the local volatility surface $\sigma_{\text{LV}}(K, T)$ from market-observed European option prices $C(K, T)$:

$$\sigma^2_{\text{LV}}(K, T) = \frac{\frac{\partial C}{\partial T} + qC - qK\frac{\partial C}{\partial K} + rK\frac{\partial C}{\partial K}}{\frac{1}{2}K^2\frac{\partial^2 C}{\partial K^2}}.$$

Usefully, the model fits the entire implied volatility surface perfectly at calibration time. However, the deterministic volatility implies no stochastic features like volatility clustering or future smile dynamics. Additionally, as remarked on [15], several authors tested the hedging performance on local volatility models, with generally poor performances, in some cases even worse than the Black & Scholes one, therefore concluding that the assumption of a deterministic spot volatility may be too simplistic and that stochastic volatility models would be more realistic.

- **Stochastic Local Volatility Model**: The Heston model [13] introduces stochastic variance and can be extended with local volatility scaling, forming the basis for the Stochastic Local Volatility (SLV) model.

  The SLV dynamics combine local and stochastic volatility:

$$\begin{cases} dS_t = (r - q)S_t dt + \sqrt{v_t}\, L(S_t, t)S_t dW^S_t, \\ dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t}dW^v_t, \\ dW^S_t dW^v_t = \rho dt, \end{cases}$$

  where:

  - v_t is the stochastic variance,
  - κ is the rate of mean reversion,

- θ is the long-term variance,
- ξ is the volatility of variance,
- ρ is the correlation between the asset price and variance processes,
- $L(S_t, t)$ is the local volatility adjustment factor.

The local volatility factor $L(S_t, t)$ is chosen such that the model reproduces the observed market implied volatility surface while retaining the stochastic volatility features for realistic future dynamics.

This formulation captures stochastic volatility features (volatility clustering, term structure, skew dynamics) while fitting the implied volatility surface. However, it requires more complex calibration and implementation, with robust numerical methods (e.g., Monte Carlo or PDE solvers).

The different features of the two models are summarized in the below table.

Feature	**Dupire Local Volatility**	**Heston SLV Model**
Volatility Process	Deterministic	Stochastic + Local Adjustment
Smile Fit	Exact at calibration	Exact at calibration, realistic dynamics forward
Volatility Clustering	Absent	Present
Skew Dynamics	Static	Dynamic
Calibration Complexity	Moderate	High

Table 1.1: Summary of Local Volatility Model features vs. Stochastic Local Volatility.

1.7. Conclusions

The first Chapter of this PhD Thesis is devoted to a general overview on Variable Annuities products, with specific focus on the risks arising when dealing with such products. Due to their complex nature and structuring flexibility, no unique risk management approach can be defined for VAs, but it has to be rather tailor made based on the specific risks faced and the objectives set by the Insurance Company. From this point of view, determining the most appropriate hedging strategy is of utmost importance in order to avoid unpleasant surprises over time, since a wrong decision today may be hard to unwind at a reasonable cost in the future. The Chapter does not just focus on hedging, but also covers the main areas impacting the life-cycle of VA products. From this point of view, the structuring of a simple GMAB prototype with different types of possible guarantees is illustrated, as well as other important considerations on non-hedgeable risks and regulatory aspects.

Lastly, also model choice implications are discussed, comparing different solutions, with their respective pros and cons.

2 | Long-Term Options: Hedging Under a Displaced Heston Model

The second chapter is focused on the analysis of how the seller of long-dated put options may effectively put in place a sound pricing and risk management framework in order to address the issues related to missing market data and tradable instruments to protect its balance sheet from undesired losses.

2.1. Introduction

Financial options are nowadays available with a very high degree of liquidity for a wide variety of equity indexes. However, it is quite rare to find liquidity to trade options with long-term expiries. Generally speaking, options up until 5 years expiry are reasonably feasible to trade, more rarely until 10 years. Going further beyond such tenor is quite complex and, although perhaps doable, quite expensive due to bid-ask costs. This lack of market completeness for such long-term tenors may be quite challenging for life-insurance companies dealing with long-dated insurance products ($>$ 10 years) with underlying optionalities, reasons being in a first place the possibility to choose and calibrate proper pricing models due to the lack of market data and secondly, but not less importantly, to source for the best financial instruments to hedge those insurance liabilities.

The most popular model used to price Financial Options is the Black-Scholes [9]. The model relies on the following assumptions:

- The underlying asset price follows a geometric Brownian motion with constant volatility and drift:

$$dS_t = \mu S_t \, dt + \sqrt{v} S_t \, dW_t \tag{2.1}$$

- There are no arbitrage opportunities.

- Markets are frictionless, meaning there are no transaction costs or taxes.
- Trading of the underlying asset is continuous.
- The risk-free interest rate is constant and known.

Based on a portfolio replication approach, the model provide the following closed form solution for the pricing of a generic European call option:

$$C(S,t) = S_0\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2), \tag{2.2}$$

where Φ is the cumulative distribution function of the standard normal distribution, K is the strike price, and T is the time to maturity. The terms d_1 and d_2 are defined as:

$$d_1 = \frac{\ln(S_0/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \tag{2.3}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}. \tag{2.4}$$

Although the model is still widely used as it provides an easy to implement and quick pricing framework, it is based on a too simplistic assumptions, especially for what concerns the volatility which is normally expected to change over time.

A more appropriate representation of reality is provided by [13], with a model which incorporates stochastic volatility. Unlike the Black-Scholes model, which assumes constant volatility, the Heston model allows volatility to fluctuate over time according to a stochastic process. The dynamics of the underlying asset price S_t and its variance v_t use the following system of stochastic differential equations (SDEs):

$$\begin{aligned} dS_t &= \mu S_t\,dt + \sqrt{v_t}S_t\,dW_t^S, \\ dv_t &= \kappa(\theta - v_t)\,dt + \sigma\sqrt{v_t}\,dW_t^v, \end{aligned} \tag{2.5}$$

where:

- μ is the drift rate of the asset price.
- κ is the rate at which v_t reverts to the long-term mean θ.
- θ is the long-term mean of the variance.
- σ is the volatility of the variance process (also known as the vol of vol).
- W_t^S and W_t^v are two correlated Wiener processes with correlation ρ.

The Heston model offers several advantages over the Black-Scholes model:

- It captures the stochastic nature of volatility observed in financial markets.
- It can fit the implied volatility surface more accurately.
- It accommodates the correlation between the asset price and its volatility.

2.2. Option Pricing with the Heston Model

As explained in [16], the European call option price under the Heston model may be expressed in a Black-Scholes-like form as the discounted expected value of the payoff under the risk-neutral measure $\mathbb{Q}$:

$$C(K) = e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+] \tag{2.6}$$

where S_t is a non-dividend paying stock, K is the option strike and $\tau = T - t$ is the residual time to maturity. We can further decompose the above equation as follows:

$$\begin{aligned} C(K) &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[(S_T - K)\mathbf{1}_{S_T>K}] \\ &= e^{-r\tau}\mathbb{E}^{\mathbb{Q}}[S_T\mathbf{1}_{S_T>K}] - Ke^{-r\tau}\mathbb{E}^{\mathbb{Q}}[\mathbf{1}_{S_T>K}] \\ &= S_tP_1 - Ke^{-r\tau}P_2 \end{aligned} \tag{2.7}$$

where $\mathbf{1}$ is the indicator function. The quantities P_1 and P_2 each represent the probability of the call expiring in-the-money, conditional on the value S_t of the stock and on the value v_t of the volatility at time t. Hence

$$P_j = \Pr(\ln S_T > \ln K) \tag{2.8}$$

for $j = 1, 2$. These probabilities are obtained under different probability measures. It can be demonstrated that:

$$C(K) = S_t\mathbb{Q}^S(S_T > K) - Ke^{-r\tau}\mathbb{Q}(S_T > K). \tag{2.9}$$

The measure $\mathbb{Q}$ uses the risk-free bond as the numeraire, while the measure $\mathbb{Q}^S$ uses the stock price S_t.

The quantities P_1 and P_2 can be each computed through the log-stock price characteristic function at time T:

$$P_j = \Pr(\ln S_T > \ln K) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \Re\left[\frac{e^{-i\xi \ln K} f_j(\xi; x, \nu)}{i\xi}\right] d\xi. \tag{2.10}$$

As per the original Heston [13] formulation, the characteristic functions for the log-stock price in T, $x_T = \ln S_T$, are of the log linear form:

$$f_j(\xi; x_t, \nu_t) = \exp(C_j(\tau, \xi) + D_j(\tau, \xi)\nu_t + i\xi x_t) \tag{2.11}$$

where,

$$C_j(\tau, \xi) = r\xi i\tau + \frac{a}{\sigma^2}\left\{(b_j - \rho\sigma\xi i + d)\tau - 2\ln\left[\frac{1 - ge^{d\tau}}{1 - g}\right]\right\}$$

$$D_j(\tau, \xi) = \frac{b_j - \rho\sigma\xi i + d}{\sigma^2}\left[\frac{1 - e^{d\tau}}{1 - ge^{d\tau}}\right],$$

$$g = \frac{b_j - \rho\sigma\xi i + d}{b_j - \rho\sigma\xi i - d}, \qquad d = \sqrt{(\rho\sigma\xi i - b_j)^2 - \sigma^2(2u_j\xi i - \xi^2)}.$$

And:

$$u_1 = \frac{1}{2}, \quad u_2 = -\frac{1}{2}, \quad a = \kappa\theta, \quad b_1 = \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda.$$

As shown in [17], the above formulation may not always behave smoothly, therefore creating numerical integration issues.

[18] and [17] then defined a different formulation equivalent to the Heston's initial one, but with a much better behavior to facilitate numerical integration:

$$\begin{aligned} C_j(\tau, \xi) &= r\xi i\tau + \frac{a}{\sigma^2}\left\{(b_j - \rho\sigma\xi i + d)\tau - 2\ln\left[\frac{1 - ge^{d\tau}}{1 - g}\right]\right\} \\ D_j(\tau, \xi) &= \frac{b_j - \rho\sigma\xi i + d}{\sigma^2}\left[\frac{1 - e^{d\tau}}{1 - ge^{d\tau}}\right], \\ g &= \frac{b_j - \rho\sigma\xi i + d}{b_j - \rho\sigma\xi i - d}, \qquad d = \sqrt{(\rho\sigma\xi i - b_j)^2 - \sigma^2(2u_j\xi i - \xi^2)}. \end{aligned} \tag{2.12}$$

2.3. The Heston Model with Displacement

Although the Heston model provides a useful way to account for stochastic volatility, it still shows some limitations. In particular, it is worth to mention that the variance process v_t may get to 0. This can be avoided by ensuring that the so called Feller condition $2\kappa\theta > \sigma^2$ is satisfied which, however, may deeply reduce the model capability to fit market data. In order to overcome such limitation, here we consider a specific case where we add to the standard Heston model a deterministic displacement factor. As proposed in [19], the new dynamics looks as follows:

$$dS_t = \mu S_t\,dt + \sqrt{v_t + \phi_t} S_t\,dW_t^S, \tag{2.13}$$

$$dv_t = \kappa(\theta - v_t)\,dt + \sigma\sqrt{v_t}\,dW_t^v, \tag{2.14}$$

the model diverges from the standard Heston for the ϕ_t parameter, which is supposed to be non-negative and sufficiently smooth to avoid integration issues. Additionally, it is imposed that $\phi_0 = 0$ and that W_t^S and W_t^v are correlated Wiener processes with time-dependent instantaneous risk-neutral correlation:

$$\operatorname{corr}(dW_t^S, dW_t^v) = \rho\sqrt{\frac{v_t}{v_t + \phi_t}} \tag{2.15}$$

where $\rho \in [-1, 1]$. Such form of correlation guarantees the linearity of the pricing partial differential equation (PDE). $|\rho| \leq 1$ and $\phi_t \geq 0$ ensure that the correlation does not exceed 1 in absolute value. The extension provided in [19] is particularly useful because it basically allows to enhance the standard Heston model with a parameter that can be seen as a floor for the instantaneous variance without the need to impose constraints on the model parameter to satisfy the Feller condition. Furthermore, the additional parameter would increase the flexibility of the model to fit Market option prices. Although the displacement parameter can be defined as time dependent (but deterministic), for our analysis we assume a deterministic and flat parameter to be interpreted as a long run volatility floor.

This particular approach should make the Heston model more realistic while requiring just minor adjustments to the original characteristic function which would then look as follows:

$$\begin{aligned} f_j^{\phi}(\xi; x_t, \nu_t) &= f_j(\xi; x_t, \nu_t) \exp\left(\frac{1}{2}\xi(a_j i - \xi) I_\phi(\tau)\right) \\ &= \exp\left(C_j(\tau, \xi) + D_j(\tau, \xi)\nu_t + i\xi x_t + \frac{1}{2}\xi(a_j i - \xi) I_\phi(\tau)\right) \end{aligned} \tag{2.16}$$

where:

$$I_\phi(\tau) = \int_t^T \phi_u\,du \quad \text{for } 0 \leq t \leq T \tag{2.17}$$

represents the deterministic integrated variance between t and T ($\tau = T - t$).

2.4. Setting-up the Hedging Problem

The derivation of the PDE for the Heston model follows a similar argument than the one used to derive the Black & Scholes PDE. However, in the Heston case the source of risk is coming not just from the underlying's price, but also from the instantaneous variance v_t. To properly apply a portfolio replication approach, we must therefore account for this additional source of randomness. Considering an option with price V, let's form a portfolio Π consisting into Δ units of the underlying stock S and ξ units of another option U which should offset the instantaneous variance risk of the target option V:

$$\Pi_t = \Delta S_t + \xi U_t$$

We would like to set the proper units to buy/sell of the underlying S and the option U such that the price changes dV are completely offset by the changes in the replicating portfolio Π:

$$\begin{aligned} dV &= d\Pi \\ &= \Delta dS + \xi dU \end{aligned} \tag{2.18}$$

To achieve such goal, we shall focus on dV. Differentiating V with respect to the variables S, v and t applying a second-order Taylor series expansion we find that:

$$\begin{aligned} dV = &\frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2 \\ &+ \frac{1}{2}\frac{\partial^2 V}{\partial v^2}(dv)^2 + \frac{\partial^2 V}{\partial S \partial v}dSdv. \end{aligned} \tag{2.19}$$

Under a complete market, it can be demonstrated that the condition $dV = \Delta dS + \xi dU$ can be achieved and that the price of the optimal portfolio Π should therefore match the price of V so that the no-arbitrage condition is satisfied. This portfolio replication argument works perfectly under a complete market assumptions where every security can be infinitesimally traded in continuous time and with no trading costs. This is hardly observable in practice. Most commonly trading takes place at discrete times, being them daily, intra-day, monthly and so on, it would be impossible to trade continuously not just because of trading costs, but also because markets are not continuously open for trading. In the following sections we will build up the underlying's simulation scheme under both $\mathbb{Q}$ and $\mathbb{P}$ measures, to then investigate some possible hedging solutions.

2.5. Data Construction

The first step to perform is to work on a robust framework to implement the Heston model with deterministic and flat displacement and to both, simulate the underlying under $\mathbb{Q}$ and $\mathbb{P}$ as well as to build the necessary pricing/risk sensitivities formulas trough the characteristic function to produce fast metrics and avoid the need to build a nested Monte Carlo scheme.

2.5.1. Simulation Scheme

The underlying's simulation is based on the original Heston formulation with the adjustments explained in section 2.3 necessary to account for the deterministic volatility displacement. It is therefore imposed that $\phi_0 = 0$ and that $\text{corr}(dW_t^S, dW_t^v) = \rho\sqrt{\frac{v_t}{v_t+\phi_t}}$. The simulation uses a Milstein scheme for simulation and a truncates v_t to $1E-5$ in case the variance gets below that level. Here below a pseudo-code for the simulation is reported:

Algorithm 2.1 Heston Model Simulation

```
Input: numPaths, rho, S_0, V_0, T, dt, kappa, theta, sigma, r, q, phi, sim
Initialize variables and set seeds
S[0,:] ← S_0
V[0,:] ← V_0
for t_step ← 1 to num_time do
    Generate random variables Zv and Zs_temp
    Update V[t_step] based on scheme
    V[t_step] ← max(V[t_step], 0.00001)
    Adjust ρ_adj[t_step] based on φ
    Calculate Zs[t_step] using Cholesky decomposition
    Update S[t_step] under risk-neutral measure
end for
Return: S, √V, ρ_adj
```

Apart from ϕ and q which are set arbitrarily, the parameters choice follows the one used in [17], i.e. they are not (yet) calibrated to actual market quotes. As done in the paper, it is remarked that this choice does not satisfy the Feller condition. The list of model parameters and simulation settings is reported in Table 2.1. A test case using market calibrated parameters is provided in Section 3.4.

Parameter	**Symbol**	**Value**
Initial stock price	S_0	100
Time to maturity (years)	T	20
Risk-free rate	r	0.025
Volatility	σ	$\sqrt{0.0175}$
Deterministic displacement	ϕ	0.05^2
Number of time steps per year	freq	48
Total number of steps	nsteps	48×20
Time step size	Δt	$\frac{20}{48\times 20}$
Number of simulations	nsim	100,000
Dividend yield	q	0.015
Rate of mean reversion	κ	1.5768
Long run variance	θ	0.0398
Volatility of volatility	ϵ	0.5751
Correlation between Wiener processes	ρ	-0.5711

Table 2.1: Model Parameters and Simulation Settings.

2.5.2. Monte Carlo and Characteristic Funtion Pricing

We would then investigate how the CF price diverges from the MC one in terms of equivalent Black & Scholes implied volatility and how the displacement may help to increase the flexibility to fit the implied volatility surface.

As shown in Figure 2.1, considering ϕ as set out in Table 2.1, Monte Carlo pricing produces very close results to the ones obtained through the Characteristic Function, especially for long end tenors (3 years onwards), while it under-prices the left side of the moneyness for short end tenors. Even though the model parameters are not calibrated on specific market data, the behaviour of the volatility slices seems to show an interesting pattern with a smile for short end tenors which becomes a skew for long end ones. Also, the model with displacement generally pushes the implied volatilities slightly higher for all the tenors, with particular impact on the area between 90% and 150% for the shorter term tenors.

In order to improve the MC accurancy, we normally simulate a rather granular scheme (e.g. weekly), to then only pick the monthly observations for the hedging exercise (assuming that the rebalancing horizon is monthly).

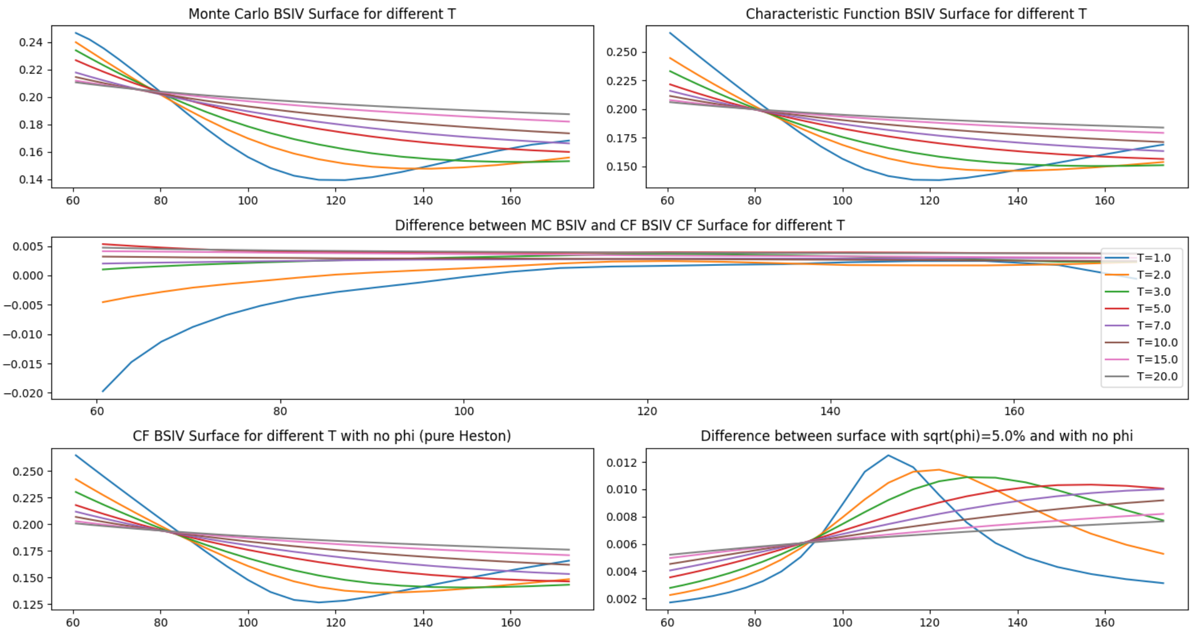


Figure 2.1: Comparison between Monte Carlo and Characteristic Function Implied Volatilities for several tenors and between displaced Heston and Standard Heston.

As per the pricing through Characteristic Function, we use the formulation defined in 2.12 which is supposed to be better behaved and suited for numerical integration than the original Heston formulation. However, it can still happen that the integrand

$$\Re\left[\frac{e^{-i\xi\ln K}f_j(\xi;x,\nu)}{i\xi}\right]$$

is not well behaved and may therefore create numerical integration issues.
There may be normally three main issue with numerical integration:

- The integrand is not defined at the point $\xi = 0$ even though the integration range is $[0,\infty)$ and therefore the integration shall start slightly after 0. This may lead to inaccuracies in case the intgrand is very steep around 0.
- The integrand may show discontinuities.
- The integrand may oscillates considerably and struggle to converge. This would require a very granular grid for numerical integration until a high value of ξ.

The construction of the pricing/partial derivatives calculation framework is of utmost importance in that we need to create a dataset which simulates the underlying and the instantaneous variance for T years and we have to then be able to construct for each scenario node

2.5.3. Tackling Numerical Integration Issues

The main objective of this work is to build a hedging strategy to neutralize the financial risks arising from long-term equity derivatives for which there are no tradable liquid derivatives matching their time to expiry, therefore creating some sizeable rollover risk. In order to achieve this goal, we need to build a big enough dataset where to model, not just the evolution of the underlying and the stochastic volatility, but may also need the price evolution and the risk metrics of the target derivative plus those of a certain amount of financial instruments which can be used for the hedging purpose besides the underlying itself. The purpose of the work is to also keep things as simple as possible, so only a limited number of plain vanilla put options are considered for the hedging objective. Nevertheless, the pricing/risk sensitivity calculation may become considerably tricky due to integration issues, producing substantially different results depending on the effective quadrature method used to approximate the integrals. In general, the main pricing/risk sensitivities calculation issues were found for options with short maturity and deep in/out of the moneynesses. In some cases, for such options, both prices and risk metrics resulted in a highly unstable and unreliable figures. In order to tackle this, we implemented a twofold approach where options with maturity less than 3 years are priced using a more accurate and computationally intense scheme. The 3 years threshold has been arbitrarily chosen based on empirical observation and to safely guarantee that no pricing errors arise in the dataset. The numerical integration scheme is defined as follows:

- **Options with less than 7 years to expiry**: we use Newton-Cotes integration rules, where the domain of integration $[a, b]$ is partitioned through the Simpson's rule to define the abscissas. The points a and b are used as abscissas and the integral domain $(0, \infty)$ is approximated defining $[\xi_{min}, \xi_{max}]$ in a way that ξ_{min} is small enough (normally $1E-5$) and ξ_{max} is big enough to properly capture the integrand (usually we set 200) and to not to lose too much accuracy. As per the Simpson's rule, it uses the following quadratic polynomials to approximate the integral:

$$\int_a^b f(x)\,dx \approx \frac{h}{3}f(x_1) + \frac{4h}{3}\sum_{j=1}^{N/2-1} f(x_{2j}) + \frac{2h}{3}\sum_{j=1}^{N/2} f(x_{2j-1}) + \frac{h}{3}f(x_N), \tag{2.20}$$

  where N represents the number of integration points (usually we set $N = 2500$) and $h = (b-a)/(N-1)$.

- **Options with more than 7 years to expiry**: In this case, we observe that numerical integration should be more stable and accurate, so we speed up considerably the calculations using a Gauss-Laguerre quadrature scheme, which is particularly suited for integrals over the integration domain $(0, \infty)$. Defining N as the number of quadrature points, we have that the abscissas $(x_1, \ldots, x_N)$ are the roots of the Laguerre polynomial $L_N(x)$ of order N defined as

$$L_N(x) = \sum_{k=0}^{N} \frac{(-1)^k}{k!} \binom{N}{k} x^k \tag{2.21}$$

where $\binom{N}{k}$ is the binomial coefficient. There are N roots in all and the weights $(w_1, \ldots, w_N)$ can be obtained through the derivative of $L_N(x)$ evaluated at each of the N abscissas

$$L'_N(x_j) = \sum_{k=1}^{N} \frac{(-1)^k}{(k-1)!} \binom{N}{k} x_j^{k-1} \quad \text{for } j = 1, \ldots, N. \tag{2.22}$$

Each weight is then defined as:

$$w_j = \frac{(n!)^2 e^{x_j}}{x_j [L'_N(x_j)]^2} \quad \text{for } j = 1, \ldots, N. \tag{2.23}$$

It is also remarked that the Laguerre polynomial in Equation 2.21 has $N+1$ terms, but its derivative 2.22 has N terms, which is the correct number of terms required for the approximation of the integral:

$$\int_a^b f(x)\, dx \approx \sum_{k=1}^{N} w_k f(x_k). \tag{2.24}$$

There may be also other schemes to explore for the pricing/risk sensitivities calculation objective. To this extent, a very clear and detailed explanation, together with Matlab code, is provided on Chapter 5 of [16].

The same approach is applied to the calculation of risk sensitivities (Delta, Gamma, Vega, Volga, Vanna, Theta), with the only additional caveat that in case the time to expiry of the option is less than five years and the put option is deep out of the money ($\frac{S_t}{K} \geq 10$), the sensitivities are zeroed to again avoid numerical instability since the put option price should converge towards 0, but it may happen that the integration on the sensitivity calculation results in unreasonably high numbers.

2.6. Hedging Through Risk Sensitivities

Here we would like to analyse how a hedging strategy based on risk sensitivities would perform assuming that the portfolio is rebalanced periodically (e.g. monthly), on a subset of the scenarios generated under the physical probability measure $\mathbb{P}$. In addition to the parameters reported on table 2.1, we set a real world drift $\mu = 0.05$, a market price of volatility risk $\lambda = 0.25$, which then results in adjusted κ and θ

$$\kappa_{RW} = \kappa - \lambda = 1.3268,$$

and

$$\theta_{RW} = \theta \frac{\kappa_{RW} + \lambda}{\kappa_{RW}} = 0.047299.$$

The log-returns distribution of the two simulation schemes is shown on figure 2.2. As expected the real-world one is slightly right-shifted, but it also exhibits a higher yearly standard deviation (25.15% vs 22.52%) more negative skew (-0.83 vs -0.73) and higher kurtosis (1.48 vs 1.17).

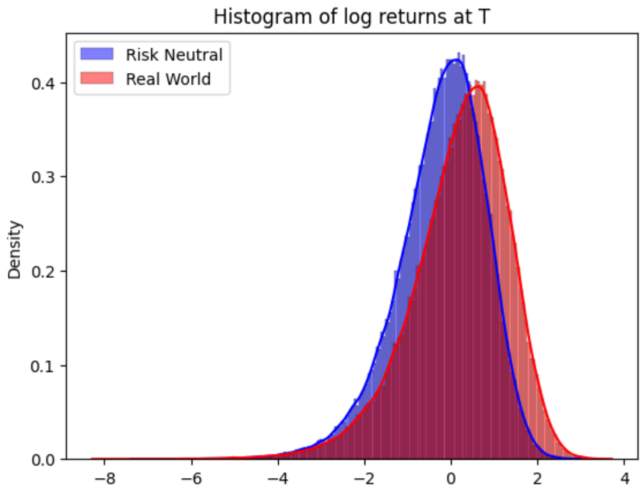


Figure 2.2: Histogram of log-returns from Risk Neutral and Real World scenarios with kernel density approximation.

We are interested into the simulation under the physical measure because we would like to design and assess how a hedging strategy would perform in a real world where the hedge portfolio manager is going to rebalance its positions on a monthly time horizon after observing, at each step, the underlying's price as well as the price of a set of financial securities which may be used to improve the quality of the hedge itself. These sets of prices form the list of state variables and the hedge portfolio manager has the freedom to choose how to allocate its financial resources into such instruments, therefore controlling the number of units to buy/sell (control variables). We perform this analysis on a subset of 1000 scenarios sampled randomly from the generated scenarios.

2.6.1. Theoretical Risk Sensitivity Hedging

The first form of analysis that we provide is a theoretical exercise where we imagine that, somehow, it is possible to perfectly offset, at each rebalancing time, all the sensitivies based on the second-order Taylor series expansion reported on equation 2.19 which we report here below for convenience:

$$
\begin{aligned}
dV = & \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2 \\
& + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}(dv)^2 + \frac{\partial^2 V}{\partial S \partial v}dSdv.
\end{aligned}
$$

This means that we are going to build a theoretical hedge portfolio which targets (and perfectly offsets) Delta (Δ), Gamma (Γ), Vega ($\mathcal{V}$), Volga, Vanna and Theta (Θ). Where:

$$
\Delta = \frac{\partial V}{\partial S}, \quad \Gamma = \frac{\partial^2 V}{\partial S^2}, \quad \Theta = \frac{\partial V}{\partial t}, \quad \mathcal{V} = \frac{\partial V}{\partial v}, \quad Volga = \frac{\partial^2 V}{\partial v^2}, \quad Vanna = \frac{\partial^2 V}{\partial S \partial v}.
$$

As per $\mathcal{V}$, in our model the source of randomness is only coming from the instantaneous variance v_t, so this is the variable that we target directly. This implicitly assumes that the long term variance, the mean reversion speed, the volatility of volatility and the correlation between the stochastic processes are not changing over time.

As mentioned before, we imagine that it is possible at every time step to build a hypothetical hedge portfolio which perfectly tracks the target derivative sensitivities and for which the price change is perfectly matching with the previous period sensitivities scaled by the movement of each single risk factor.

We compound the discretized PnL of this theoretical portfolio over time, as well as the portfolio based on pure Delta hedging, and we compare them with the final payoff of the target derivative.

Statistic	**All Greeks Hedging Portfolio**	**Delta Hedging Portfolio**
Mean	-2.291657193271448	-3.6338812927482387
Median	-1.9997340169551304	-2.716878675790813
Std	1.493172709725844	8.732982926426503
Skew	-0.663709055092253	-0.5777833300713289
Kurtosis	0.5538201189401392	-0.04206571058752615
VaR 95%	-5.0862239943169545	-19.387386039068616
VaR 99%	-6.288215709209006	-27.10207841366426
ES 95%	-5.898315162869906	-23.809433701361395
ES 99%	-6.986027049747871	-30.33148151004287
Min	-8.483038645280647	-36.86767749531609
Max	2.7253949713424586	12.755574696073062

Table 2.2: Summary statistics for the terminal payoffs between the All-Greeks and Delta hedging portfolios

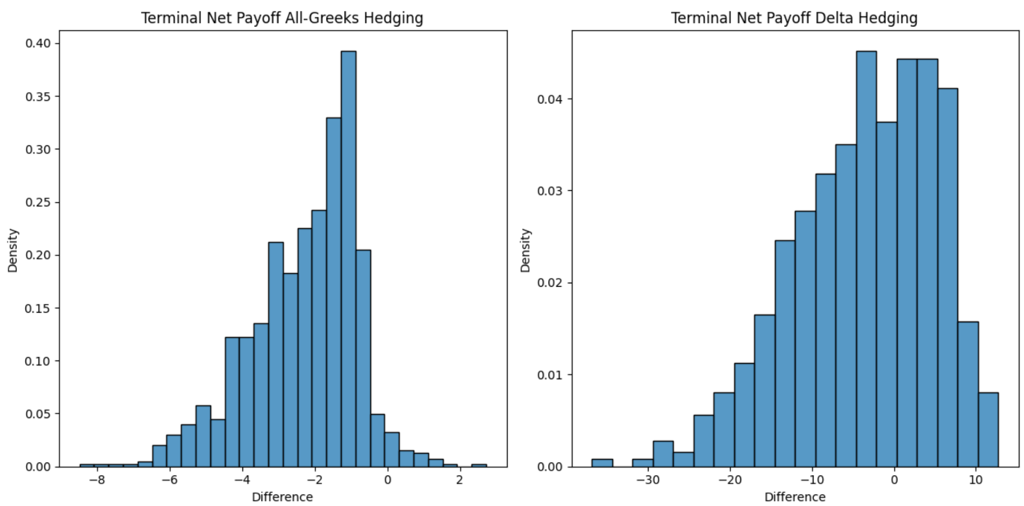


Figure 2.3: Terminal PnL of the All-Greeks theoretical and Delta hedging strategies.

Finally we show the evolution of the net portfolio value at each rebalancing time for the five best and worst performing scenarios (Figure 2.4 and 2.4). Interestingly, although the Target option PV (in the plots referred also as Liability) moves dramatically over the life of the trade, in some cases more than doubling to the drop considerably, the theoretical hedge portfolio manages to smooth out the net changes, resulting in a much more balanced PnL over time.

This view is insightful because it does not just focus on the net payoff at expiry, but also considers what we can call the "Balance Sheet" evolution of the option seller over time. Let's for example imagine a life insurance Company selling such kind of protective put with long dated expiry over a certain fund portfolio investment. The Company wouldn't just focus on the final payoff of the insurance policy, but it would also like to immunize its Balance Sheet from the ongoing PnL as otherwise there may be the strong need to have regular capital injections in order to compensate for the fluctuations of the market consistent valuation of the insurance liabilities.

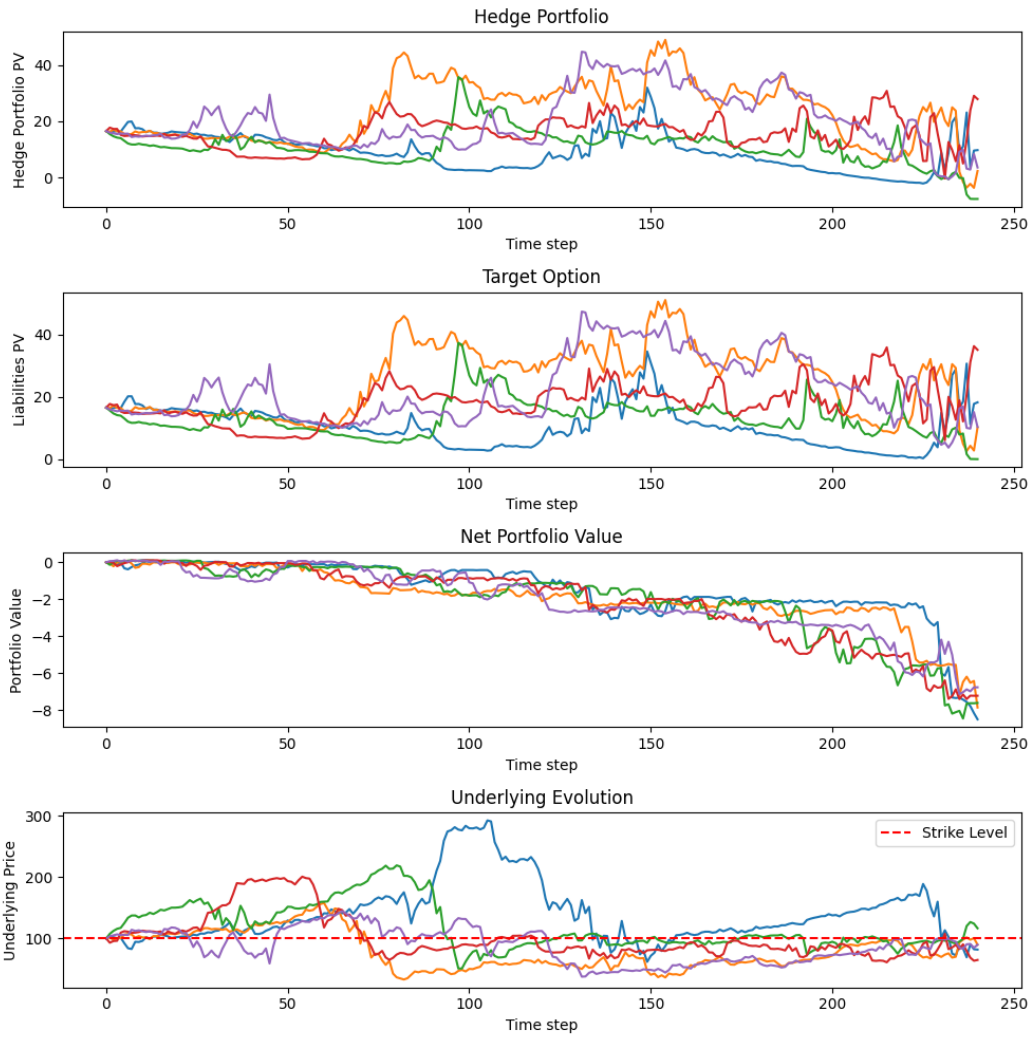


Figure 2.4: Time evolution of PVs for Hedge Portfolio, Target Option, Net Portfolio and Underlying's price for the 5 worst performing scenarios.

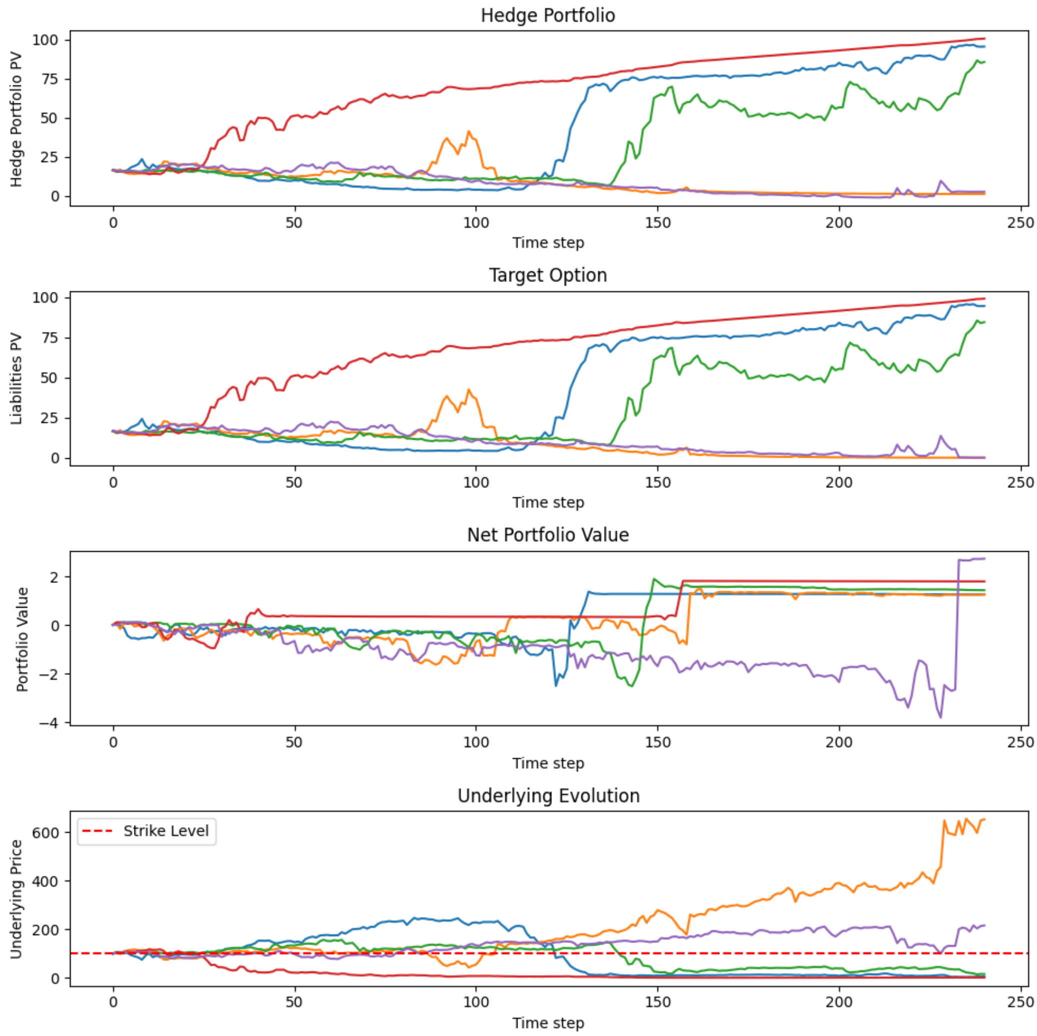


Figure 2.5: Time evolution of PVs for Hedge Portfolio, Target Option, Net Portfolio and Underlying's price for the 5 best performing scenarios.

The analysis, although theoretical for the all-greeks portfolio, shows how enhancing simple Delta hedging by targeting also other first and second order sensitivities may help to effectively reduce the PnL noise at expiry. Although the average PnL result of the All-Greeks based hedging would still result in a loss of 2.29, the standard deviation of the terminal PnL distribution is considerably reduced by around 80%, with a worst case result passing from -36.9 to -8.48 (i.e. around 70% reduction).

This exercise, although theoretical for the All-Greeks case, leads to two main conclusions:

- Targeting first and second order sensitivities may considerably help to reduce PnL noise even under real-world generated scenarios with discrete time rebalancings;
- On the other hand, even assuming that these sensitivities may be perfectly matched, this does not assure that (at least) on average the target derivative payoff will be

neutralized by the hedge portfolio.

2.6.2. Risk Sensitivities Hedging Using Actual Financial Instruments

Having shown that a theoretical portfolio targeting a broader base of risk sensitivities (rather than just the option delta) would produce considerably better results in terms of PnL noise reduction (Figure 2.3 and Table 2.2), we start exploring how to effectively build a portfolio in a way that these additional factors are integrated in the hedging strategy and how to choose the appropriate financial instruments to achieve the goal.
Clearly we cannot base our hedge only on the underlying as this would not give us the needed Gamma/Vega/Volga/Vanna exposures required to hedge the target put. Our goal is to immunize our portfolio from this exposures using the plainest derivatives that can be found on the market, so we consider "potentially liquid" put options that can be easily traded.
By "Potentially liquid" we refer to options for which it would be likely to find a counterparty willing to trade and without the need to pay an extremely high bid-ask spread. Generally, this means that we are no able to trade hedging options with a such long maturity as our target option, since normally liquidity on option markets can be found until 2, 5 or maximum 10 years depending on the underlying.
This translates into the fact that such hedging instruments will not last for the entire life of the target trade to be hedged, consequently creating the need to progressively roll-over the hedge portfolio with new instruments as far as the initial ones approach their expiry. In a real life market environment we have to then tackle the two following issues which can create a considerable basis risk:

- The equivalent Black-Scholes Implied Volatility between the target option and an option with shorter maturity will probably be different and will not necessary change with the same movements over time. For example, if we would hedge a 20 years ATM put option with a 5 years one, we would probably observe from market quoted/extrapolated vols that the two implied volatilities are not the same in t_0 (see for example Figure 2.1) and that in general they will not necessarily move perfectly together over time.

- We can purchase a hedge put with say 5 years and dynamically readjust the appropriate level of units (together with the underlying) to keep the net portfolio immunized from the risk factors, but this may create mismatches due to the fact that we purchase/realize a different volatility than the target one and that in general we will need to buy at a certain future date a new 5 years volatility at the then future levels which may be higher or lower than when the target put is sold/the initial hedge portfolio is created.

We adopted the Heston model with instantaneous variance displacement to capture this potential effects coming from the random behavior of v_t. Although we are not yet testing for cases in which parameters such as the long-term variance, speed of mean reversion and vol of vol change over time which may effectively have an impact on our hedge effectiveness, we start by designing and assessing some option hedging strategies (still based on risk sensitivities neutralization for the time being) to identify the best candidates to achieve our long-term derivative hedging goal.

2.7. Risk Sensitivities Based Hedging Using Financial Options

We now focus on the practical construction of our hedge portfolio. Knowing that risk sensitivities may be used for such goal, we need to extend our trade instrument perimeter to build hedge portfolios which compress the future net PnL distribution and therefore ensuring that the option seller will not incur into huge losses to meet the liability that will have to be funded through equity injections or result into the seller's bankruptcy. This Section assesses few risk sensitivities based hedging strategies as summarized here below:

- Hedging through the underlying plus 1 month ATM puts;
- Hedging through the underlying plus a strategy based on the so called Last Liquid Point (LLP) with same strike as the Target Put;
- Hedging through the underlying plus an ATM strategy based on the so called Last Liquid Point (LLP);
- Hedging through the underlying plus a combined ATM/ITM/OTM strategy based on the LLP.

2.7.1. Hedging Through the Underlying and a 1 Month ATM Puts

The first strategy to be tested comes from the idea that we would like to find the most simple trade to use to complement the underlying to also capture some Gamma/Vega/Volga/Vanna. We consider for this objective a strategy based on the roll-over of very short term (1 month) ATM put options.

Basically the hedger is creating a replicating portfolio composed of the underlying and the 1 month put. Then after the month has passed, the remaining portfolio value plus the payoff of the expired put is used to fund the rebalancing and repurchase of the newly quoted 1 month ATM put.

As remarked earlier, using such short expiring instrument may generate numerical integration instability which would provide inaccurate results for both option prices and risk sensitivities. We therefore use the Newton-Cotes integration for such options, but we use an integration limit of 150 (instead of 200) and a number of integration points of 1500 (instead of (2500) as we all the options to be priced in this case are, by definition, ATM. The pricing and risk sensitivities accuracy for both the hedge portfolio and the target option are key since, otherwise, the optimization of the sensitivities, as well as, the ongoing calculation of the hedge portfolio PnL would also be inaccurate, therefore misleading the assessment of the results. Once all the data (prices and risk sensitivities) have been built, we move forward with the optimization by finding the units of both the underlying asset and the 1M ATM put options which minimize the net portfolio sensitivities. We directly target only Delta, Gamma and Vega as we consider them to be the most relevant greeks and we define our objective function as follows:

$$\begin{aligned}\hat{S^u_{i,j}}, \hat{P^u_{i,j}} = \arg \min_{S^u_{i,j}, P^u_{i,j}} &[(\Delta_{i,j} - S^u_{i,j} - P^u_{i,j} \times \Delta^{1M}_{i,j})^2 + (\Gamma_{i,j} - P^u_{i,j} \times \Gamma^{1M}_{i,j})^2 \\ &+ (\mathcal{V}_{i,j} - P^u_{i,j} \times \mathcal{V}^{1M}_{i,j})^2],\end{aligned} \tag{2.25}$$

for each $i = 1, \ldots, N^{sim}$ and $j = 1, \ldots, N^{steps}$ and where $S^u_{i,j}$ and $P^u_{i,j}$ represent respectively the units to buy/sell of the underlying and of the 1M ATM put option. In the calculation of the net Delta, $S^u_{i,j}$ is implicitly multiplied by 1 as the underlying is a Delta-One asset, while $S^u_{i,j}$ does not appear for Gamma and Vega has it gives no sensitivity to those factors.

Also the pseudo-code is provided here below:

Algorithm 2.2 Compute Hedge Units for Risk Sensitivities Hedging

```
1: Input: hedge_units, i, j
2: NetPtf_Delta ← hedge_units[0] × 1 + hedge_units[1] × Puthedge_Delta_t[i, j] -
   Delta_Target_t[i, j]
3: NetPtf_Gamma ← hedge_units[0] × 0 + hedge_units[1] × Puthedge_Gamma_t[i,
   j] - Gamma_Target_t[i, j]
4: NetPtf_Vega ← hedge_units[0] × 0 + hedge_units[1] × Puthedge_Vega_t[i, j] -
   Vega_Target_t[i, j]
5: error ← (NetPtf_Delta^2) + (NetPtf_Gamma^2) + (NetPtf_Vega^2)
```

As it would happen on a real-life case, the minimization is performed at each rebalancing step for each scenario. Basically the hedge portfolio manager observes the underlying and the quoted puts prices, calculates the target put price and sensitivities for both the target put and the (in this case) one and only chosen as possible hedging candidate and then optimizes the holdings in order to minimize the net portfolio sensitivities.

The optimization is ran using the Limited-memory BFGS algorithm. Figure 2.6 summarizes the net sensitivities per risk factor at each rebalancing time.

Delta-Gamma-Vega are directly targeted in the optimization, so it is not surprising that, besides some few cases, the net exposure is pretty low. As per the other risk factors not directly targeted in the optimization, Vanna looks to be reasonably small, net Volga is systematically negative which is probably due to the fact that for a longer dated option the impact on Vega of a change in volatility gives a higher effect. But what is most concerning is the fact that Theta is deeply and consistently negative. This is a key fact, in that it signals that short maturing options are probably not a good hedging solution as they carry a too strong time decay effect which does not really match with the one of the target put.

We remark that for easiness of interpretation, the net Volga exposures reported in 2.6 and onwards are scaled by 100.

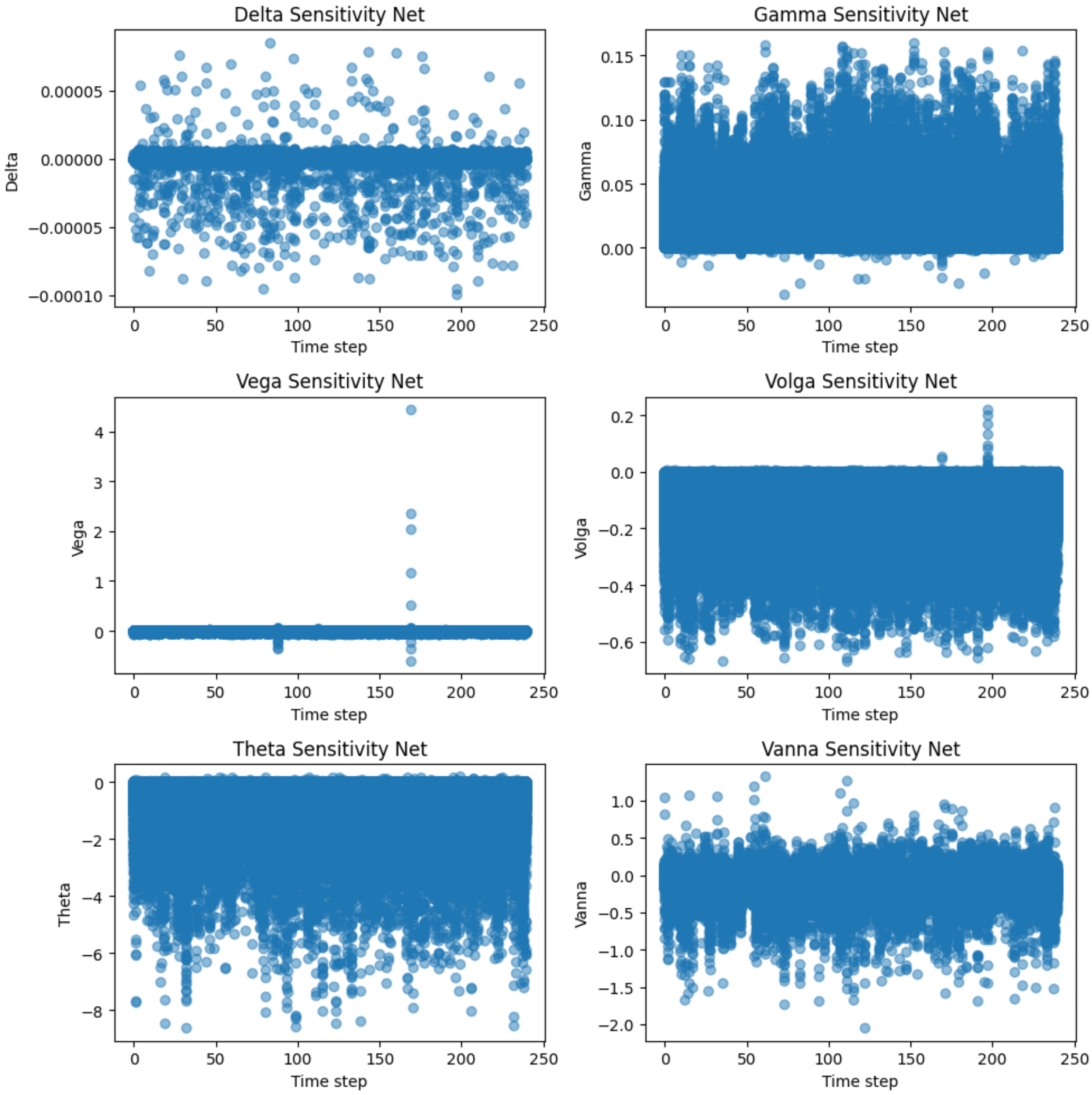


Figure 2.6: Net portfolio sensitivities using a roll-over strategy of 1 month ATM puts.

The terminal results of the 1 month ATM hedging strategy is summarized on Table 2.3. As already hinted, although we managed to create an almost perfect Delta-Gamma-Vega hedged portfolio from a net sensitivities minimization perspective, the inappropriate choice of the hedging instrument produces some extremely bad effectiveness. As a matter of fact, pure Delta hedging as reported on Table 2.2 does considerably outperform such strategy in terms volatility of the terminal PnL distribution, as well as in terms of left-tail events. Not only that, such 1M ATM puts based strategy also looks possibly worse than not to hedge at all. In fact, although the standard deviation of the terminal PnL is slightly below the unehdged one, the 99% Expected Shortfall is bigger than in the unhedged case, while the expected final result of the unhedged case is considerably bigger than the hedged case and actually in more than 50% of the cases, not hedging at

all results in a final profit of 27.3 monetary units, while hedging with this approach would lead us to big terminal profits, but only on a very limited number of cases.
For completeness, the terminal unehdged portfolio PnL is calculated as the target option premium paid in t_0 compounded at the risk free rate until the option expiry (therefore assuming that the premium is parked in some bank account) minus the terminal payoff of the target option. In short, it can be reasonably concluded that this hedging approach

Statistic	**1M ATM Δ-Γ-$\mathcal{V}$ Optimized Portfolio**	**Unhedged Portfolio**
Mean	0.012070626144582122	5.78718877388337
Median	0.11641305982392147	27.27486199371994
Std	22.420677445406678	30.825130551035326
Skew	-0.15038240837330358	-1.098914466347664
Kurtosis	1.36742784824543	-0.28695918304103163
VaR 95%	-35.24033680247617	-57.712694861118905
VaR 99%	-61.339990854002906	-69.18771220906017
ES 95%	-50.878085695334214	-65.26697530870337
ES 99%	-76.0687506803684	-70.33244364069995
Min	-89.05327286887395	-71.68685176453504
Max	78.04505236141672	27.27486199371994

Table 2.3: Summary statistics for the Delta-Gamma-Vega optimized and unhedged portfolios

would be deeply sub-optimal even though, on paper, it may look good from a net sensitivities perspective; and that rather than going for this way, it may make sense to consider to not to hedge at all. Further evidence is offered on Figure 2.7 where the terminal distribution of the PnL is reported.

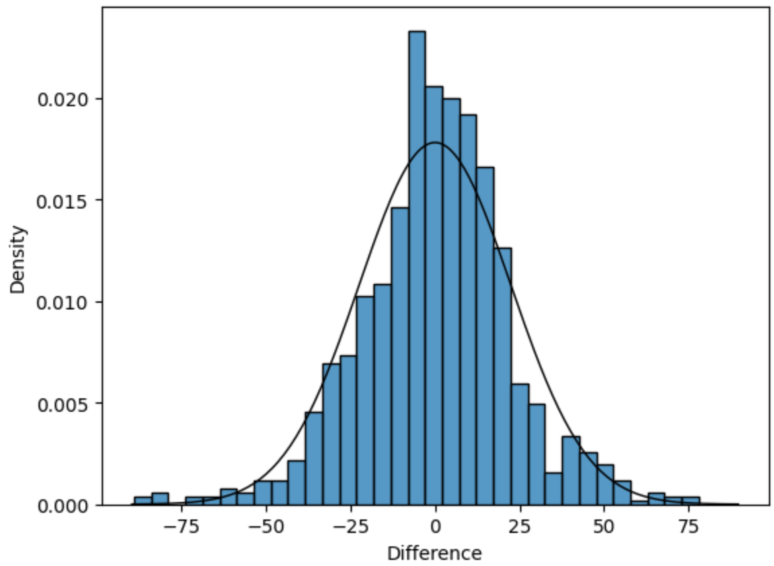


Figure 2.7: Terminal net PnL using a roll-over strategy of 1 month ATM puts.

2.7.2. Hedging Through the Underlying and a Strategy Based on LLP

The analysis conducted on Subsection 2.7.1 has shown how, besides the underlying, the choice of the proper hedging instrument is of utmost importance in order to avoid undesirable effects arising from a potentially good risk sensitivities matching but generally inappropriate fit. Based on the finding that the option time decay and the stochastic volatility dynamic play a major role on the hedge effectiveness of the strategy, we now consider a hedging approach based on the idea of Last Liquid Point (LLP). In a nutshell, we test a strategy which combines the underlying with the longest ATM maturing option available on the market. This assessment is pretty tricky and has to be based on a sound operational experience on financial markets and depends strongly on the option underlying asset.

In our case, we consider the LLP to be 5 years.

In addition to that, as we learnt that both the volatilities are not necessarily moving along across the whole surface and that time decay plays a major role, we don't simply buy and hold the hedging option until its maturity date just rebalancing the units, but we roll-over the option at each rebalancing time, trying to keep a constant maturity and exposure to the (in this case) 5 years volatility tenor. This until the target option reaches the 5 years to maturity, approaching therefore the LLP itself and finally reaching a point where liquidity is decent enough to match the target option tenor with tradable options. At that point, we adapt the time to expiry of the hedging trades with the one of the target.

As a first example, we assess how a strategy based on a 5 years put with the same strike as the target one and rolled over every month would work. We perform two tests:

1. **We buy 1 unit of the hedge put for every unit sold of the target put at every time step without trading the underlying**.

   Knowing that we can only trade up to 5 years tenor, an agent may be tempted to roll-over a hedging strategy using such tenor and matching the strike of the target put using the very same notional. This assumes that buying 1 unit of the strategy based on LLP with same strike as the target put can be enough to reduce our net portfolio risk for 1 unit of the target put. We can think about this strategy as a sort of "super-hedging" where instead of taking position on the underlying, we do it through an option strategy. Although this solution provides a considerable upside in the terminal payoff distribution, this comes with a very high cost on most of the cases. Looking for example at the median of the distribution, we can see that in more than 50% of the cases, this strategy would cost us 5.8 monetary units. We can conclude that this approach is not efficient neither in terms of net sensitivities matching nor in terms of terminal PnL distribution.

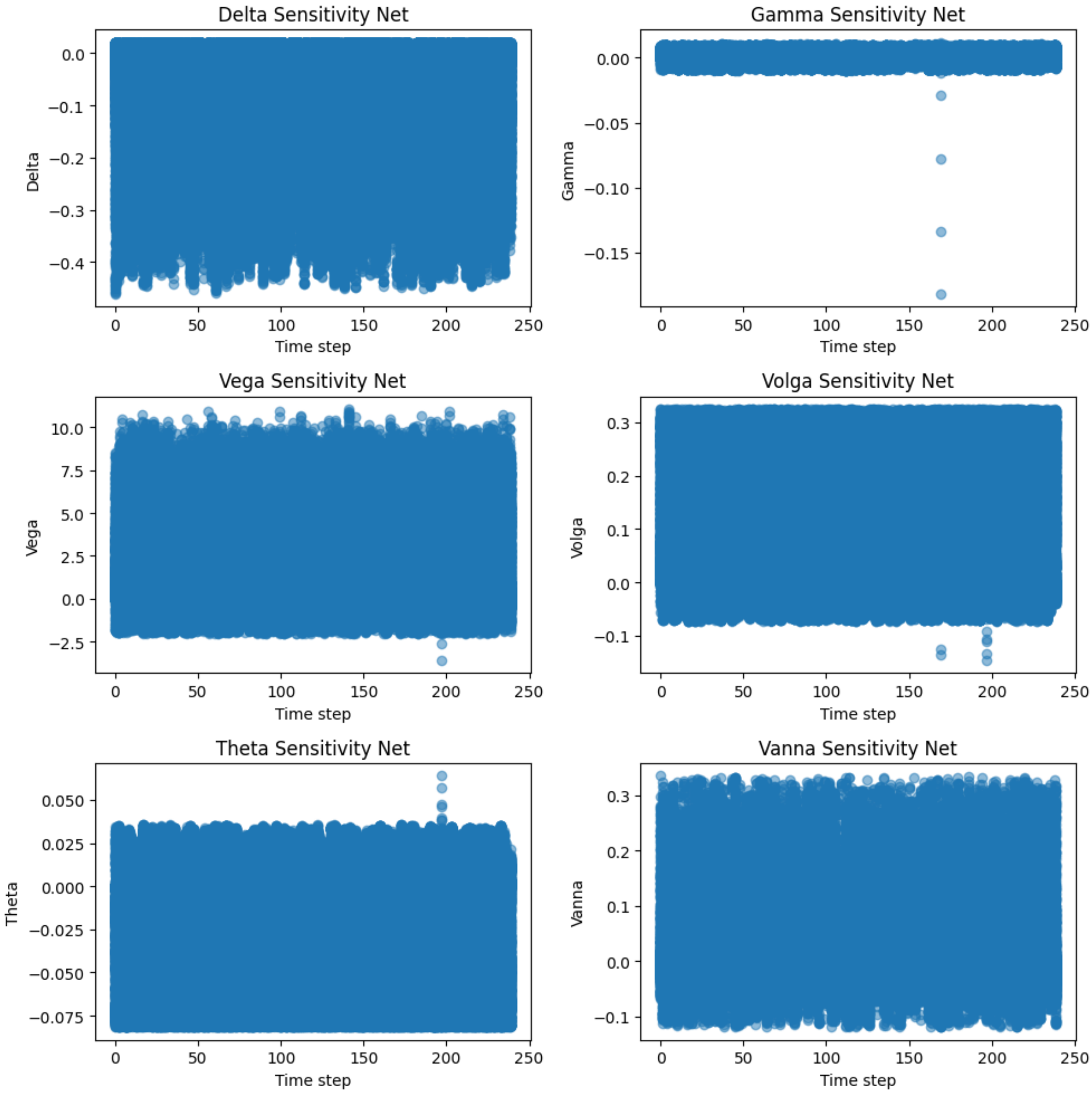


Figure 2.8: Net portfolio sensitivities using a roll-over strategy of 5 years puts with same strike as the target put until the residual maturity of the target put remains above 5 years and matching it afterwards.

Statistic	LLP 5Y With Fixed Strike
Mean	0.062386442298579425
Median	-5.78553157956887
Std	13.596582730250658
Skew	1.405347179472754
Kurtosis	0.9219412580690798
VaR 95%	-12.030315280298975
VaR 99%	-13.096254858158781
ES 95%	-12.781245373640813
ES 99%	-13.47730388395976
Min	-13.992327603228361
Max	43.81432888889326

Table 2.4: Summary statistics for the 5y LLP Strategy with same strike as the target derivative and fixed units.

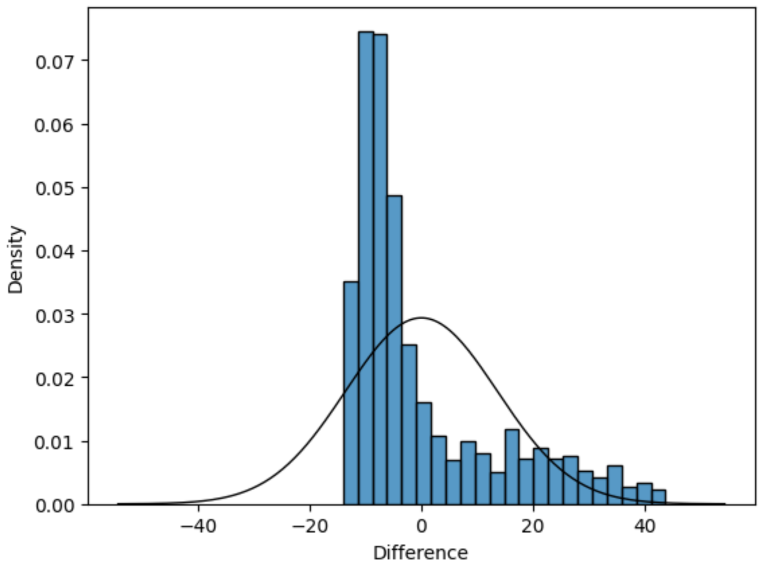


Figure 2.9: Terminal net PnL using a roll-over strategy of 5 years puts with same strike as the target put until the residual maturity of the target put remains above 5 years and matching it afterwards.

2. **We optimize the trade units for both the hedge option and the underlying to minimize the net Delta-Gamma-Vega**.
   We apply a net sensitivities minimization strategy to the above case with the aim to find a strategy to properly combine the underlying with the hedge put. In this case, we obtain some very interesting results, with considerably low standard deviation of the the terminal payoff distribution.

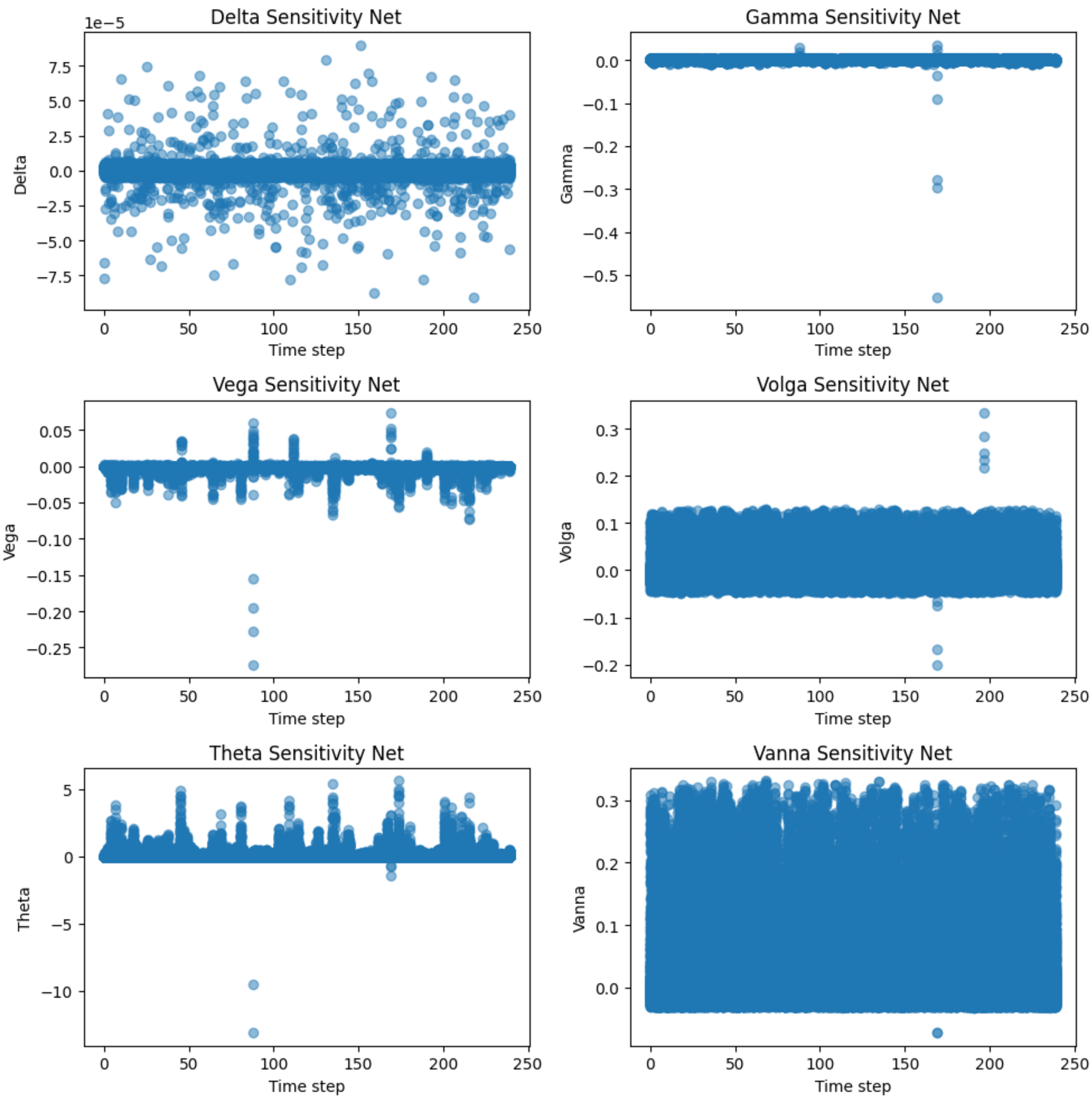


Figure 2.10: Net portfolio sensitivities using a roll-over strategy of 5 years puts with same strike as the target put until the residual maturity of the target put remains above 5 years and matching it afterwards. Units of underlying and the put are set by minimizing the net Delta-Gamma-Vega.

Statistic	LLP 5Y With Fixed Strike Δ-Γ-$\mathcal{V}$ Optimized
Mean	-0.0353180010429264
Median	-0.018944710407527838
Std	0.5304343555078018
Skew	-0.8240003533610307
Kurtosis	7.567697651445833
VaR 95%	-0.8947508623690523
VaR 99%	-1.8521002645284796
ES 95%	-1.4417835461235766
ES 99%	-2.397490570308032
Min	-4.149632050310984
Max	2.3511769741769655

Table 2.5: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the 5 years LLP Strategy with same strike as the target option.

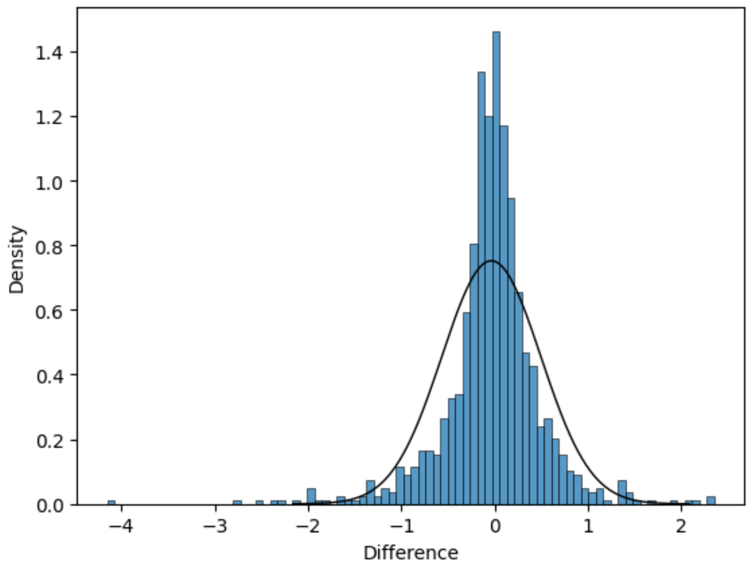


Figure 2.11: Terminal net PnL using a roll-over strategy of 5 years puts with same strike as the target put until the residual maturity of the target put remains above 5 years and matching it afterwards. Units of underlying and the put are set by minimizing the net Delta-Gamma-Vega.

2.7.3. Hedging Through the Underlying and an ATM Strategy Based on LLP

Subsection 2.7.2 provides an insightful result using a LLP based strategy with a fixed strike that reflect the one of our target put option. However, trading the same strike across the entire life of the derivative may be quite difficult, therefore we consider in this

section a more likely tradable strategy using a series of ATM options still based on the last liquid point.
The units selection approach is still based on a Delta-Gamma-Vega net sensitivities minimization.
The pricing and risk sensitivities calculation uses a Gaussian Quadrature for options with more than 3 years expiry and the usual Newton-Cotes otherwise.
Here below, as for the previous cases, net ongoing exposures, summary statistics and the histogram for the terminal PnL distribution are provided. The evidence shows some very good results, with a really high effectiveness in reducing the PnL volatility. Even better than in the theoretical case summarized on Table 2.2.

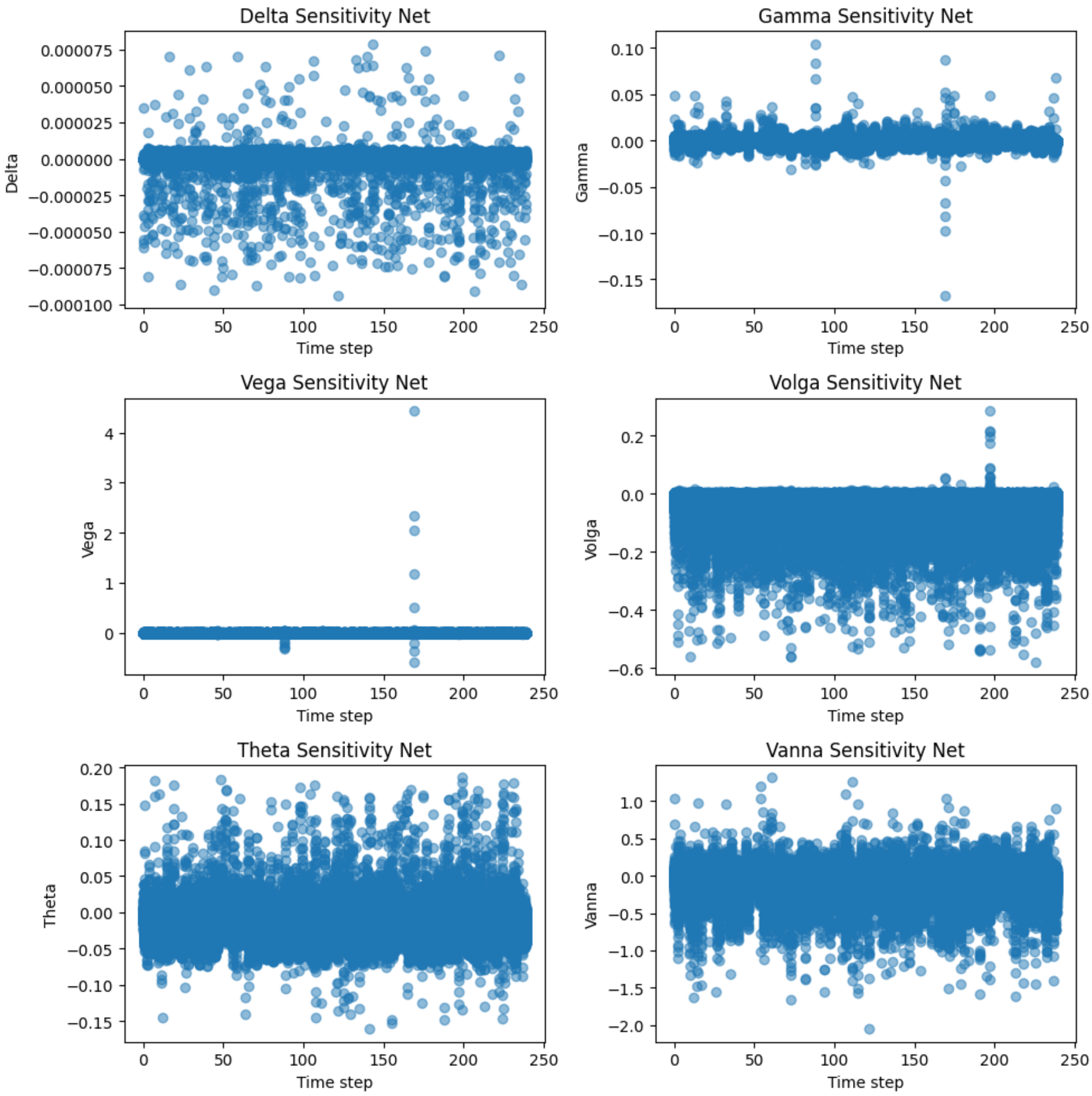


Figure 2.12: Net portfolio sensitivities using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards.

Statistic	5Y ATM LLP Δ-Γ-$\mathcal{V}$ Optimized Portfolio
Mean	-0.05468294513210767
Median	0.0024582539479880783
Std	0.877813711222651
Skew	-0.7765527836784711
Kurtosis	3.5381862253833236
VaR 95%	-1.5163570359061294
VaR 99%	-3.0417943779446377
ES 95%	-2.378295644228037
ES 99%	-3.7335124488565596
Min	-4.458541589237738
Max	3.5893615167902055

Table 2.6: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the 5 Years ATM LLP Strategy

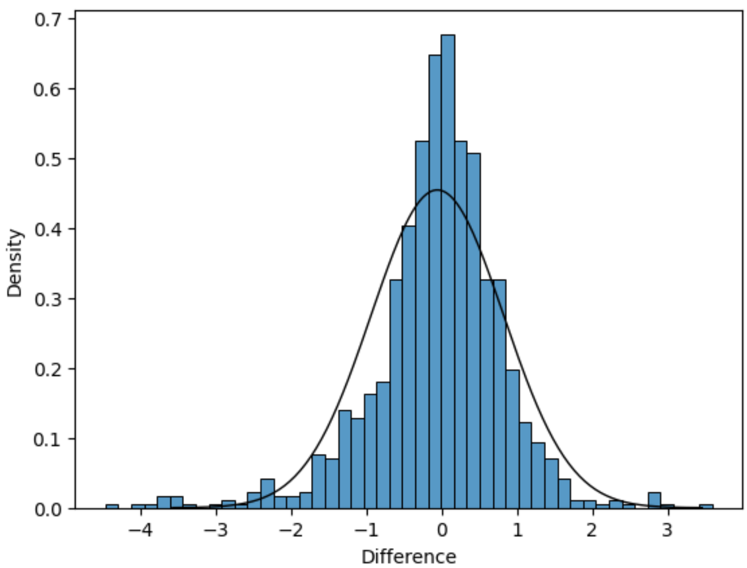


Figure 2.13: Terminal net PnL using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards.

2.7.4. Hedging Through the Underlying and an ATM/ITM/OTM Strategy Based on LLP

After the very promising results obtained on Subsection 2.7.3, we would like to further enhance our possible trade perimeter including also some In the Money and Out of the Money options following the same logic as per the the ATM case.

The idea here is that our risk sensitivities based optimization could be made better by capturing additional parts of the volatility smile/skew related to the LLP. The structure of the hedging problem remains the same, but in this case the optimization variables are the underlying and the ATM/ITM/OTM options units. We also constrained these possible units to remain between the -1/+1 boundaries in order to avoid unreasonable combinations of tens or more hedge puts bought on one side and sold on the other. Again the results in terms of both sensitivities matching and terminal PnL distribution are provided here following.

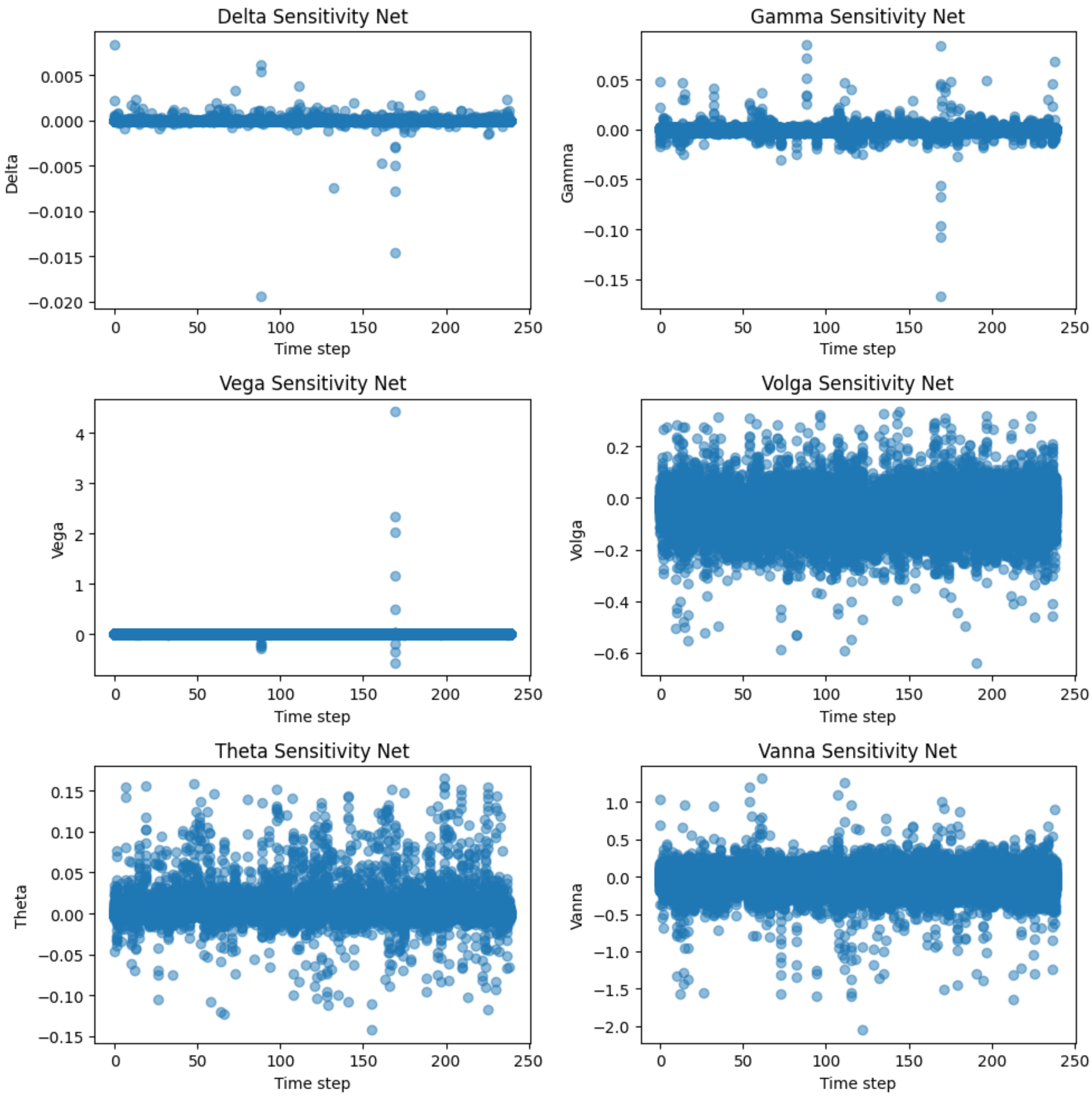


Figure 2.14: Net portfolio sensitivities using a roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards.

Statistic	5Y ATM/ITM/OTM LLP Δ-Γ-$\mathcal{V}$ Optimized Portfolio
Mean	-0.029283118428525413
Median	0.00819456704006882
Std	0.7709353008154314
Skew	-0.4035915562583632
Kurtosis	2.997245297857718
VaR 95%	-1.3798023211211694
VaR 99%	-2.4255521326274776
ES 95%	-1.9668512906021507
ES 99%	-2.9454042357003067
Min	-3.684422601997838
Max	3.557338188448142

Table 2.7: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the 5 Years ATM/ITM/OTM LLP Strategy

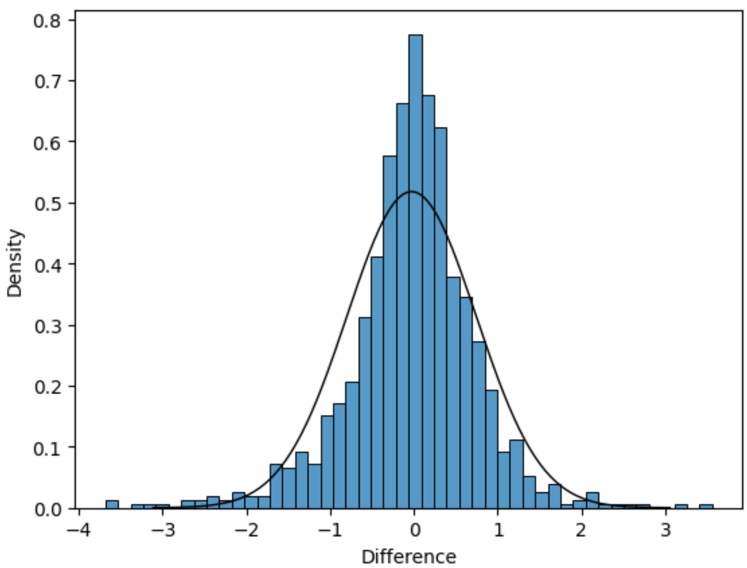


Figure 2.15: Terminal net PnL using a roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards.

Looking at the shape of the terminal PnL distribution, we see that the standard deviation is now considerably below 1 and that the basically over 1000 scenario tested, with a 99% confidence level we should not lose more than 2.5 monetary units. As a reference, the initial fair premium of the target put option calculated using the characteristic function was 16.54 monetary units. Although this experiment does not take into account any explicit form of trading costs, the seller of the put option may factor into the pricing of the derivative this analysis and define a proper mark-up in line with its Risk-Appetite.

The last insight that we provide regarding the hedging strategy based on risk sensitivities optimization is how this approach effectively translates into hedge units.
Here we would like to understand how the hedging strategy effectively behaves over time and over the possible states of the underlying price/stochastic volatility.

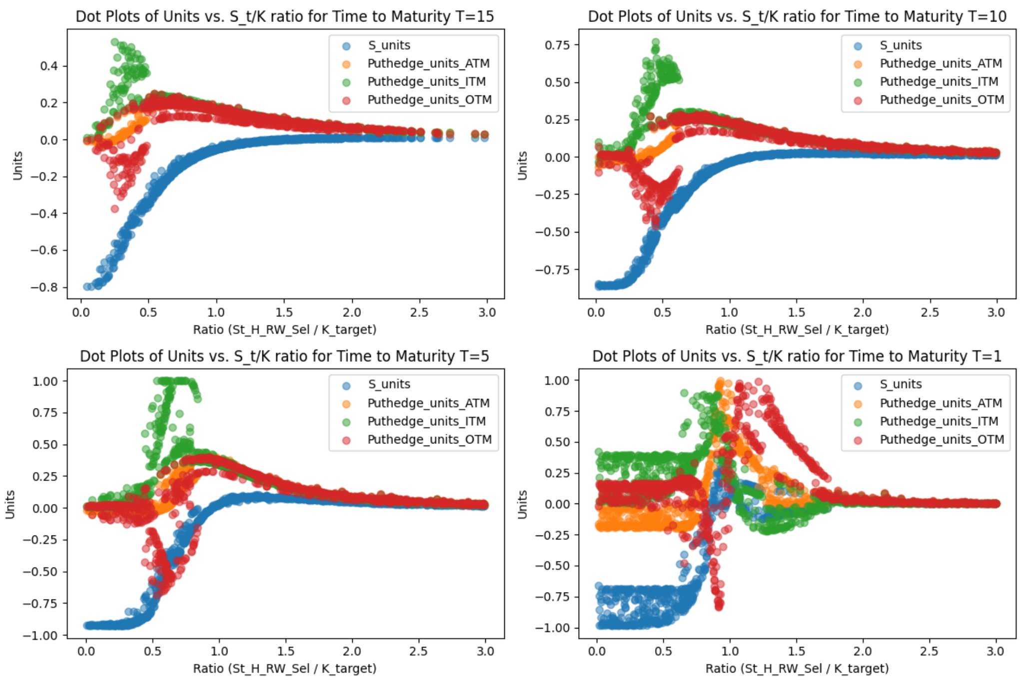


Figure 2.16: Dynamic evolution of the hedge trades units for certain t and considering the ratio S_t/K.

As shown on Figure 2.16, the dynamic of the hedging strategy seems to be reasonably stable when the time to maturity of the option is still far, generally combining a short investment into the underlying which increase sharply as the underlying price falls and the target put increases in value, while it progressively drifts towards zero as far as the target put becomes worthless.
As per the hedge options, they are combined with long and short positions, especially when the underlying starts to decrease and they turn slightly positive when the underlying price approaches the at-the-money level.
The pattern gets more and more confused as time passes, but remains quite clear for the underlying units.
Sometimes (especially when the time to expiry is low) we observe clouds of dots around the same level of the S_t/K ratio. This is likely due to the fact that although the level of S_t may be similar across scenarios, the sensitivities - and therefore the hedge units - are also impacted by the "observed" level of v_t.

2.8. Ongoing Net PnL Analysis

As an additional test, we are interested to know how the hedge portfolio is tracking the target put on an ongoing basis. This is an important test because we not only want to ensure that the terminal PnL stays close to 0 with low variance, but we also want to ensure that the hedge portfolio performs well over time and does not require considerable ongoing equity injections to avoid distressed situations.
We measure this by calculating for each scenario at each point in time the Net Portfolio Value (Hedge Portfolio PV minus Target Put PV) and then we take the first order difference, in order to check how much the net portolio value is changing over time. Normally, it would be good if such differences, which we may also call slippages, would stay around 0 with low standard deviation.

We first consider the case shown in Subsection 2.7.2 where we used to consider a put option based on LLP and a fixed strike to match the target put. This case, although still quite scholastic, made possible to understand that such kind of hedging will not normally work when the distance between the target put time to expiry is still much bigger than the one of the hedge strategy.
We can particularly appreciate such behaviour from Figure 2.17. There we can see that although the mean fluctuates around 0 (recall Table 2.4), the median stays quite below 0 at each rebalancing time more or less for the first 5 years. Also the standard deviation starts very high, to then progressively decrease until finally getting to 0 as soon as the residual time to maturity of the target put approaches the LLP. From then onwards we have perfect replication as we may have expected.

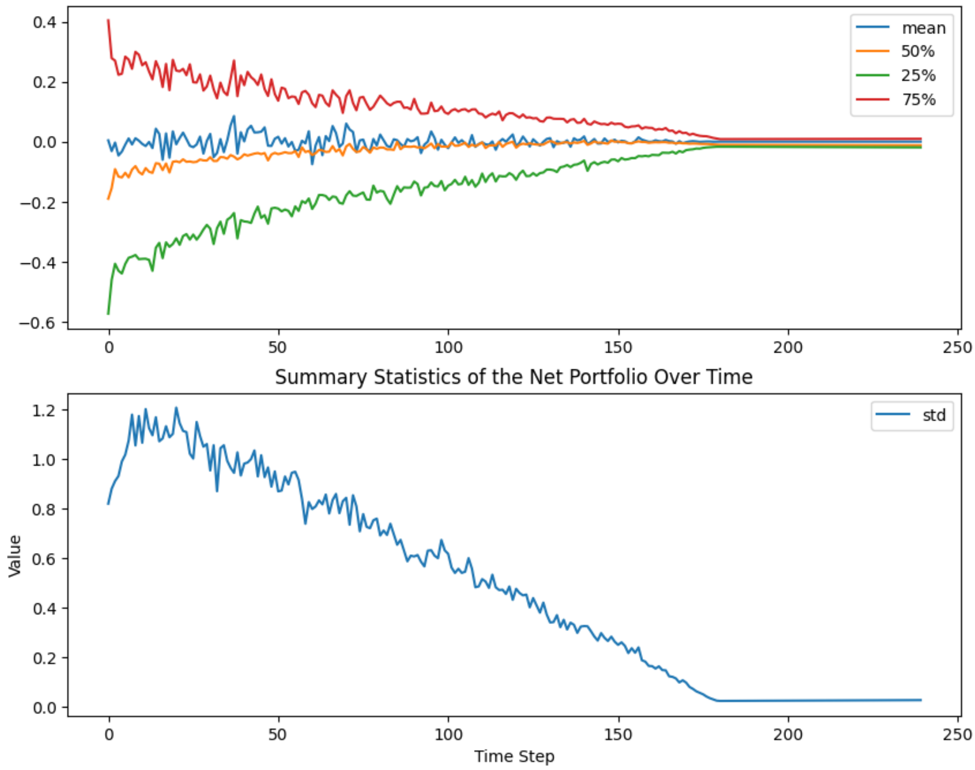


Figure 2.17: Ongoing Net PnL using the 5y LLP strategy with same strike as the target put.

Let's now consider the best case we achieved so far with the 5 year ATM/ITM/OTM options. We obtain the following:

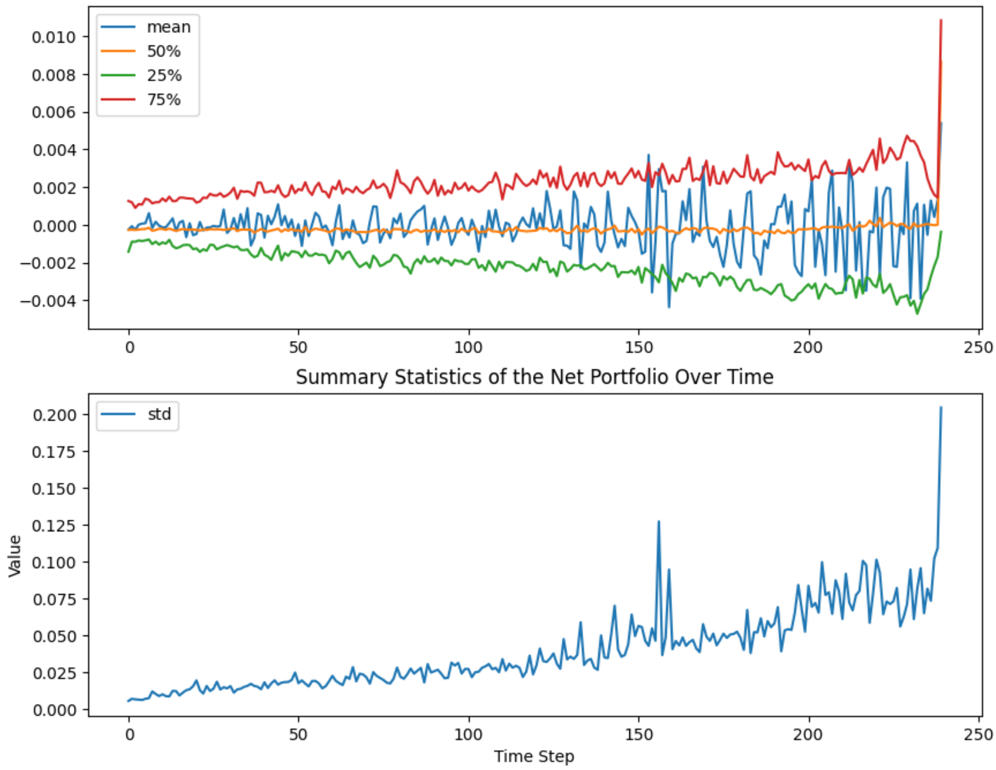


Figure 2.18: Ongoing Net PnL using the 5y LLP strategy with ATM/ITM/OTM options.

In this case we manage to keep mean and median both close to 0, but interestingly, we see the opposite phenomenon as the previous case. Here we start from a very low level of standard deviation, which progressively increase as far we approach the target option maturity.

Interpreting such results is not straightforward.

In a first place, it would look that the hedging at the beginning (say within the first 5 years) would work better than the mid-term one in terms of ongoing standard deviation. We explain this with the fact that at the beginning the underlying price and volatility both start from the same level, while going forward with the projection time we would observe many more levels. This may hint that the more we can capture the volatility skew/smile for the LLP, the better.

Secondly, it would seem that the biggest deviations would occur at the end of the projection, especially as soon as the target option approaches its maturity. This can be explained by the fact that at that point, the quality of the hedge becomes very sensitive to the effective strike of the target put, since at a certain point we are supposed to use that one to meet our liabilities.

With the evidences reported above, we propose a last analysis where we combine the ATM/ITM/OTM strategy based on the 5 years LLP and we match the target option strike as soon as we get to the 5 years left to maturity. This is normally hard to put in place in case the target put is deep ITM/OTM, but it may insightful for illustration purpose. This would be the best strategy achievable so far.

Statistic	5Y ATM/ITM/OTM LLP Optimized + Fixed Strike
Mean	-0.019664358412409603
Median	0.002405000702257354
Std	0.5191422992351714
Skew	-0.8339091749067185
Kurtosis	6.080724327231941
VaR 95%	-0.8455982991199078
VaR 99%	-1.5439796099691638
ES 95%	-1.3181012631151114
ES 99%	-2.223737362989934
Min	-3.5116861339805525
Max	2.2279951326209613

Table 2.8: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the LLP Strategy and ATM/ITM/OTM options until time to expiry of target put above the LLP; fixed strike matching target put afterwards.

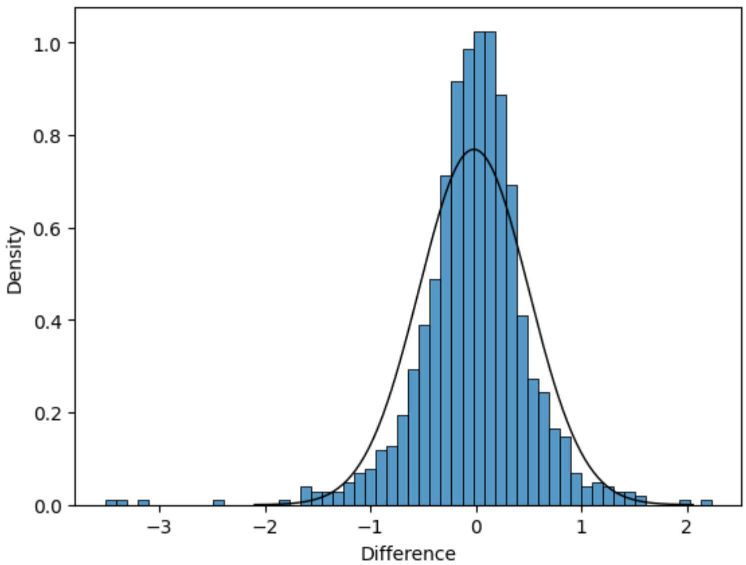


Figure 2.19: Terminal net PnL using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards with same strike as the target put.

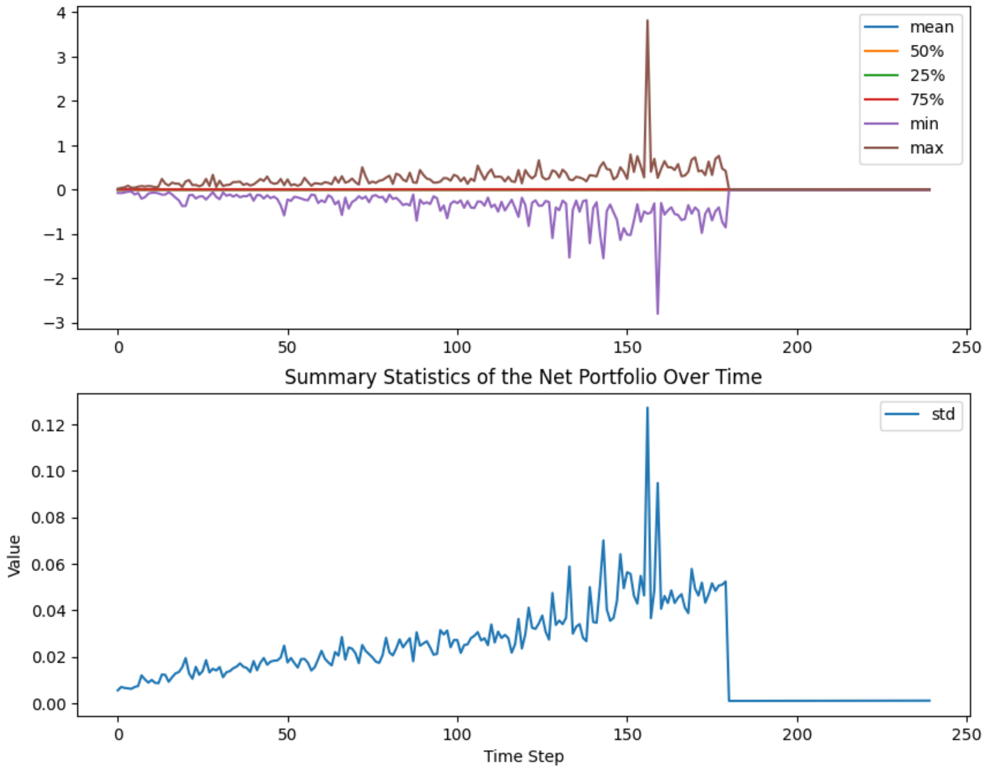


Figure 2.20: Ongoing Net PnL using the 5y LLP strategy with ATM/ITM/OTM options until the remaining time to maturity of the target put approaches the LLP and matching it with same strike afterwards.

2.9. Impact of the LLP on the Hedge Effectiveness

As an agent selling a put option with long-term expiry, we would also like to know how the LLP could potentially affect our ability to hedge our position. In a nutshell, we would like to quantify the impact on the hedging strategy in case the LLP would be shorter/longer than 5 years. We consider for this example the 2 and 10 years cases. Normally we would expect that the longer we can get with the LLP, the better hedge results we should be able to achieve.

As per the 2 years LLP case, we can see that, as expected, the hedge effectiveness deteriorates, although not dramatically.

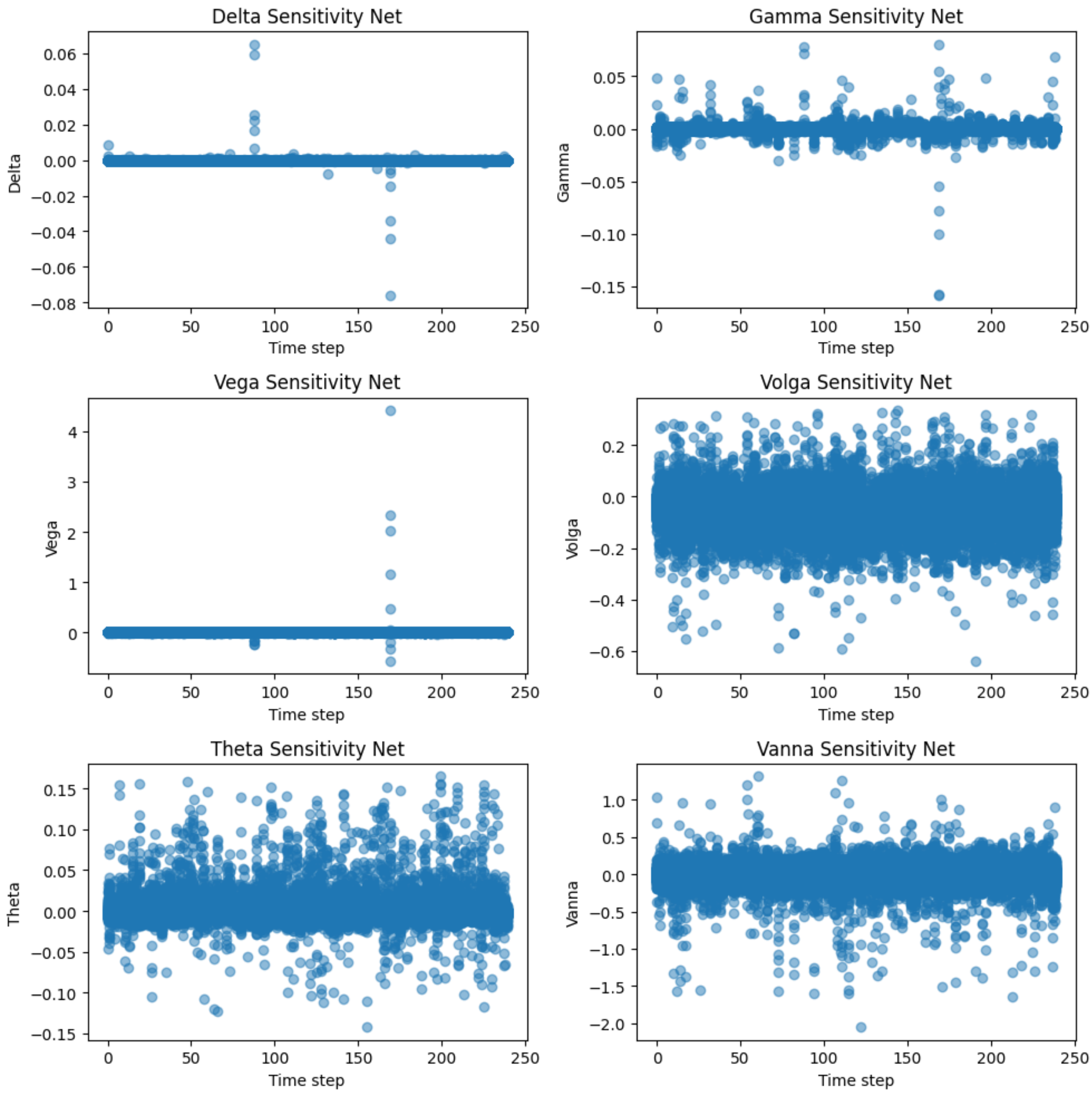


Figure 2.21: Net portfolio sensitivities using a roll-over strategy of 2 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 2 years and matching it afterwards.

Statistic	2Y ATM/ITM/OTM LLP Δ-Γ-$\mathcal{V}$ Optimized Portfolio
Mean	-0.04135835939766332
Median	-0.03087058915674845
Std	0.8722941248438615
Skew	-0.38110591933964544
Kurtosis	2.526669751110021
VaR 95%	-1.5345920206167107
VaR 99%	-2.649158487559372
ES 95%	-2.196878552027806
ES 99%	-3.297915334188282
Min	-4.59274047991926
Max	3.370861724459448

Table 2.9: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the LLP Strategy and ATM/ITM/OTM options with 2 years as LLP.

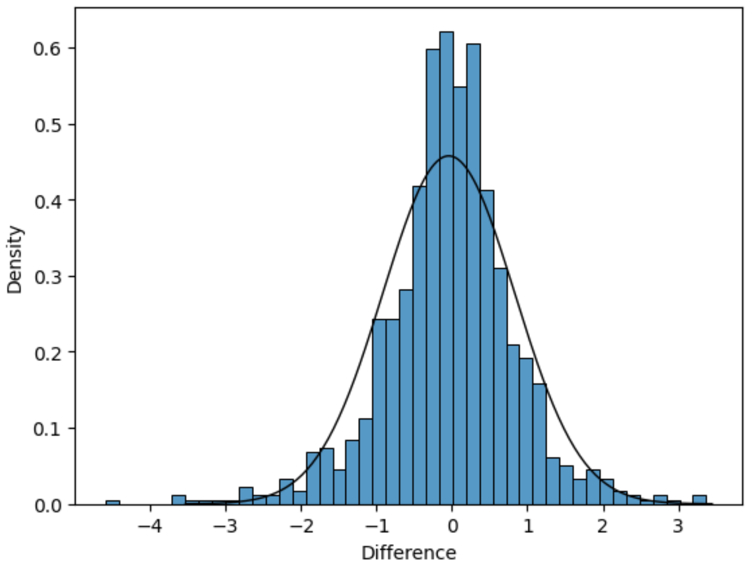


Figure 2.22: Terminal net PnL using a roll-over strategy of 2 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 2 years and matching it afterwards.

Ultimately, we do perform the same test using 10 years hedging options. In this case we observe some improvement with respect to the 2 years case, but very marginal if compared with the 5 years case.

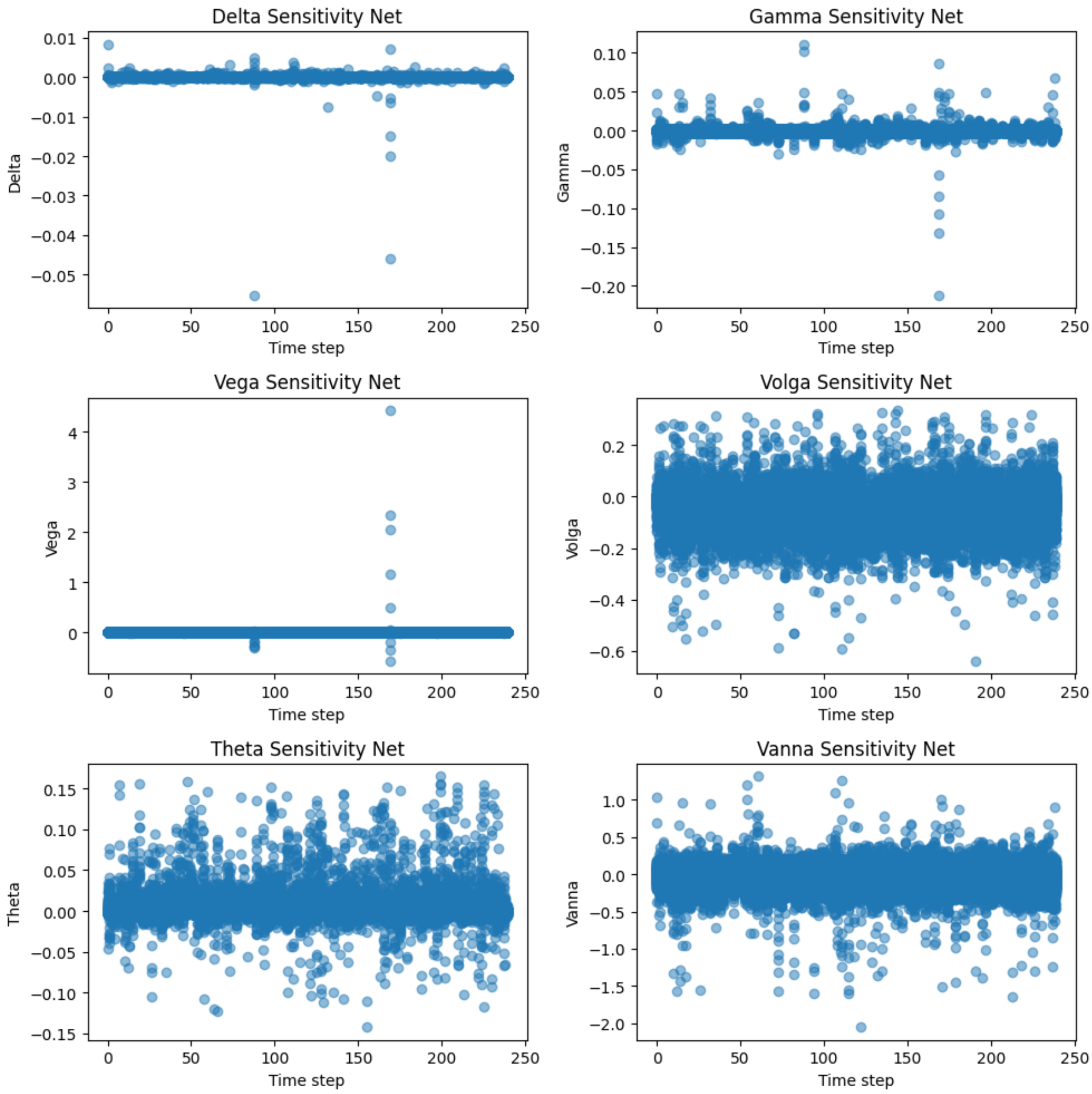


Figure 2.23: Net portfolio sensitivities using a roll-over strategy of 10 years ATM/IT-M/OTM puts until the residual maturity of the target put remains above 10 years and matching it afterwards.

Statistic	**10Y ATM/ITM/OTM LLP Δ-Γ-$\mathcal{V}$ Optimized Portfolio**
Mean	-0.009768483090032126
Median	0.02291611243958837
Std	0.7623631803319204
Skew	-0.3576604655761683
Kurtosis	2.8637045437164876
VaR 95%	-1.2733658559109182
VaR 99%	-2.338802663342751
ES 95%	-1.9205727411000186
ES 99%	-2.8287167388977794
Min	-3.6661666187347604
Max	3.688805452453405

Table 2.10: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the LLP Strategy and ATM/ITM/OTM options with 10 years as LLP.

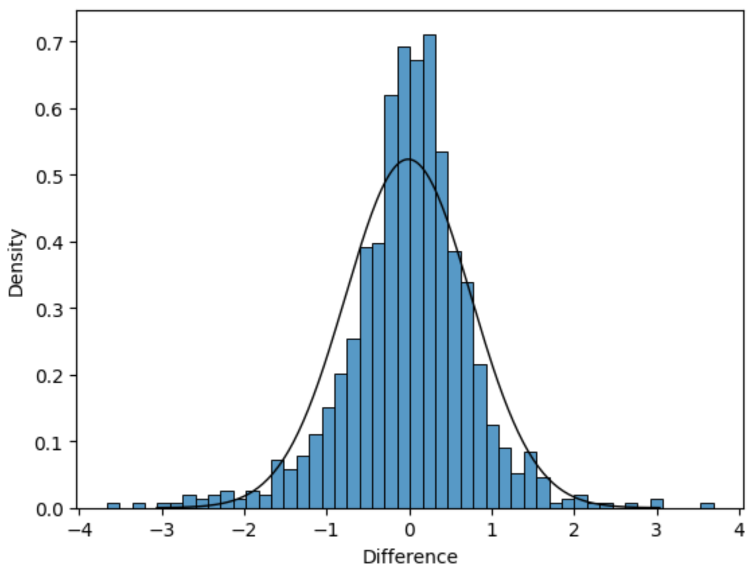


Figure 2.24: Terminal net PnL using a roll-over strategy of 10 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 10 years and matching it afterwards.

Based on the provided evidences, it would seem that the closer we can get to the effective expiry of the target put, the better hedge result we should achieve. Interestingly though the benefit of getting very long expiries seems to progressively fade away as far as manage to get some minimum requirement which does not impact us too much from a time decay perspective. It would seem that for this specific case, an LLP of 2 years would already provide a decent enough fit.

2.10. Greeks Hedging Under Risk Limit Constraints

Although providing very good results, the hedging strategies put in place so far using options would be quite expensive to be implemented on a real life case. We therefore consider a case where we only use a 5y ATM option which is only rebalanced yearly, unless the underlying price would overtake a +/-10% level over the strike price, or the vega exposure would exceed a certain threshold. The delta is then adjusted monthly by only using the underlying. In a real world environment this may help us the Company save quite some trading costs.

In case the put hedge rebalancing is triggered before the yearly occurrence, the strategy would close the existing put hedge position and open a new one with new ATM strike but unchanged time to expiry if the stock price deviates considerably from the strike price, while we would simply adjust the existing put hedge units if the intra-year rebalancing is triggered by the overtaking of the max Vega exposure.

This strategy would aim to reproduce an hedging policy where the vega/gamma/volga/-vanna sensitivities are targeted and adjusted only once per year, unless some specific events occur (underlying price moving far from the strike or vega exposure increasing too much), while the delta is adjusted every month using the underlying. This would considerably reduce the need to trade options regularly, but still trying to ensure a decent enough protection from the non-delta exposures.

We here present two cases. In the first we set a low Vega limit (i.e. 2 monetary units for a single option), while in the second we try a wider limit (i.e. 8).

As for other tested cases, the possible units of the hedging put are constrained between -1 and 1 per target option.

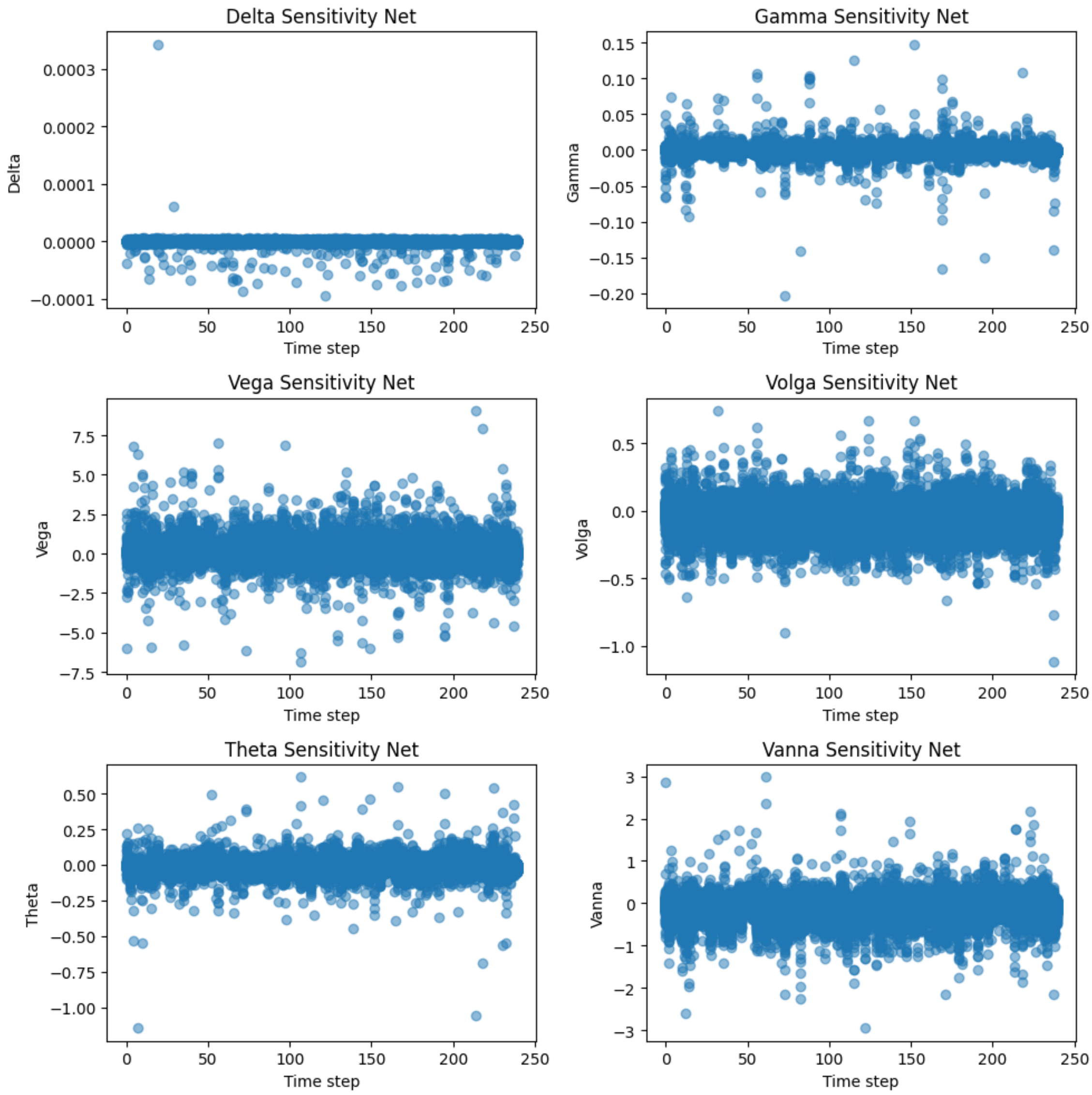


Figure 2.25: Net portfolio sensitivities using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. Option rebalancing done yearly unless stock price goes above/below 10% from the strike or if vega limit (2) is breached.

Statistic	5Y ATM Δ-Γ-$\mathcal{V}$ Optimized Portfolio - Low Vega Limit
Mean	-0.057697505323022924
Median	-0.0020827473793769502
Std	1.0344639608564172
Skew	-0.9021533875439864
Kurtosis	4.670606271889536
VaR 95%	-1.6516227505216527
VaR 99%	-3.3747544204155404
ES 95%	-2.8235144673251
ES 99%	-4.4575659126573335
Min	-6.657303330331739
Max	4.027979044569509

Table 2.11: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the LLP Strategy and rebalanced yearly/if certain events occur.

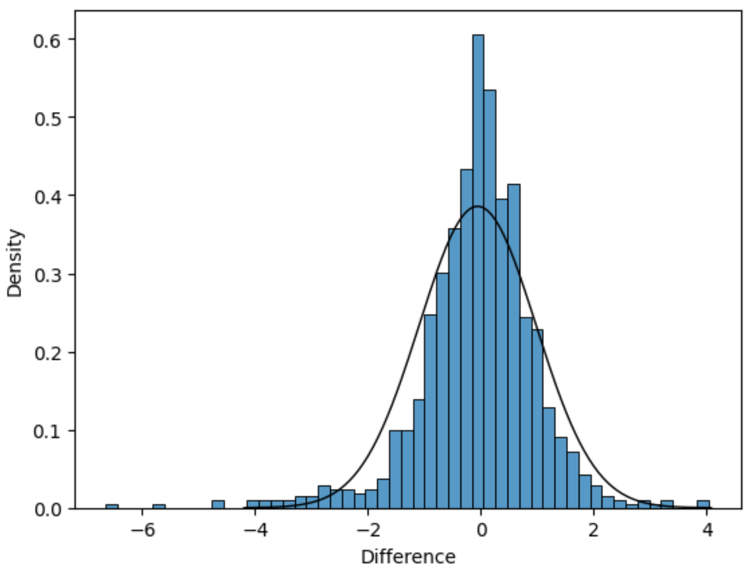


Figure 2.26: Terminal net PnL using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. Option rebalancing dona yearly unless stock price goes above/below 10% from the strike or if vega limit (2) is breached.

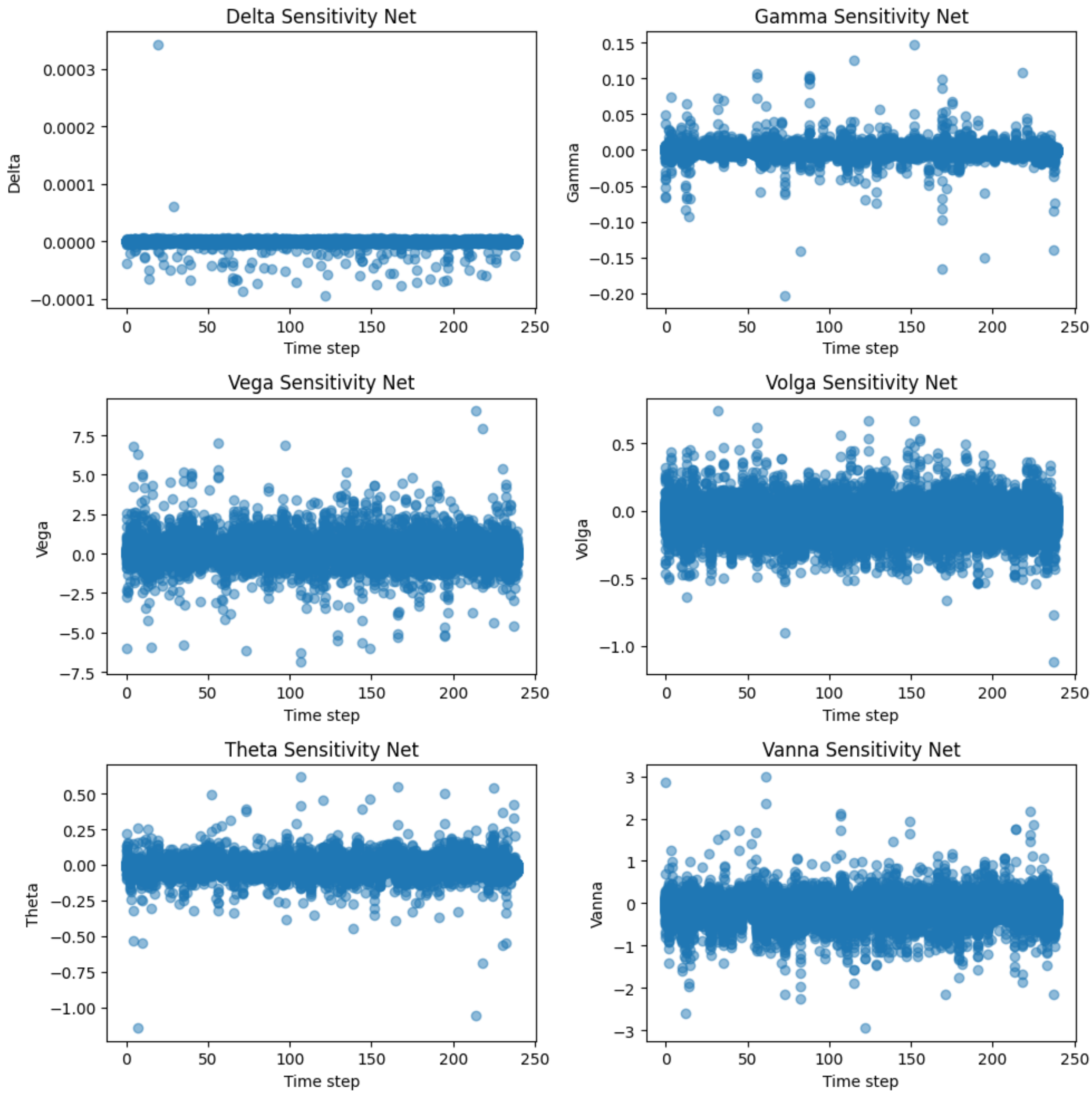


Figure 2.27: Net portfolio sensitivities using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. OPtion rebalancing dona yearly unless stock price goes above/below 10% from the strike or if vega limit (2) is breached.

Statistic	5Y ATM Δ-Γ-$\mathcal{V}$ Optimized Portfolio - High Vega Limit
Mean	-0.05902035327328246
Median	0.03241456011964908
Std	1.116788991302527
Skew	-0.8338647378123742
Kurtosis	4.292068288831388
VaR 95%	-1.9053151855943116
VaR 99%	-3.4952516174074186
ES 95%	-2.9239721128622023
ES 99%	-4.904390528170632
Min	-6.878477707967214
Max	4.431647732317248

Table 2.12: Summary statistics for the Delta-Gamma-Vega optimized portfolio with the LLP Strategy and rebalanced yearly/if certain events occur.

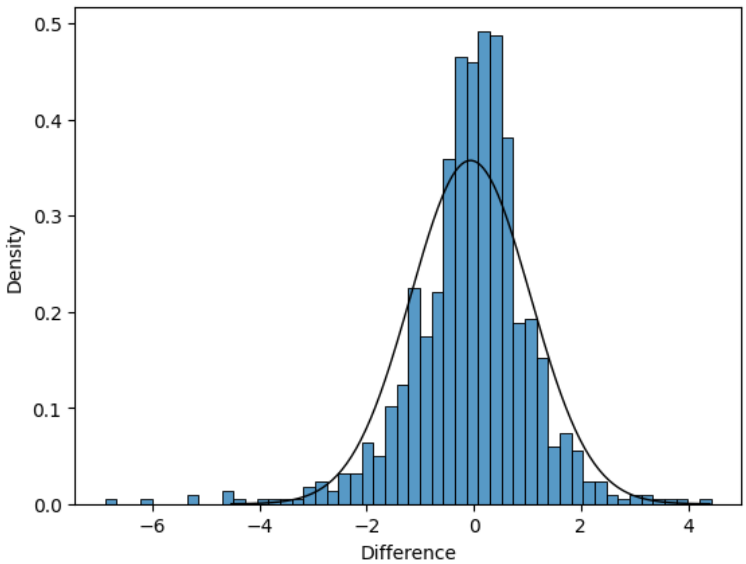


Figure 2.28: Terminal net PnL using a roll-over strategy of 5 years ATM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. Option rebalancing dona yearly unless stock price goes above/below 10% from the strike or if vega limit (8) is breached.

Interestingly, it would appear that lowering considering the vega limit from 8 to 2, while not necessarily adjusting the put hedge instrument, would increase considerably the trading activity on the option, while resulting in a very low benefit from an overall risk reduction perspective. It would rather seem that even if the rebalancing is triggered, the optimizer would often struggle to bring the vega exposure below the desired threshold (see Figure 2.25. This is due to the fact that the put hedge units boundaries may prevent

the possibility to reach the same vega as the target option if the vega of the target put exceeds considerably the one of the hedging one.
Results would improve in case the rebalancing barrier of the ration between the stock price w.r.t. the strike would be lowered (e.g. from 10% to 5%), but there again the move from a high vega limit to a low one wouldn't play a major role.

2.11. Conclusions

In this Chapter, we set the ground to design and test on *ex-ante* basis an efficient hedging strategy based on risk sensitivities minimization under a Displaced Heston Model. We initially defined the model, the simulation scheme and the pricing logic, as well as the dataset creation rules to provide reliable results which would not provide misleading outcomes.
We first analysed from a theoretical point of view the possibility to use risk sensitivities to solve the hedging problem, then we tested several approaches using a variety of instruments/strategies.
We discovered that the pure risk sensitivities cannot work without identifying the appropriate instruments to use. Particularly insightful from this point of view is the case of 1M ATM options which, although providing a decent fit in terms of Delta-Gamma-Vega matching, results in a terminal PnL distribution from a certain point of view (left-tail cases) even worse than the completely unhedged case. The Chapter then defines a specific hedging strategy based on the Last Liquid Point concept, which shows immediately very good results by only using the ATM level and improves further in case also OTM/ITM strategies are introduced.
The Chapter also provides an extensive analysis on the ongoing hedge slippage to further identify the best strategy. It turns then out that the ATM/ITM/OTM strategy effectiveness deteriorates as far as time passes and the time to maturity of the target put reduces. We think this may be due to the fact that the dynamically reset ATM/ITM/OTM strikes may become more and more far from the strike level of the target put. We test a last strategy where we match the target put features as soon as its expiry reaches the LLP. This would be the best hedging strategy found so far (although not really feasible in practice).
We then compared strategies based on 2 and 10 years LLP. For such cases, we observe that the 2 years already provides good hedge results which improve decently if we go with 5 years. We don't observe the same level of improvement between 5 and 10 years though, sign that we would expect the hedge results to stabilize if a minimum expiry requirement for the hedge puts is satisfied.

Finally, a real-life like exercise is assessed in order to reduce as much as possible (both in terms of times over the year and quantities) the option trading. This is done by imposing a yearly rebalancing only of the put hedge units, while monthly adjusting the delta hedging, unless certain specific events occur (vega limit breach/big movements of the underlying price with respect to the strike). Here we find that triggering new trades based on a low vega limit versus a higher one does not massively improve the hedge results (much stronger the impact of lowering the price movement barrier).

3 | Risk Minimization Hedging for Long-Term Contingent Claims

After extensively show how to define efficient hedging strategies in an incomplete market through the usage of risk sensitivities under a displaced Heston model, we would now focus on a different way to approach the hedging problem without using risk sensitivities, but looking for an optimal hedging policy which would allow us to minimize the terminal net PnL (Liabilities minus Assets) of our portfolio at expiry. We would also consider more complex target payoffs than the plain protective put used for testing the risk sensitivities hedging.

3.1. Risk Minimization Hedging

When it comes to contingent claims, the ultimate purpose of hedging is to make sure that the operator dealing with them will have, at the claims expiry time, enough capital to meet any potential liability arising from them.

Under a complete and frictionless market, it should normally be possible to set in place a hedging strategy which would perfectly replicate the claim's payoff at expiry by continuously readjusting the holdings of the elementary securities (e.g. the risk-free bond and the underlying stock) based on the sensitivity of the claim itself to the relevant financial risk drivers. This makes the portfolio only instantly risk free, meaning that if the portfolio rebalancings are only performed within discrete time intervals, then the portfolio immunization will almost for sure not last until the next rebalancing time. Natural consequence is that it would not be possible to completely hedge the intrinsic risk carried by the claim since it will not be possible to perfectly replicate it.

With this in mind, we would like to define an "optimal" hedging strategy to minimize some particular measure of risk within a finite set of hedging times t_0, $t_1, ...,$ T. We may also extend such approach to factor in trading costs in the definition of our optimal strategy.

Let's consider an agent who has entered into a short position on an European contingent claim with payoff at time T given by an $\mathcal{F}_T$-measurable square-integrable random variable H.
The goal for for the agent would be to ensure that the value of its portfolio at time T, net of all capital injections, will be close to H.

Following a notation similar to [20], we define our trading strategy as a combination of stochastic processes $\theta := (\theta^{(B)}, \theta^{(S)}, \theta^{(D_j)})$, with $j = 1, ..., J$ representing then number of derivatives on the underlying asset to be possibly used for hedging purposes, such that $\theta^{(B)} := \{\theta_t^{(B)}\}_{t\in\mathcal{T}}$ is an adapted process and $\theta^{(S)} := \{\theta_t^{(S)}\}_{t\in\mathcal{T}}$ and $\theta^{(D_J)} := \{\theta_t^{(D_j)}\}_{t\in\mathcal{T}}$ are a predictable processes. The $\mathcal{F}_t$-measurable random variables $\theta_t^{(B)}$, $\theta_t^{(S)}$ and $\theta_t^{(D_j)}$ represent, respectively, the number of bond asset shares, the number of shares of stock and units of derivatives with the stock as underlying held over the time period $[t, t+1]$.

We define the value, for $t \in \mathcal{T}$, of the hedge portfolio V_t^θ under the trading strategy θ as:

$$V_t^\theta := \theta_t^{(B)} B_t + \theta_t^{(S)} S_t + \sum_{j=1}^{J} \theta_t^{(D_j)} D_t^j. \tag{3.1}$$

We now introduce the variable G_t^θ as the discounted hedging gain process from time 0 to t, based on the trading strategy θ, and set $G_0^\theta := 0$. For $t \in \mathcal{T} \setminus \{0\}$, we have:

$$G_t^\theta := \sum_{n=1}^{t} \theta_{n-1}^{(S)} \left(B_{n-1}^{-1} S_n - B_{n-1}^{-1} S_{n-1}\right) + \sum_{n=1}^{t}\sum_{j=1}^{J} \theta_{n-1}^{(D_j)} \left(B_{n-1}^{-1} D_{nj} - B_{n-1}^{-1} D_{(n-1)j}\right). \tag{3.2}$$

The discounted hedging cost from time 0 to t is defined by

$$C_t^\theta := B_t^{-1} V_t^\theta - G_t^\theta, \tag{3.3}$$

with $C_0^\theta = V_0^\theta$.

Finally, the incremental discounted hedging cost at time t is defined as $C_t^\theta - C_{t-1}^\theta$. In fact, we can easily see that:

$$\begin{aligned} C_t^\theta - C_{t-1}^\theta &= B_t^{-1} V_t^\theta - B_{t-1}^{-1} V_{t-1}^\theta - \theta_{t-1}^{(S)} \left(B_t^{-1} S_t - B_{t-1}^{-1} S_{t-1} \right) \\ &- \sum_{j=1}^{J} \theta_{t-1}^{(D_j)} \left(B_{t-1}^{-1} D_{tj} - B_{t-1}^{-1} D_{(t-1)j} \right), \\ &= \left(\theta_t^{(B)} - \theta_{t-1}^{(B)} \right) B_t^{-1} + \left(\theta_t^{(S)} - \theta_{t-1}^{(S)} \right) B_t^{-1} S_t \\ &+ \sum_{j=1}^{J} \left(\theta_t^{(D_j)} - \theta_{t-1}^{(D_j)} \right) B_t^{-1} D_{tj}, \end{aligned} \tag{3.4}$$

corresponds to the total outlay needed at time t (in discounted terms) to rebalance the hedging portfolio.

With the above definitions, we can now formalize the terminal financial result for the agent as the difference between the claim's payoff and the terminal value of the hedging portfolio net of incremental costs:

$$\Lambda_T := H - \left(V_T^\theta - B_T \sum_{t=1}^{T} (C_t^\theta - C_{t-1}^\theta) \right), \tag{3.5}$$

This deficit represents the *terminal hedging error*, where θ represents the chosen trading strategy, V_T^θ represents the value of our hedge portfolio at expiry under the hedging strategy θ, $B_T = \exp(rT)$ and C_t^θ is our discounted hedging cost process under the trading strategy θ.
We can equivalently formalize 3.5 as the difference between the final payoff and the sum of the initial hedge portfolio value plus the sum of the discounted hedging gains, compounded for T:

$$\Lambda_T := H - B_T(V_0^\theta + G_T^\theta) \tag{3.6}$$

The objective of the agent can theoretically be defined in many ways with respect to Λ_T, but referring to it as the hedging error naturally implies the fact that the agent would try to make it as close as possible to 0, by properly choosing the optimal hedging policy θ, as well as by charging the counterparty with an appropriate price V_0^θ to fund it. This way, rather than trying to speculate to achieve a positive Λ_T, the agent can then apply a mark-up on the theoretical price V_0^θ based on its risk-appetite/utility function in order to make the business profitable.

With the theoretical framework just defined, we see that the quadratic minimization of the terminal hedging problem can be essentially approached in two ways:

1. **Global Quadratic Hedging**
   We try to find a self-financing hedging strategy θ and an initial capital V_0^θ that solve
   $$\arg\min_{(V_0^\theta,\theta)} \mathbb{E}\left[\Lambda_T^2\right]. \tag{3.7}$$
   Where by self-financing it is required that, under the trading strategy θ, the cost process C_t is constant for all $t \in \mathcal{T}$, i.e. $C_t^\theta = V_0^\theta$.
   This implies that:
   $$\Lambda_T = H - V_T^\theta \tag{3.8}$$
   and
   $$\theta_t^{(B)} = \theta_{t-1}^{(B)} - \left(\theta_t^{(S)} - \theta_{t-1}^{(S)}\right) B_t^{-1} S_t - \sum_{j=1}^{J} \left(\theta_t^{(D_j)} - \theta_{t-1}^{(D_j)}\right) B_t^{-1} D_{tj}, \tag{3.9}$$
   for $t = 1, \ldots, T$ and with $j = 1, \ldots, J$.
   From a practical interpretation point of view, this is to say that no capital injections/withdrawals are allowed during the lifetime of the trade. Basically, after trading the required hedge targets in the underlying and derivative the capital in excess will have to be parked in the risk-free cash account, whereas if our capital is not sufficient to trade the required hedge targets, we need to borrow money at the risk-free rate to fund our hedge portfolio.
   This last implication may be further assessed since it could be possible that the hedging agent may not be able to generally borrow money, or at least not to do so at the risk-free rate. It may be then desirable to design a specific trading strategy aiming to avoid the need to borrow cash to fund the hedge portfolio.

2. **Local Quadratic Hedging** Similarly to what shown in [20] and [21], a more tractable way to solve the terminal payoff minimization problem is to impose that $V_T^\theta = H$. With this approach, Λ_T becomes:
   $$\Lambda_T = B_T \sum_{t=1}^{T} (C_t^\theta - C_{t-1}^\theta), \tag{3.10}$$
   meaning that we rather focus on a series of local optimizations in order to minimize the incremental hedging costs at each trading period.
   This approach requires the cash account units $\theta_t^{(B)}$ to be freely determined at each time $t \in \mathcal{T}$, *de facto* breaking the budget constrain. It may therefore be possible to need capital injections or withdrawals during the life of the trade. The aim is to work recursively backward to minimize the expected squared incremental hedging

costs at each trading period. That is, for $t = T, T-1, \dots, 1$ we need to find the positions $(\theta_{t-1}^{(B)}, \theta_{t-1}^{(S)}, \theta_{t-1}^{(D_j)})$ at time $t-1$ that minimize the following expectation:

$$\mathbb{E}\left[\left(C_t^\theta - C_{t-1}^\theta\right)^2 \middle| \mathcal{F}_{t-1}\right]. \tag{3.11}$$

The positions $(\theta_t^{(B)}, \theta_t^{(S)}, \theta_t^{(D_j)})$ should be used to obtain V_t^θ and satisfy the optimality condition at time t, with $V_T^\theta = H$ imposed for $t = T$. The strategy θ found with this approach will not necessarily be self-financing like the Global Quadratic Hedging, but will rather be mean self-financing, since, by construction, the cost process should be a $\mathbb{P}$-martingale.

Both Global and Local Quadratic Hedging represent viable options to complement hedging strategies based on risk sensitivities (*greeks*). As remarked in [22] the limitation of the *greeks* approach is that it relies on the assumption that markets are complete and that the hedge portfolio can be continuously rebalanced to perfectly replicate the instantaneous price change of the target claim. As soon as these requirements are not met anymore, and hedging can be only take place at discrete times, the choice of a portfolio based on offsetting the risk sensitivities would only produce an instantaneous hedge against the target claim which will most likely not last until the next rebalancing time.
The main consequence of this condition is that it would not possible to totally hedge the intrinsic risk carried by options that cannot be exactly replicated and that we may rather look to find an "optimal" hedging strategy chosen to minimize a particular measure of this risk. On the other hand, Global and Local Quadratic Hedging are not as straightforward as the risk sensitivities approach to implement. Theoretical solutions to solve these problems involve the use of dynamic programming as shown in for example in [23] and again [20].
The recent trend in the financial industry innovation is making considerable usage of artificial intelligence and machine learning. Such techniques look particularly attractive also to cope with such dynamic programming problems where we need to find a "hidden" or "mysterious" rule to take optimal decisions based on the observation of a set of state variables.

3.2. Neural Network Definition

The definition of the Neural Network architecture to use has to be assessed in relation to the kind of problem to solve. In our case, following the same arguments detailed in [24], we consider Long Short-Term Memory (LSTM) neural networks. This specific NN type belongs to the category of recurrent neural networks (RNN), which differs from the feed forward NNs (FFNN) in that they map input sequences to output sequences rather than input vectors to output vectors. Similarly to the FFNN, this is done by applying successive affine and nonlinear transformations to inputs through hidden layers, but adding up self-connections. In practice, the RNN hidden layer is a function of both an input vector from the current time-step and an output vector from the hidden layer of the previous time-step. This results in the dependency of each output to past inputs. As shown by [25] this approach should be more suitable to be applied to the risk mitigation of path dependent contingent claims.

3.2.1. Neural Network Architecture

The model architecture is composed of the following components:

1. **LSTM Layer**: A multi-layer LSTM network that takes sequential input data and processes it across time steps.
2. **Fully Connected Layer (FC)**: A linear layer that maps the hidden states of the LSTM to the final output size.

The model is trained using the Adam optimizer. Early stopping is implemented to halt training if the validation loss does not improve for a certain number of consecutive epochs.

The best model (based on validation loss) is saved during training and used for evaluation on the test set.

3.2.1.1. LSTM Layer

The LSTM layer is defined by:

$$\mathbf{h}_0 = \mathbf{0} \in \mathbb{R}^{N_{\text{layers}} \times B \times H}$$
$$\mathbf{c}_0 = \mathbf{0} \in \mathbb{R}^{N_{\text{layers}} \times B \times H}$$

where B is the batch size, H is the hidden size, and N_{layers} is the number of LSTM layers. The input sequence X has dimensions $[B, T, D]$, where T is the sequence length and D is the input feature size.

The LSTM operation computes hidden states $\mathbf{h}_t$ and cell states $\mathbf{c}_t$ for each time step t as:

$$(\mathbf{h}_t, \mathbf{c}_t) = \text{LSTM}(X_t, (\mathbf{h}_{t-1}, \mathbf{c}_{t-1}))$$

The output of the LSTM at each time step is used as input to the next layer or the fully connected layer.

3.2.1.2. Fully Connected Layer

The fully connected layer applies a linear transformation to the LSTM outputs. For each time step t, the hidden state $\mathbf{h}_t$ is mapped to an output $\mathbf{y}_t \in \mathbb{R}^O$, where O is the output size, defined by:

$$\mathbf{y}_t = \mathbf{W}\mathbf{h}_t + \mathbf{b}$$

where $\mathbf{W} \in \mathbb{R}^{H \times O}$ is the weight matrix and $\mathbf{b} \in \mathbb{R}^O$ is the bias vector.
The combination of LSTM and fully connected layers should enable the model to capture temporal dependencies and output multi-dimensional predictions for each time step.

3.2.2. Forward Pass

The forward pass of the model consists of the following steps:

1. Initialize the hidden states $\mathbf{h}_0$ and cell states $\mathbf{c}_0$ to zero tensors.
2. Pass the input sequence X through the LSTM layers to obtain the hidden states $\mathbf{h}_t$.
3. Apply the fully connected layer to each time step's hidden state to produce the output sequence.
4. Reshape the output to have dimensions $[B, T \times O]$, where $T \times O$ is the flattened output size for each sample in the batch.

Mathematically, the output sequence Y for the input sequence X is given by:

$$Y = \text{FC}(\text{LSTM}(X, (\mathbf{h}_0, \mathbf{c}_0)))$$

3.2.3. Data Input and Pre-processing

The following state-variables are used to train the model:

1. Stock Price Evolution across each scenario for each time-step S_{tn} for $t = 1, \ldots, T$ and $n = 1, \ldots, N$;
2. Time t price evolution of each possible hedge instrument per each scenario D_{jtn}, where $j = 1, \ldots, J$ represents the number of tradable hedging instruments. For simplicity, only plain vanilla european put options are used for our purposes;
3. Residual time to maturity for both the target derivative $T - t$ and hedging instruments T_t^H. For the hedge instruments, the same time to maturity applies for each option and each scenario;

4. Strikes of both the target derivative K_{tn} (time and scenario dependent for cliquet options for example) and of the hedge ones K^H_{jtn};
5. Ultimately, we define our target variable, which in the case of risk minimization hedging is the terminal payoff of the claim, H_n.

As per the construction of the dataset, the input data is prepared by stacking the input features along the last dimension. Given N_{samples} and sequence length T, the input data X is a tensor of shape $[N_{\text{samples}}, T, D]$, where D is the total number of features concatenated from multiple data sources.

The dataset is split into training, validation, and test sets using the following proportions:

- 80% of the dataset is used for training and validation, with a 75%-25% split between the two.
- 20% of the dataset is reserved for testing.

3.2.4. Loss Function

Once the architecture of the NN is set up, we focus on the definition of a suitable loss function that would meet our goal. We follow the Global (Quadratic) Risk Minimization approach as set out in 3.7. A pseudo-code to set up the loss function is provided in 3.1. In this case we took for simplicity the case where the hedging is made using the underlying and one hedging option/option strategy. This can be extended by adding more options/option strategies. In our setup, we leverage on the findings shown in 2.7.3, where the strategy of rolling over, on a regular basis (monthly in our case), an ATM option matching the last liquid point maturity proved to be very effective. The loss function is based on the following variables:

- **Control_Vars**: The object used to input the control variables which should be then optimized by the NN, i.e. the time 0 target derivative price and the ongoing units to buy/sell on the underlying stock price and the hedge derivative (in this case the ATM option strategy based on LLP).
- **Put_target_T**: The payoff of the target derivative at expiry.
- **St**: The evolution of the stock price over time.
- **Puthedge_Price_t**: Time t price of the option strategy.
- **Puthedge_Price_t1**: Price in $t+1$ of the option traded in t and to be liquidated and replaced with the new option matching the new last liquid point.

Algorithm 3.1 Custom Loss Function for Global Risk Minimization Hedging

Input: Control_Vars, Put_target_T, St, Puthedge_Price_t, Puthedge_Price_t1, r, dt, q
Extract $Put\_target\_price\_t0$, S_units, $Puthedge\_units$, from $Control\_Vars$
Initialize $HedgePtf[0] \leftarrow Put\_target\_price\_t0[0]$
for $j = 1$ to seq_length **do**
  Update $HedgePtf[j]$:
    $(HedgePtf[j-1] - (Puthedge\_units[j-1] \times Puthedge\_Price[j-1])) \times \exp(r \times dt)$
    $+(Puthedge\_units[j-1] \times Puthedge\_Price\_t1[j])$
    $+(S\_units[j-1] \times (St[j] - St[j-1]))$
    $-S\_units[j-1] \times St[j-1] \times (\exp((r-q) \times dt) - 1)$
end for
$error \leftarrow \sum(HedgePtf[seq\_length] - Put\_Target\_T)^2$
return $error$

- **r, dt, q**: Respectively, the risk-free rate, the time discretization interval (1/12 for monthly case) and the stock's dividend yield.

After extracting the control variables, the function initializes the value in $t = 0$ of the hedge portfolio, which should match the premium exchanged with the counterparty for trading the derivative. Then the scenario-time iteration starts and the hedge portfolio value starts to evolve. In $t+1$, the hedge portfolio value can be expressed as the sum of the following components:

- We first compound by the risk-free rate the spare cash remained after buying/selling the hedge derivative. Here we do not include directly the underlying as it is treated separately;
- Secondly we compute the residual value of the hedge derivative at time $t+1$ times the units traded in t. With this component we capture the PnL of the derivative used for hedging purposes;
- The third element is the pure PnL given by the change in the stock price scaled by the units trade;
- The stock PnL has to be adjusted for the "positive/negative" cost of funding and the dividend yield. Assuming we can borrow/invest at the risk-free rate r and that we have no short-selling cost, if we (short) sell the stock price, we can then invest the amount received from the sale at the risk-free rate, but we need to deduct the dividends payed by the stock which should then be transferred back to the counterparty from whom we borrowed it. Should we buy the stock, this piece of

function would then treat r as the funding cost we will have to pay to borrow the money, netted by the dividend yield produced by the stock.

The iteration scheme goes forward for each scenario up to the expiry of the target derivative. There, the final quadratic hedging error is calculated as the squared difference between the T value of the hedge portfolio and the target derivative payoff and finally summed over the whole number of used scenarios.

With this loss function, the Neural Network is going to optimize the control variables in order to minimize the terminal quadratic hedging error. Also, the function can be edited in many ways to meet different purposes. Few more examples can be:

- Using a linear hedging error measure instead of a quadratic one. As shown in [22], we may observe considerably differences in the selected hedging strategy comparing linear vs. quadratic risk minimization, normally due to the fact that the quadratic case will tend to penalize more the extreme residuals reducing the quality of the fitting under most scenarios.
- As shown in [24], the loss function can also be adjusted to only penalize cases where the error is negative. This would also be easily implementable in our loss function as it simply requires to not to include in the error calculation the cases where the terminal value of the hedge portfolio is higher than the payoff of the target claim.
- The loss function could also be fine tuned to ensure a minimum return for the seller of the target derivatives. Say for example that we would like to target a certain level of profit at expiry, being it in absolute term or relative to some other parameter, then we can factor this in into our terminal hedging error and the time t_0 price of the target derivative as well as the overall hedging strategy may be (efficiently) adjusted to meet such goal.

3.3. Numerical Results

We move now to the analysis of the results obtained with the NN based hedging. The training and testing is based on simulated data as defined in 2.5 using the displaced Heston model with stochastic volatility and displacement, the interest rate and dividend yield dynamics are still deterministic and constant. We keep testing our strategy simulating under $\mathbb{P}$ and pricing over time the plain vanilla options used for hedging purposes through the inverse Fourier transform as described in 2.2 (see also [26] and [17] for more details).

3.3.1. NN Hedging Results for Plain Vanilla European Put Option

Let's start over from the test performed in 2.7.4 where ATM, ITM and OTM put options with 5 years expiry (plus the underlying) have been used to build the hedging strategy through the net *greeks* exposure minimization. We consider this case as our performance benchmark as it represented the most reasonable balance between a complex enough hedging strategy using several hedging instruments with a relatively straightforward methodology to set it in place. This comparison represents a starting point, in that we would shift from the pure European plain vanilla case to some more complex payoffs. Ideally we would also develop a framework which could be possibly applied directly to a banking/insurance book (i.e. without the need to train a model for each single derivative as this would massively increase the computation time).

3.3.1.1. Simplified Training using the Known Model Price as V_0^θ

We start by slightly simplifying the hedging problem by adjusting the loss function 3.12 by imposing that $V_0^\theta = V_0^{model}$, therefore trying to test if the NN would be able to find a risk minimization strategy which would be effective enough under the pure model pricing. The purpose of this test is twofold: it serves to test the capability of the NN to adapt to our problem and to make a first comparison with the *greeks* hedging. As we have seen, the latter form of hedging proved to be quite effective, so the hope would be that the NN hedging strategy would "learn" to hedge in a way at least as good as the sensitivities approach. For completeness of notation, the new loss function becomes:

$$\Lambda_T := H - B_T(V_0^{model} + G_T^\theta) \tag{3.12}$$

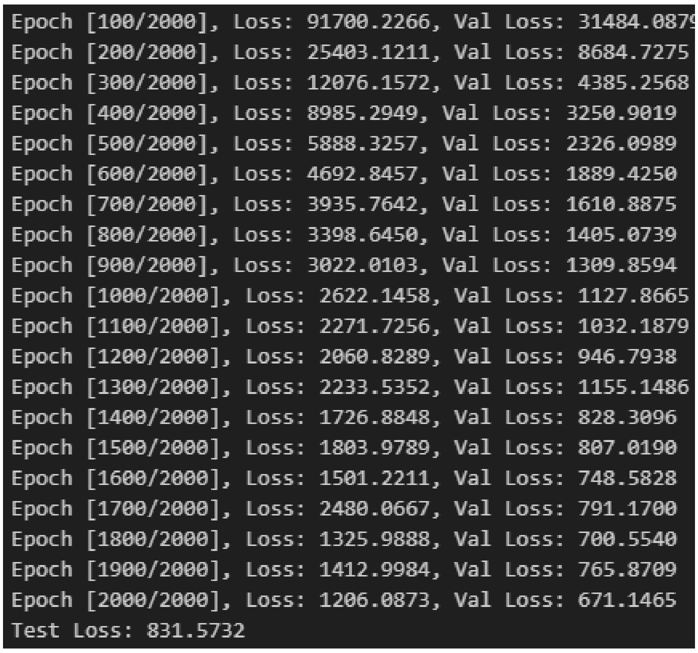
which we are going to minimize by finding the appropriate θ strategy.

$$\arg\min_\theta \mathbb{E}\left[\Lambda_T^2\right]. \tag{3.13}$$

The Long-Short Term Memory Neural Network is implemented with 3 layers of 64 neurons each, for a total number of 85'252 parameters which are trained on a training set of 3'000 scenarios with 240 observation/rebalancing steps each (720'000 total datapoints) for 2'000 epochs.

The training is validated over a 1'000 scenarios and finally tested on additional 1'000 scenarios.

The NN is trained to work with both the underlying and the 3 ATM/ITM/OTM strategies based on LLP as illustrated before.
With this amount of data, the training gets completed in about 5 hours on an i7, 3.00GHz, 3001 Mhz, 4 Cores, 8 Logical Processors, leading to the training steps as reported in 3.1. As per the other hyperparameters, the Adam optimizer is used with a learning rate of 0.001 and a batch size of 32. The detailed code used is provided in Appendix B.1.



```
Epoch [100/2000], Loss: 91700.2266, Val Loss: 31484.087
Epoch [200/2000], Loss: 25403.1211, Val Loss: 8684.7275
Epoch [300/2000], Loss: 12076.1572, Val Loss: 4385.2568
Epoch [400/2000], Loss: 8985.2949, Val Loss: 3250.9019
Epoch [500/2000], Loss: 5888.3257, Val Loss: 2326.0989
Epoch [600/2000], Loss: 4692.8457, Val Loss: 1889.4250
Epoch [700/2000], Loss: 3935.7642, Val Loss: 1610.8875
Epoch [800/2000], Loss: 3398.6450, Val Loss: 1405.0739
Epoch [900/2000], Loss: 3022.0103, Val Loss: 1309.8594
Epoch [1000/2000], Loss: 2622.1458, Val Loss: 1127.8665
Epoch [1100/2000], Loss: 2271.7256, Val Loss: 1032.1879
Epoch [1200/2000], Loss: 2060.8289, Val Loss: 946.7938
Epoch [1300/2000], Loss: 2233.5352, Val Loss: 1155.1486
Epoch [1400/2000], Loss: 1726.8848, Val Loss: 828.3096
Epoch [1500/2000], Loss: 1803.9789, Val Loss: 807.0190
Epoch [1600/2000], Loss: 1501.2211, Val Loss: 748.5828
Epoch [1700/2000], Loss: 2480.0667, Val Loss: 791.1700
Epoch [1800/2000], Loss: 1325.9888, Val Loss: 700.5540
Epoch [1900/2000], Loss: 1412.9984, Val Loss: 765.8709
Epoch [2000/2000], Loss: 1206.0873, Val Loss: 671.1465
Test Loss: 831.5732
```

Figure 3.1: Training and Validation Loss evolution per epoch and final Test set Loss.

The results of the strategy on the test set are summarized in Table 3.1 and compared with those of the *greeks* approach.

Statistic	**LSTM NN Terminal PnL Distribution**	**Δ-Γ-$\mathcal{V}$ Hedging**
Mean	-0.020089003668762737	-0.029283118428525413
Median	-0.00515821001447847	0.00819456704006882
Std	0.9121412437502409	0.7709353008154314
Skew	0.6296091554407799	-0.4035915562583632
Kurtosis	7.960533102689574	2.997245297857718
VaR 95%	-1.4275061185891829	-1.3798023211211694
VaR 99%	-2.523058980367784	-2.4255521326274776
ES 95%	-2.089862716129953	-1.9668512906021507
ES 99%	-3.27431072822869	-2.9454042357003067
Min	-5.038591384887695	-3.684422601997838
Max	7.297592952995316	3.557338188448142

Table 3.1: Summary statistics for the risk minimized portfolio with NN on the Test Set and compared with the Δ-Γ-$\mathcal{V}$ Hedging results as shown in 2.7.4.

The terminal PnL distribution is graphically reported in 3.2. Although the results are not outperforming the *greeks*'ones, we do observe a reasonably close outcome, with an average result quite close to 0 and a small standard deviation.

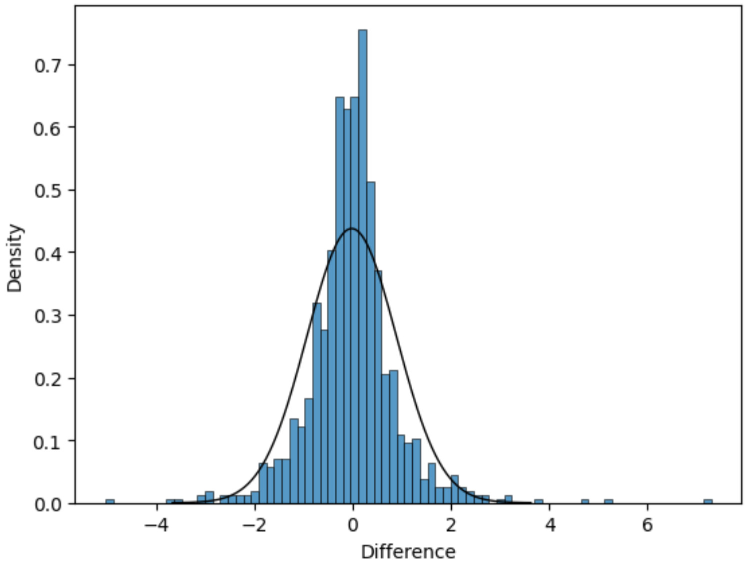


Figure 3.2: Terminal net PnL using a LSTM NN based roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards.

Interestingly, the NN seems to work very well with the model price used as input, ultimately converging to a strategy really close to the one based on the sensitivities matching (see Figure 2.16), but showing some smoother dynamics.

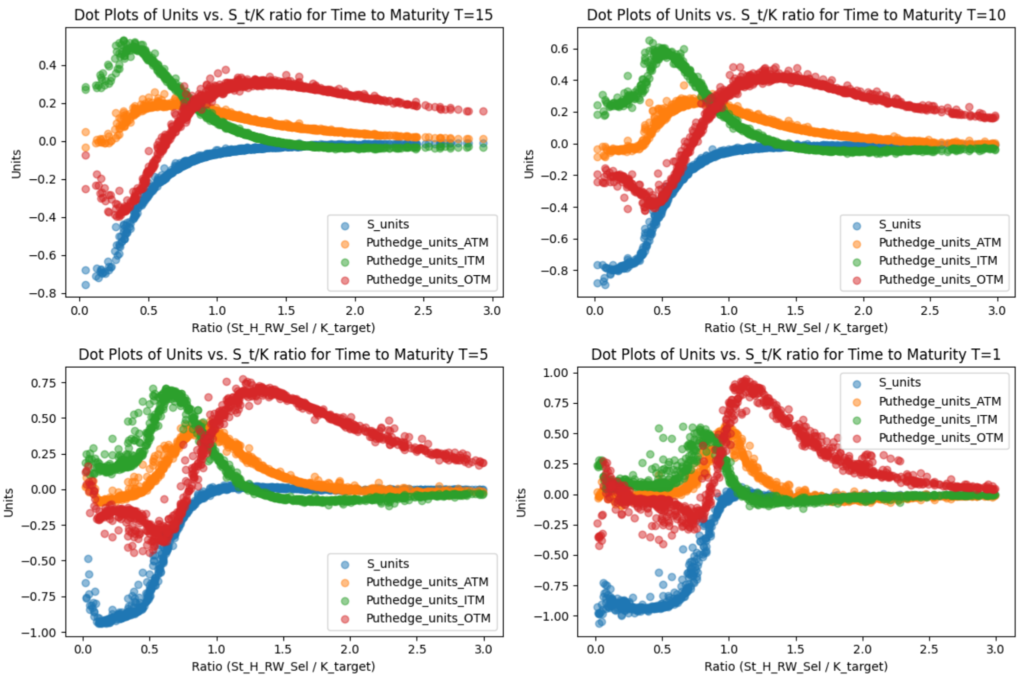


Figure 3.3: Dynamic evolution of the NN based hedge trades units for certain t and considering the ratio S_t/K.

3.3.1.2. Extended NN Training with V_0^θ Calibration

We now move to the case where the Neural Network should do "everything" by itself, in the sense that given a set of data which include the evolution of the underlying and the prices of the hedging derivatives it should be able to get trained to find a model independent way to come up with an appropriate hedging strategy θ and related V_0^θ price for the target derivatives which jointly minimize the loss function 3.7.

We perform this test on the European Put case already used in the previous section. The simulation scheme is still based on the Heston model with volatility displacement and same parameters as the previous cases. Although still applied to a simple target derivative, this test is particularly useful because it allows us to assess the flexibility of the NN to converge towards an appropriate solution, which we now it exists in this case as we are able to fully benchmark the exercise with the results provided by the model pricing and the *greeks* hedging.

It is therefore of utmost importance to understand where the NN is going to converge under this simplified example before applying the methodology to more complex and real-life cases (e.g. most complex payoff structures with the introduction of bid-ask costs) where we could not fully benchmark the outcome of the NN strategy.

As a first result, let's assess the "fair" price estimated by our Neural Network to trade the target derivative under the hedging policy defined to meet the target set out in 3.7. We can compare this with the model price as calculated using the methodology in 2.2 and 2.3 as well as with the one calculated through our risk neutral Monte Carlo simulation. For the ATM Put option with expiry 20 years we get:

$$\begin{aligned} V_0^{\theta} &= 16.573 \\ V_0^{CF} &= 16.543 \\ V_0^{MC} &= 16.834 \end{aligned} \tag{3.14}$$

The observed result for the price calculated by the NN V_0^{θ} gets extremely close to the model and the Monte Carlo prices, with actually the NN Price getting closer to the "true" price than the Monte Carlo generated with high granularity (100K scenarios with weekly frequency and Milstein discretization scheme).

As per the terminal hedging PnL, we see the following:

Statistic	**NN Terminal PnL with V_0^{θ}**
Mean	-0.017975424542920657
Median	0.004916248843073845
Standard Deviation	0.746625951642729
Skewness	-0.7963303895697817
Kurtosis	12.021433878127668
VaR 95%	-1.1223201260899032
VaR 99%	-2.140927869781419
ES 95%	-1.9138554875447755
ES 99%	-3.500461510080414
Minimum	-5.5455827713012695
Maximum	5.768661812095658

Table 3.2: Summary statistics for the risk minimized portfolio with NN on the Test Set with the estimation of V_0^{θ}.

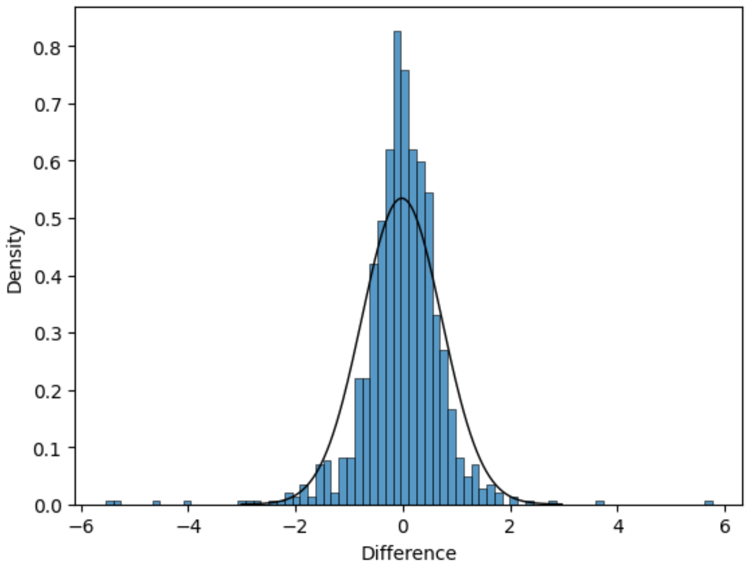


Figure 3.4: Terminal net PnL using a LSTM NN based roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. The NN is also used to calculate V_0^θ.

In this case the results look very good in terms of terminal PnL Average, Median and Standard Deviation. Although, we observe a slightly higher kurtosis, risk metrics such as the Value-at-Risk and Expected Shortfall are still in line with the previous test where the option price was given as an input, as well as the Δ-Γ-$\mathcal{V}$ hedging case.

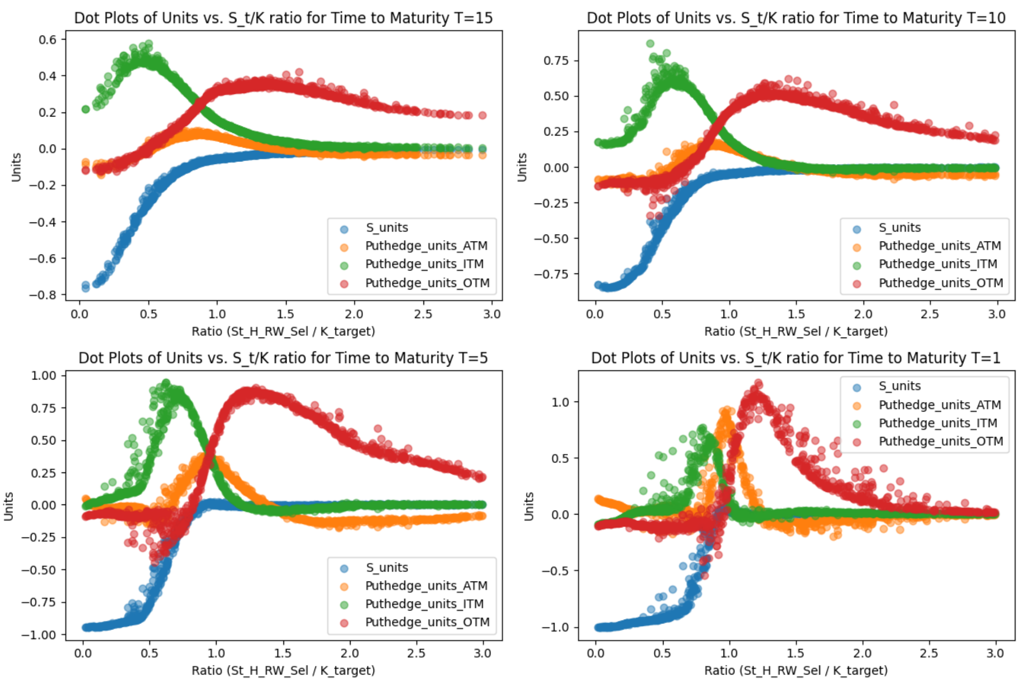


Figure 3.5: Dynamic evolution of the NN based hedge trades units for certain t and considering the ratio S_t/K. The NN is also used to calculate V_0^θ.

A quick look at the evolution of the hedging strategy over time based on the moneyness ratio S_t/K (Figure 3.5), shows again a behavior consistent with the simplified case where V_0 was provided as an input as well as with the *greeks* hedging case.

3.3.2. Checking the Ongoing PnL Volatility - Assessment on the NN Hedging Performance

Subsection 3.3.1.2 has shown how Neural Networks could well perform in the minimization of the terminal hedging error at expiry.
To have a full understanding of the effective performance of this approach we should also be confident that our hedging strategy is not only working well from a terminal PnL point of view, but also that the ongoing slippage (change in the Assets minus change in the liabilities) remains small enough to prevent the Company from running into troubles due to equity losses. Taking for example the case of an Insurance Company which is selling guarantees on a certain portfolio of financial securities, it may well be that our hedging strategy would perform well at expiry, but leaving the Company exposed to huge PnL swings over time. That may be as bad as not hedging the liabilities at all in that it may cause severe ongoing losses to be absorbed with capital injections. We would therefore need to assess the hedge efficiency also over the life of the claim. Since we are considering monthly rebalancing steps, we assess the monthly volatility of the hedge results. As detailed in 1.5, there may be many ways to assess the hedge performance (see [12] and [27]). We focus on the metric defined in 1.8, the simplest and most intuitive KPI for Hedge Effectiveness:

$$HE = 1 - \left[\frac{VARIANCE_{HedgedPortfolio}}{VARIANCE_{UnhedgedPortfolio}}\right]$$

Intuitively, the closest this indicator gets to 1, the better our hedging strategy is performing. On the contrary, a value of HE close to 0 would suggest that our hedging may be ineffective in terms of variance reduction.
Finally, we may witness negative values for HE in case the variance of the hedged portfolio is actually higher than the variance of the unhedged position. This may not be necessarily an indication that our hedging is bad, but may arise when for example different and negative correlated risk drivers move the price of our liabilities and we try to hedge them all separately, but the effective residual exposures that we get after hedging actually creates bigger effects than the initial unhedged ones aggregated.

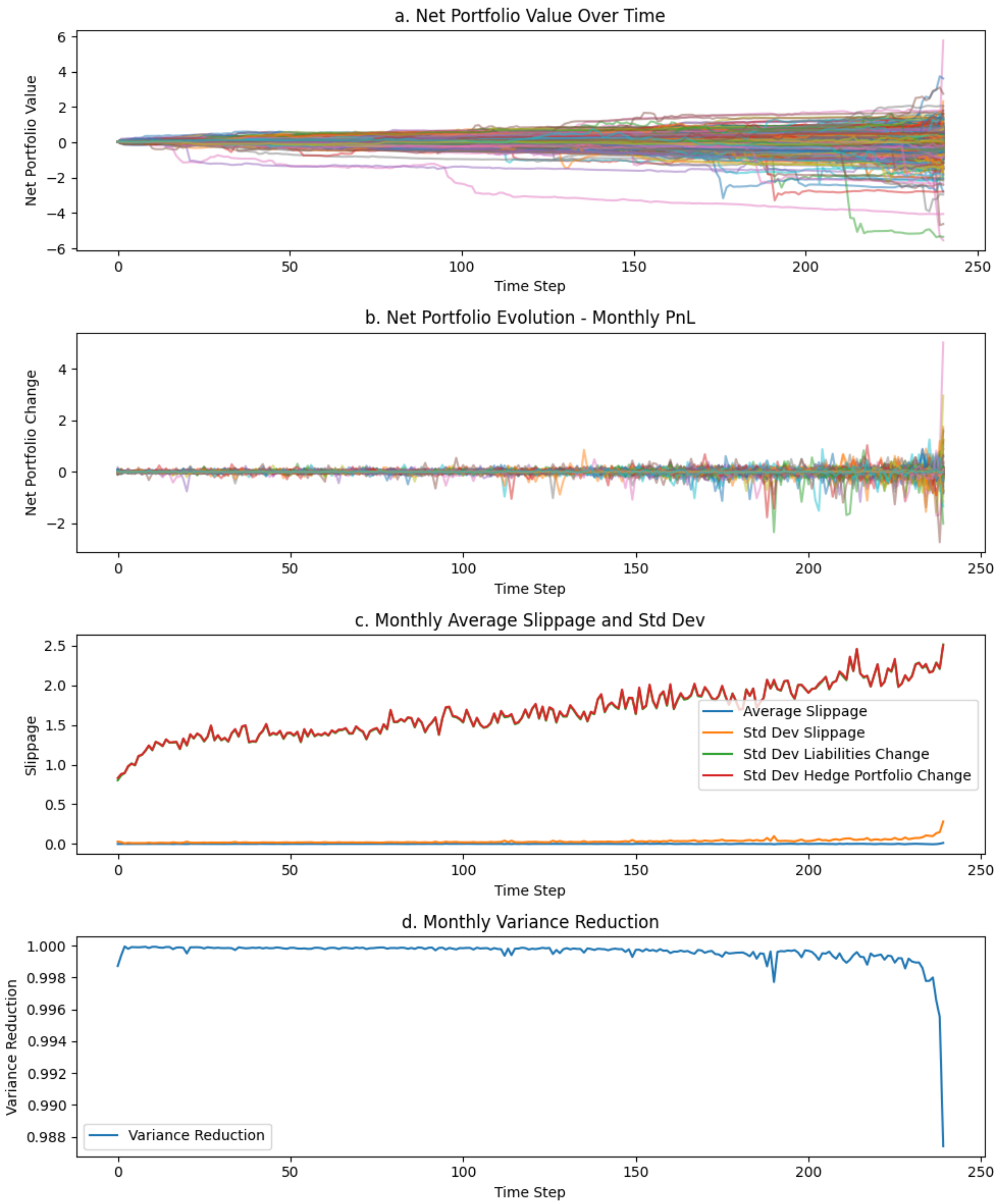


Figure 3.6: Hedging results analysis: **a**. shows the net portfolio value over time. **b.** shows the net portfolio changes over time; **c**. reports the slippage average and standard deviation, together with the standard deviation of the standalone liability and hedge portfolio; **d**. shows the variance reduction achieved with the Neural Netwrok based hedging.

The outcome of the analysis is summarized in Figure 3.6. In a first place (Sub-fig. **a.**),

the overall evolution of the portfolio value is reported. The starting point of the graph is at 0 since the value of the hedge portfolio in $t = 0$ is supposed to be the same as V_0^θ for this exercise. However this may not be necessarily true in case the operator would like to add an additional margin on top of the 'fair' derivative price. Reasonably we can see that as time passes the distribution of the net portfolio value widens as it can be expected that hedging errors cumulates over time. Ultimately, the terminal PnL distribution converges to what shown already in Figure 3.4.
The following Sub-figures are also extremely useful:

b. Shows the monthly net portfolio changes. Here we can visually understand if the hedging is working on an ongoing basis. We can really appreciate how the strategy works well in the first years, while getting a bit more volatile as soon as the time to maturity decreases. The most interesting element here is the PnL spike at the very end. Intuitively this may arise from the fact that at every step the strike of the hedge puts is refreshed with the usual ITM,ATM,OTM scheme, while the strike of the target derivative remains unchanged to the starting value. This may create discrepancies at the very end when the terminal value coincides with the actual payoff on both assets and liabilities.

c. Here we see how both the average and standard deviation of the net portfolio (monthly) changes stay reasonably around 0. This is the goal we aim at, in that we would like to achieve a perfect matching between assets and liabilities (which have a pretty similar evolution of the monthly changes standard deviation). Worth to mention the volatility spike in the monthly slippage at the end of the hedging horizon, which has been already discussed on the previous point. The statistics here are calculated by aggregating all the scenarios for each single time.

d. Ultimately we have our measure of efficiency based on the summary statistics previously computed under the methodology defined in **??**. Here we can appreciate the extremely high effectiveness of our strategy, which makes possible to reduce almost completely the variance of the ongoing monthly results. The efficiency slightly deteriorates at the end, but nevertheless remaining in the order of 99% variance reduction.

3.3.3. Main Limitations of the Current Approach

As shown in 3.3.1.1, 3.3.1.2 and 3.3.2, hedging through Neural Networks for long term equity derivatives performs reasonably well in absolute terms and relatively to the sensitivities based hedging. Additionally, it gives extreme flexibility in the definition of the custom loss function to minimize, which could then incorporate operational constraints and more realistic assumptions such as assumptions on the costs of trading.

However, some limitations affecting our implementation have to be highlighted:

1. **Training Time**: The NN requires a considerable amount of parameter and consequently scenarios and time steps to be trained in order to avoid overfitting. This results in a slow training process which could last several hours.

2. **Generalization Power**: Our model is trained to price and hedge a very specific and targeted derivative: a 20 years ATM Put option. For this case we get accurate and effective results which are also consistent with the model-based results. However, our implementation suffers of an undesirable drawback, in that it is not able to generalize the obtained results to claims with different features. If we try, for example, to use our trained model to price the a 20% In-the-Money Put option, we would still get a price close to the one estimated for the ATM case. Also if we would like to compute the price for the derivative with a shorter time to expiry, we may not get an accurate result as the NN is trained to compute the $t = 0$ price from a 20 years long hedging strategy. The issue arises from the fact that the NN is using the previous time value of the hedge portfolio as the starting point to adjust the hedging units, therefore making the strategy dependent on the past evolution of our hedging.

3.3.4. NN Hedging Results for a Cliquet Put Option

After the extensive assessment on the plain vanilla case, we move forward to consider a more complex and interesting product from a life-insurance perspective. Let's take the case of a potential investor who seeks to preserve his capital while benefiting from the upside given by the investment into a certain basket of financial securities/Index(es). The pure combination of the investment into the underlying(s) plus the protective ATM put option for a long maturity horizon (10 years or more) may not be optimal in case the insurer would like to make the product more appealing for the policyholder and swap the payment of the fees over the life of the insurance policy instead of charging a one-off fee (similar to an option premium) at inception.

Swapping such fee over time with a fixed strike price until maturity may be a risky move for the insurer as it may create considerable lapse behavior risk in case the guarantee starts to become deep OTM. In such case, a rationale and informed policyholder would probably stop to pay the ongoing charges, set based on a past higher moneyness, for a worthless guarantee, basically leaving the Company without the necessary revenues to fund the cumulated losses/costs of the hedge portfolio.

In order to mitigate such risk, the insurance provider may consider to add a Ratchet feature on the product structuring. Allowing the strike of the embedded put option to rest over time based on the new levels reached by the underlying price may effectively increase the appeal of the product and the willingness of the clients to remain and keep paying the ongoing charges.

Adding such feature may considerably increase the pricing of the claim, which may be hardly understood/accepted from the potential policyholder. It would make therefore sense to impose some constraints on the Ratchet mechanism so to avoid pricing overshoots. We consider the following scheme:

1. The guarantee resets at each policy anniversary.
2. The guaranteed value can reset only up to a cap which is defined as the 5% of the initial underlying price.

The potential evolution of such contingent claim is shown in Fig 3.7. Under the highlighted scenario, the strike K gets reset and increased few times although the terminal price of the underlying ends up at a slightly higher level but still very close to the starting one. Under a fixed strike price condition at 100, the guarantee would expire out of the money (final stock price of 105), while with the current Cliquet structure, the guarantee ends up deeply in the money and the policyholder pockets almost 140.

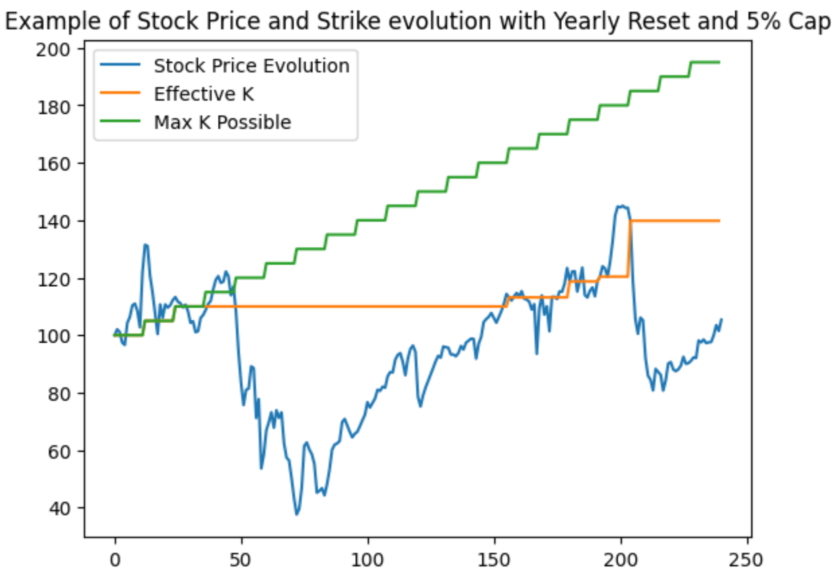


Figure 3.7: Possible dynamic evolution of the strike price under the Ratchet mechanism described in 3.3.4.

This structure should therefore improve the attractiveness of the product for a potential customer since it provides a possibly increasing guaranteed amount over time and also for the Insurance Company which aims to collect the fees on an ongoing basis since it reduces the chances that the policyholder will lapse as it also reduces the probability of the guarantee being deep out of the money.

With this structure and the usual parameters already used so far, we obtain a Monte Carlo price of:

$$P^{MC}_{Cliquet} = 27.411$$

Imposing the time dependent cap on the strike reset gets useful from a pricing attractiveness point of view for a retail/private client. If we would for example impose an annually resetting guarantee without the cap feature, then the theoretical price to charge would increase to $P^{MC}_{NoCap} = 38.991$ with a clear impact on the potential product's appeal.

As detailed above, the payoff structure here analyzed exhibits a much higher degree of complexity than the European plain vanilla case. Furthermore, the pricing and risk metrics calculation methodology presented in Section 2.2 cannot be applied for this case anymore.
Even if this reduces the benchmarking capacity, it does not affect by any means the possibility to apply our NN methodology to price and hedge this specific payoff, although we may expect a slightly higher volatility to reach the hedging target due to increased complexity.

The first element to assess is the pricing we obtain from the NN training:

$$P^{NN}_{Cliquet} = 27.1511$$

Interestingly, the price is extremely close to the Monte Carlo one as for the plain vanilla case.

Moving to the terminal hedging PnL, we observe the following results on the Test Set for the Cliquet option using the (re-trained) LSTM-NN already set up for the plain vanilla case:

Statistic	NN Terminal PnL for Cliquet
Mean	0.06732390601674745
Median	0.031049736900051528
Standard Deviation	2.637371267494814
Skewness	0.10162062311350513
Kurtosis	1.7866536307910716
VaR 95%	-4.564566146564959
VaR 99%	-6.867522175639404
ES 95%	-6.059572355785483
ES 99%	-8.047881785760666
Minimum	-9.972596168518066
Maximum	12.337162017822266

Table 3.3: Summary statistics for the risk minimized portfolio with NN on the Test Set with the estimation of V_0^θ, for a Cliquet option with yearly reset caps on 5% on the initial spot level.

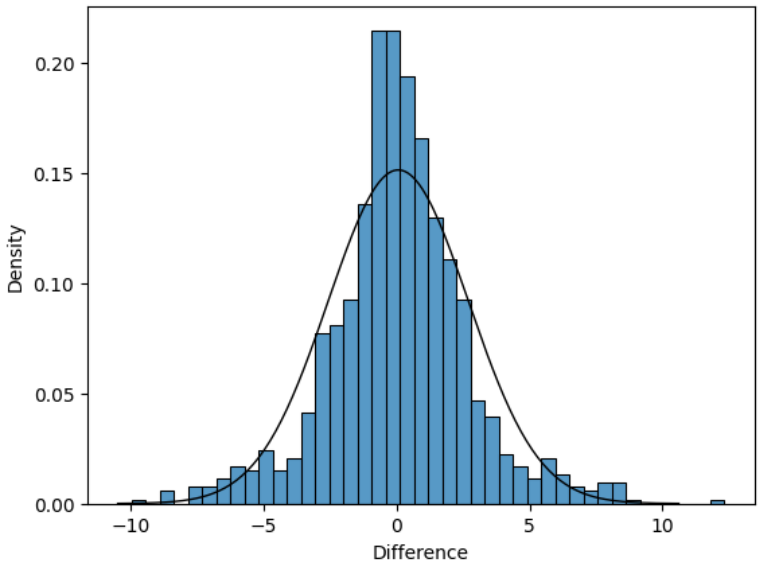


Figure 3.8: Terminal net PnL for a 20 years Cliquet option with yearly reset caps of 5% of the initial spot, using a LSTM NN based roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. The NN is also used to calculate V_0^θ.

Here again the NN is set to achieve a mean/median terminal result which is around 0, with a reasonably symmetric dynamic and a slight increase on the volatility of the results, although also the initial price is higher than the initial plain vanilla case.

We ultimately analyze the evolution of the hedging strategy in Fig. 3.9. In this case we have considerably more data points which lie in the moneyness area due to the reset feature. Compared with the hedging strategy shown in Fig. 3.5, here we observe a less stable and smooth behavior of the hedge portfolio units allocated to the allowed instruments, especially close to the moneyness area.

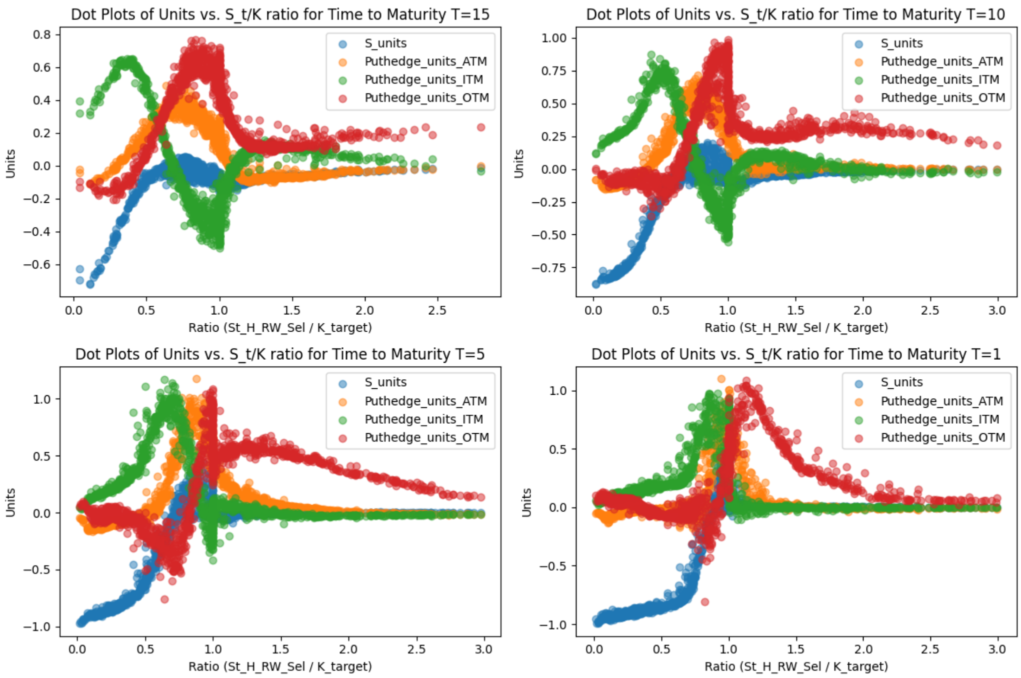


Figure 3.9: Dynamic evolution of the NN based hedge applied to the 20 years Cliquet option with yearly reset caps of 5% of the initial spot: trade units for certain t and considering the ratio S_t/K.

3.4. Hedging Performance Under Market-Calibrated Parameters

As a further insightful test, we would like to assess how the chosen model would fit some actual market data and how the NN pricing and hedging would then perform under such specific parameters.

3.4.1. Calibration of the Heston Model with Displacement

The calibration is performed as of 30th September 2024 (close of business) on the Eurostoxx50's implied volatilities as provided by Bloomberg. The target volatilities used for calibration range between 0.8 and 1.2 moneyness with tenors between 1 year and 10 years. The calibration routine aims to minimize the quadratic distance between the market implied volatilities and the model ones, tuning the set of parameters already discussed for the Heston model with displacement. The objective function first calculates the displaced Heston call option price under a certain set of parameters, then inverts it to find the corresponding Black & Scholes implied volatility which is ultimately compared with the observed one. The routine stops once the sum of squared errors is minimized and it is based on differential evolution (see [28] for reference). A pseudo code is reported on Algorithm 3.2. Boundaries are introduced for parameters to avoid unrealistic behaviors in the scenario generation.

Algorithm 3.2 Heston Model Calibration via Differential Evolution

Input: Spot prices S, Strikes K, Maturities T, Rates r, Dividends q, Observed implied vols IV_{obs}
Initialize minimalValueReached $\leftarrow \infty$
Optimize:
    $result \leftarrow$ DifferentialEvolution(ObjectiveFunction)
if $result.fun <$ minimalValueReached **then**
    optimalX $\leftarrow result.x$
    minimalValueReached $\leftarrow result.fun$
end if
function ObjectiveFunction(settings)
    Compute Heston prices: $P \leftarrow HestonPriceGaussLaguerre(S, K, T, r, q, settings)$
    Compute implied vols: $IV_{model} \leftarrow$ BSImpliedVolatilitySolver(P, S, K, T, r, q)
    Compute error: $error \leftarrow \|IV_{model} - IV_{obs}\|_2$
    return $error$
end function
Output: Optimal Heston parameters optimalX

The parameters obtained and used for the pricing and hedging tests of this Section are reported on Table 3.4. r and q are not directly calibrated and are assumed to be respectively 1.5% and 0% (μ set at 5% for real world simulation). As per the other parameters, the model fit looks to be very good, with a total 0.0242 Root Mean Square Error (RMSE) over a set of 72 volatility points. This is also visualized on Figure 3.10, where the model IV surface slices are compared with the market points. Besides the market volatilities used for calibration on the target tenors range, also the 3 and 6 months expiries are shown with the related fitted smile, which captures pretty well the market shape notwithstanding the fact that it is not directly calibrated on these points. It also seems that the model is rather good in capturing the short term smile while progressively reflecting the skew for longer dated tenors.

Parameter	Symbol	Value
Risk-free rate	r	0.015
Volatility	σ	$\sqrt{0.02597}$
Deterministic displacement	ϕ	0.004557
Dividend yield	q	0
Rate of mean reversion	κ	1.2
Long run variance	θ	0.035797
Volatility of volatility	ϵ	0.6
Correlation between Wiener processes	ρ	-0.7724

Table 3.4: Model Parameters for Eurostoxx50 as of 30th September.

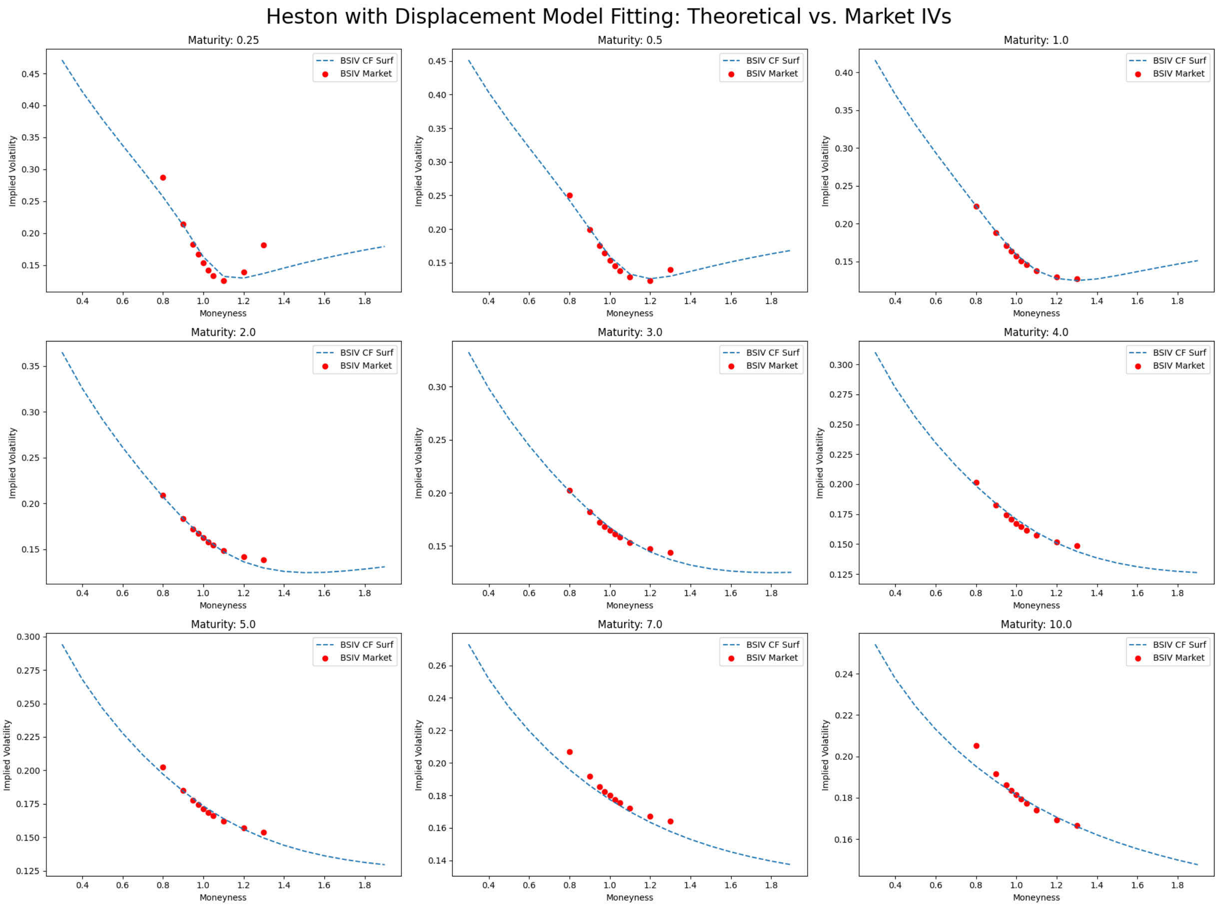


Figure 3.10: Visualization of the Heston model with displacement fitting on Implied Volatilities for the Eurostoxx50 as of 30th September 2024.

3.4.2. Pricing and Hedging Results

As done for the other test cases, we consider the plain vanilla ATM European Put option case with 20 years expiry and the NN performance in terms of pricing and hedging, together with the evolution of the hedging strategy across time and moneyness.

$$\begin{aligned} V_0^{\theta} &= 17.2031 \\ V_0^{CF} &= 17.1753 \\ V_0^{MC} &= 18.394 \end{aligned} \tag{3.15}$$

Statistic	**NN Terminal PnL on Eurostoxx50 Calibrated Parameters**
Mean	-0.0765264353357975
Median	-0.07417701929807663
Standard Deviation	0.6028427729510489
Skewness	0.0907715787014489
Kurtosis	4.765270303170036
VaR 95%	-1.0812869924508952
VaR 99%	-1.8879029762744903
ES 95%	-1.5522310126788854
ES 99%	-2.257127502546335
Minimum	-2.812966399492666
Maximum	3.7844276021886856

Table 3.5: Summary statistics for the risk minimized portfolio with NN on the Test Set with the estimation of V_0^{θ}, for a plain vanilla ATM European Put with 20 years time to expiry and stock price generated under a stochastic volatility model with displacement and calibrated on Eurostoxx50 IVs as of 30th September 2024.

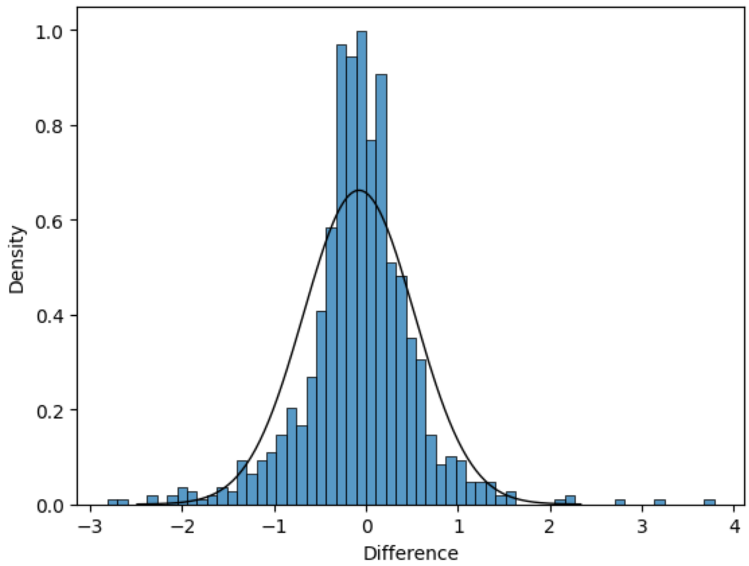


Figure 3.11: Terminal net PnL for a 20 years ATM Put option with underlying dynamic based on stochastic volatility with displacement calibrated on Eurostoxx50 IVs, using a LSTM NN based roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. The NN is also used to calculate V_0^θ.

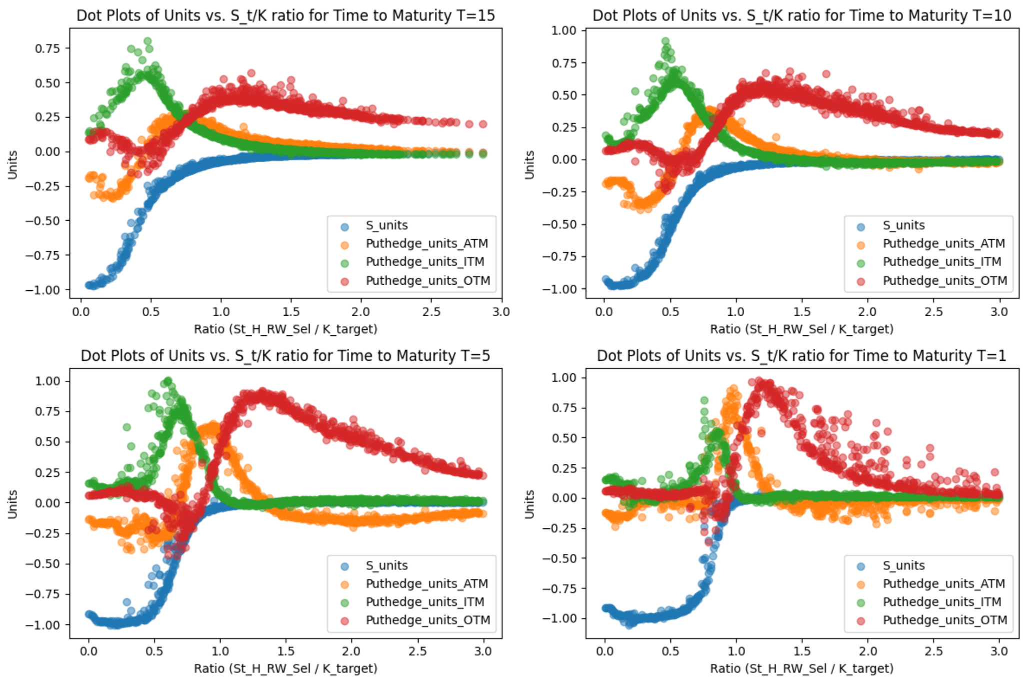


Figure 3.12: Dynamic evolution of the NN based hedge applied to the 20 years ATM Put option with underlying dynamic using stochastic volatility with displacement calibrated on Eurostoxx50 IVs: trade units for certain t and considering the ratio S_t/K.

Not surprisingly, we observe again very good results in terms of convergence to the theoretical price and of hedging results on the test set.

3.5. Hedging under Bates model with Instantaneous Variance Displacement

Let's ultimately consider a more complex model for the Stock price which adds a jump process to the stochastic volatility model with displacement considered so far.

Taking into account the Bates model [29] with the addition of a deterministic volatility displacement, we define the real-world measure $\mathbb{P}$ through the following system of stochastic differential equations (SDEs):

1. The asset price S_t follows the SDE:

$$dS_t = S_t \left(\mu \, dt + \sqrt{v_t + \phi_t} \, dW_t^S + J \, dN_t\right)$$

   where:

   - S_t is the asset price at time t.
   - μ is the drift of the asset price, representing the expected return of the asset in the real world.
   - v_t is the stochastic variance of the asset.
   - ϕ_t is the deterministic time dependent displacement factor, with $\phi_0 = 0$.
   - dW_t^S is a Brownian motion under the real-world measure $\mathbb{P}$.
   - $(1 + J)$ is the jump size, with $\ln(1 + J) \sim \mathcal{N}\big(\ln(1 + \mu_J) - \frac{\sigma_J^2}{2}, \sigma_J^2\big)$.
   - N_t is a Poisson process with intensity λ, representing the arrival of jumps: $\mathbf{P}(dN_t = 1) = \lambda dt$.

   The term $J \, dN_t$ represents jumps in the asset price, and the drift μ reflects the expected growth rate of the asset in the real world.

2. The variance v_t under the real-world measure $\mathbb{P}$ is modeled by the following mean-reverting stochastic process:

$$dv_t = \kappa(\theta - v_t) \, dt + \sigma \sqrt{v_t} \, dW_t^v$$

   where:

- v_t is the variance of the asset price at time t.
- κ is the rate of mean reversion of the variance.
- θ is the long-term mean of the variance.
- σ is the volatility of volatility (vol of vol).
- dW_t^v is a Brownian motion driving the variance, under the real-world measure $\mathbb{P}$.

The correlation between the Brownian motions dW_t^S and dW_t^v is given by:

$$\text{corr}(dW_t^S\, dW_t^v) = \rho\sqrt{\frac{v_t}{v_t + \phi_t}} \tag{3.16}$$

This specific model is well suited for our research as, besides the MC simulation, it also allows us to price plain vanilla options in a fast and accurate way through the Characteristic Function with a simple extension of the methodology shown in 2.2.

3.5.1. Hedging Performance

Let's now investigate how the NN hedging works under a price process with jumps.

In addition to the usual parameters as shown in Table 2.1, the following jump parameters are used:

Parameter	**Symbol**	**Value**
Drift of Jump	μ_j	-0.3
Volatility of Jump	σ_j	0.06
Jump Intensity	λ_j	0.02

Table 3.6: Jump Parameters.

We consider again the European plain vanilla ATM Put with 20 years to expiry as a target derivative and hedging it with the usual LLP based strategy. The trained Neural Network is, interestingly, once again converging toward the theoretical price:

$$\begin{aligned} V_0^{\theta} &= 17.1614 \\ V_0^{CF} &= 17.1466 \\ V_0^{MC} &= 17.2045 \end{aligned} \tag{3.17}$$

As per the hedging results on the test set, they are reported in Table 3.7 and are again very similar to the case without jumps when the LLP option strategy is used.

Statistic	NN Terminal PnL for Bates
Mean	-0.09651135324032983
Median	-0.060682396273321615
Standard Deviation	0.8694072618844932
Skewness	-0.5680524748933939
Kurtosis	4.133420642340933
VaR 95%	-1.5959318041801451
VaR 99%	-2.504755666255951
ES 95%	-2.2340015477494695
ES 99%	-3.400939920913589
Minimum	-6.070892810821533
Maximum	3.27364873354864

Table 3.7: Summary statistics for the risk minimized portfolio with NN on the Test Set with the estimation of V_0^θ, for a plain vanilla ATM European Put with 20 years time to expiry and stock price generated under a stochastic volatility model with displacement and jumps.

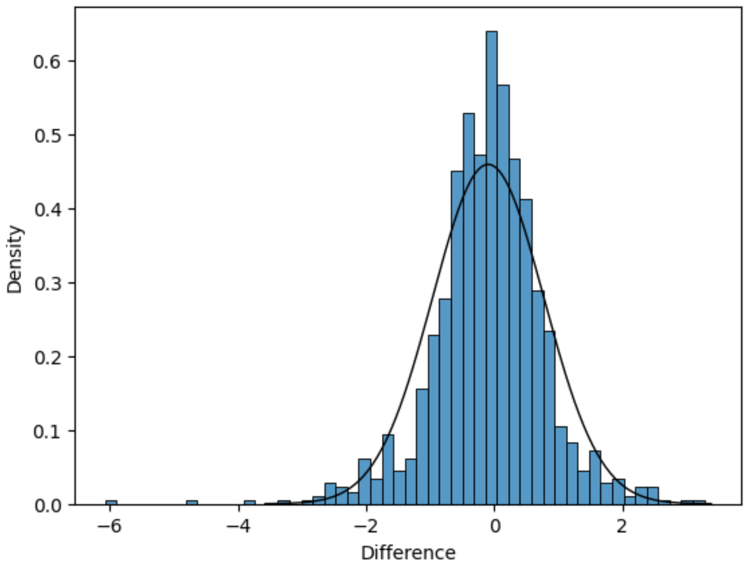


Figure 3.13: Terminal net PnL for a 20 years ATM Put option with underlying dynamic based on stochastic volatility with displacement and jumps, using a LSTM NN based roll-over strategy of 5 years ATM/ITM/OTM puts until the residual maturity of the target put remains above 5 years and matching it afterwards. The NN is also used to calculate V_0^θ.

The comparison with the "base" stochastic volatility with displacement case discussed in Subsection 3.3.1.2 is particularly interesting. It would seem that under the new dynamic, the NN based hedging would not decrease its efficiency, while properly reflecting a Put price increase in line with the new jump model used, also slightly adapting the hedging strategy to achieve the terminal target.

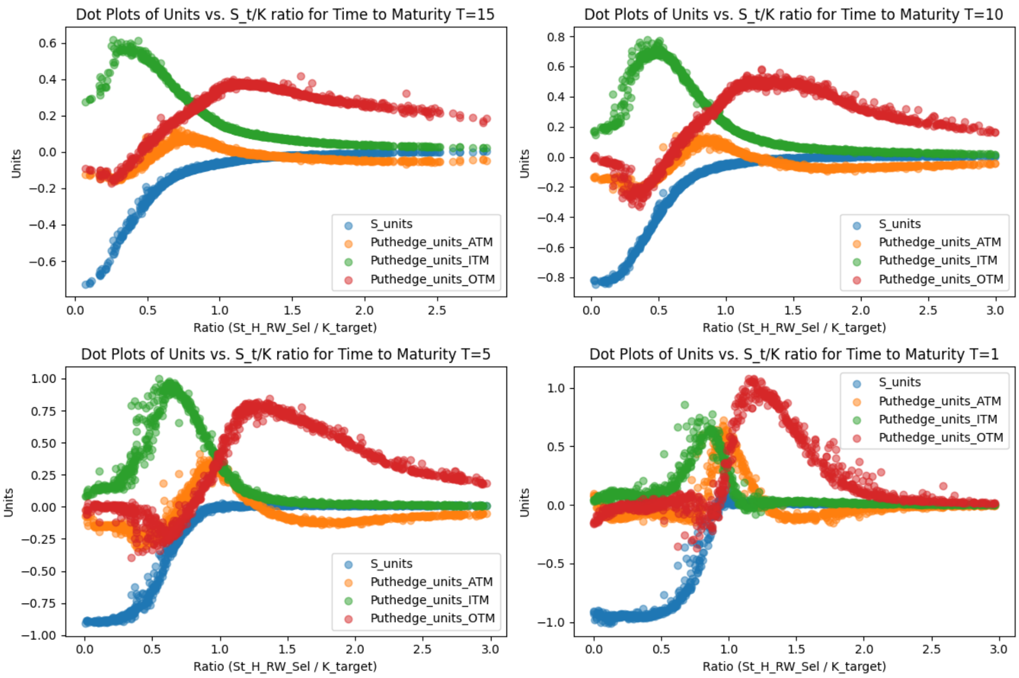


Figure 3.14: Dynamic evolution of the NN based hedge applied to the 20 years ATM Put option with underlying dynamic using stochastic volatility with displacement and jumps: trade units for certain t and considering the ratio S_t/K.

3.6. Conclusions

The final Chapter of this PhD Thesis has investigated how a financial operator could start dealing with long term equity derivatives in a non complete market, by moving away from the traditional hedging process based on risk sensitivities to the one based on global risk minimization, through the usage of Neural Networks.

The Chapter offered a general overview of the global risk minimization hedging based on the Literature available on the topic, readjusting it to be compatible with the hedging strategy based on the LLP proposed in Chapter 2.

After setting up the hedging problem, the focus was shifted on the definition of the technology to be used to solve it: a Long Short-Term Memory Neural Network with 3 layers of 64 neurons each. The structure of the customized loss function to be optimized has also been extensively detailed.

Finally, several numerical test cases to assess the performance of global risk minimization hedging with NNs have been provided:

1. Base case on a plain vanilla ATM European Put option with 20 years strike where the theoretical price is given as an input and the NN has to find the best hedging strategy;
2. More complex case where the NN does not know the theoretical price, which becomes therefore a variable to find by the NN.
3. Application of this technique to the case of an exotic Cliquet Put which may represent an attractive product to sell for a Life Insurance Company.

Additionally, test cases were also provided for a real case where the model parameters have been calibrated over Eurostoxx50 Implied Volatilities, as well as for the case where the underlying stock price process has been generated under a stochastic volatility model with displacement and jumps.

In all these test cases, the hedging results combining the Neural Network with the hedging strategy based on the last liquid point have performed very well, generally matching very closely the theoretical model price.

It is expected that With all these applications, the operator would become quite practical on how to hedge long term equity derivatives in a non complete market. Extending this approach to the case where also trading costs are taken into account should be rather feasible by properly adjusting the loss function to optimize and the basic principles of the hedging strategy in order to reduce as much as possible the rebalancings. This is left for future research.

4 | Key Takeaways and Conclusions

The main aim of this PhD Thesis is to improve the current literature and techniques used to price and manage risks related to Variable Annuities products. To do so, several aspects have been faced and analyzed.

Chapter 1 offers a general overview on VA products, together with some useful literature references. The objective is to break down the main product specificities in order to make them understandable also to the reader not necessarily involved directly into such business, and then to focus on all the key areas that affect the life-cycle of such products. More specifically, the main VA types are presented (GMAB, GMDB, GMWB, GMIB), with a specific attention to the GMAB case. In that regard, different types of structures have been illustrated and visualized (Figure 1.1).

The main VA risks have also been discussed, both from a financial and non-financial perspective, with important considerations on the best risk management practices and attention to product design.

The concept of financial Hedging has then been introduced as the key element to ensure the solvability of the insurance company. Several hedging approaches have been presented, with also a comprehensive overview on how to measure the hedge performance. The last contribution of the Chapter is a general overview on the pricing logic needed for VA products. This has a two-fold objective: first, it serves as a pricing/modeling introduction to better understand the topics faced in the following Chapters; secondly, it widens the perimeter of models presented in this Thesis, suggesting some possible alternatives to those effectively used later on.

Chapter 2 sets the ground to design and test, on *ex-ante* basis, an efficient hedging strategy through risk sensitivities minimization under a Displaced Heston Model. The model is rigorously defined, from both a theoretical and a practical perspective, with detailed explanations on the simulation scheme, pricing logic and the creation of the dataset used for optimization purpose.

The tests start with a theoretical hedging exercise comparing the effectiveness of a pure Delta hedging against one based on all the *Greeks*, with the latter showing a considerable improvement to the first. This is done for illustration purposes, then the real hedging investigation starts by focusing on strategies based on financial instruments potentially traded in the real world.

It is highlighted that pure risk sensitivities hedging cannot work without identifying the appropriate instruments to use. Particularly insightful from this point of view is the case of 1M ATM options which, although providing a decent fit in terms of Delta-Gamma-Vega matching, results in a terminal PnL distribution from a certain point of view (left-tail cases) even worse than the completely unhedged case. The Chapter then defines a specific hedging strategy based on the Last Liquid Point concept, which shows immediately very good results by only using the ATM level and improves further in case also OTM/ITM strategies are introduced.

The Chapter also provides an extensive analysis on the ongoing hedge slippage to further identify the best strategy. It turns then out that the ATM/ITM/OTM strategy effectiveness deteriorates as far as time passes and the time to maturity of the target put reduces, probably due to the fact that the dynamically reset ATM/ITM/OTM strikes may become more and more far from the strike level of the target put. A last strategy, matching the target put features as soon as its expiry reaches the LLP is also shown. This would be the best hedging strategy found so far (although not really feasible in practice).

Also strategies based on different LLPs are tested (e.g. 2 and 10 years). For such cases, we observe that the 2 years already provides good hedge results which decently improve if we go with 5 years. We don't observe the same level of improvement between 5 and 10 years though, sign that we would expect the hedge results to stabilize if a minimum expiry requirement for the hedge puts is satisfied.

Finally, a real-life like exercise is performed in order to reduce as much as possible (both in terms of times over the year and quantities) the option trading. This is done by imposing a yearly rebalancing only of the put hedge units, while monthly adjusting the delta hedging, unless certain specific events occur (vega limit breach/big movements of the underlying price with respect to the strike). Here we find that triggering new trades based on a low vega limit versus a higher one does not massively improve the hedge results.

Chapter 3 investigates how a financial operator could start dealing with long term equity derivatives in a non complete market, by moving away from the traditional hedging process based on risk sensitivities to the one based on global risk minimization, through the usage of AI techniques and, more specifically, Neural Networks.

The Chapter offers a general overview of the global risk minimization hedging based on the Literature available on the topic, readjusting it to be compatible with the hedging strategy based on the LLP proposed in Chapter 2.

After setting up the hedging problem, the focus is shifted on the definition of the technology to be used to solve it: a Long Short-Term Memory Neural Network with 3 layers of 64 neurons each. The structure of the customized loss function to be optimized is also extensively detailed.

Several numerical test cases to assess the performance of global risk minimization hedging with NNs are also provided:

1. Base case on a plain vanilla ATM European Put option with 20 years strike where the theoretical price is given as an input and the NN has to find the best hedging strategy;
2. More complex case where the NN does not know the theoretical price, which becomes therefore a variable to find by the NN.
3. Application of this technique to the case of an exotic Cliquet Put which may represent an attractive product to sell for a Life Insurance Company.

A real quotes test case is also performed, with model parameters calibrated on Eurostoxx50 Implied Volatilities, together with a case where the underlying stock price process gets generated under a stochastic volatility model with displacement and jumps.

In all these cases, the hedging results combining the Neural Network with the hedging strategy based on the last liquid point performs very well, generally matching very closely the theoretical model price.

Some useful Python code to complement the Thesis content is also provided in the two Appendices.

Hopefully, this Thesis will both serve as a good starting point and give insightful ideas to everybody interested in the VAs world, with an innovative touch to combine well established pricing and risk management techniques with new cutting-edge tools. Many topics have been illustrated and detailed, with several practical examples provided, too. There can also be interesting inspirations for future research. For example, extending this approach to the case where also trading costs are taken into account should be rather feasible by properly adjusting the loss function to optimize and the basic principles of the hedging strategy in order to reduce as much as possible the rebalancings, while still ensuring a good effectiveness. This may also serve to calculate a fair valuation adjustment to the VA pricing. Also, making the optimization problem more realistic by focusing on additional financial and non-financial risk factors may be extremely valuable from a practical perspective.

Bibliography

[1] EIOPA, *Report on Variable Annuities.* 2011.

[2] A. R. Bacinello, P. Millossovich, A. Olivieri, and E. Pitacco, "Variable annuities: A unifying valuation approach," *Insurance: Mathematics and Economics*, vol. 49, no. 3, pp. 285–297, 2011.

[3] M. De Felice and F. Moriconi, "Market based tools for managing the life insurance company," *ASTIN Bulletin: The Journal of the IAA*, vol. 35, no. 1, pp. 79–111, 2005.

[4] G. Castellani, M. De Felice, F. Moriconi, and C. Pacati, "Pricing formulae for financial options and guarantees embedded in profit sharing life insurance policies," tech. rep., Citeseer, 2007.

[5] I. Chahboun and N. Hoover, "Variable annuities: Underlying risks and sensitivities," *FRB Boston Risk and Policy Analysis Unit Paper No. RPA*, pp. 19–1, 2019.

[6] F. Ruez, "Variable annuities with guaranteed lifetime withdrawal benefits: an analysis of risk-based capital requirements," tech. rep., Working paper, Ulm University, 2017.

[7] P. Carr and L. Wu, "Static hedging of standard options," 2002.

[8] P. Banerjee, S. Iyer, and S. Jain, "Multi-period static hedging of european options," *arXiv preprint arXiv:2310.01104*, 2023.

[9] F. Black and M. Scholes, "The pricing of options and corporate liabilities," *The Journal of Political Economy*, vol. 81, no. 3, pp. 637–654, 1973.

[10] N. N. Taleb, *Dynamic Hedging: Managing Vanilla and Exotic Options.* John Wiley & Sons, 1997.

[11] N. El Karoui and M.-C. Quenez, "A super-hedging approach to stochastic volatility models," *Mathematics of Operations Research*, vol. 20, no. 4, pp. 853–876, 1995.

[12] J. Cotter and J. Hanly, "Reevaluating hedging performance," *Journal of Futures Markets: Futures, Options, and Other Derivative Products*, vol. 26, no. 7, pp. 677–702, 2006.

[13] S. L. Heston, "A closed-form solution for options with stochastic volatility with applications to bond and currency options," *The review of financial studies*, vol. 6, no. 2, pp. 327–343, 1993.

[14] B. Dupire *et al.*, "Pricing with a smile," *Risk*, vol. 7, no. 1, pp. 18–20, 1994.

[15] C. Alexander and L. M. Nogueira, "Hedging with stochastic local volatility," 2004.

[16] F. D. Rouah, *The Heston Model and Its Extensions in Matlab and C.* Wiley, 2013.

[17] H. Albrecher, P. Mayer, W. Schoutens, and J. Tistaert, "The little heston trap," *Wilmott*, no. 1, pp. 83–92, 2007.

[18] J. Gatheral, *The volatility surface: a practitioner's guide.* John Wiley & Sons, 2006.

[19] C. Pacati, R. Renò, and M. Santilli, "Heston model: shifting on the volatility surface," *Risk*, p. 54, 2014.

[20] M. Augustyniak, F. Godin, and C. Simard, "Assessing the effectiveness of local and global quadratic hedging under garch models," *Quantitative Finance*, vol. 17, no. 9, pp. 1305–1318, 2017.

[21] M. Schweizer, "A guided tour through quadratic hedging approaches," *Option pricing, interest rates and risk management*, pp. 538–574, 2001.

[22] T. F. Coleman, Y. Li, and M.-C. Patron, "Discrete hedging under piecewise linear risk-minimization," tech. rep., Cornell University, 2002.

[23] A. Černỳ, "Dynamic programming and mean-variance hedging in discrete time," *Applied Mathematical Finance*, vol. 11, no. 1, pp. 1–25, 2004.

[24] A. Carbonneau, "Deep hedging of long-term financial derivatives," *Insurance: Mathematics and Economics*, vol. 99, pp. 327–340, 2021.

[25] H. Buehler, L. Gonon, J. Teichmann, and B. Wood, "Deep hedging," *Quantitative Finance*, vol. 19, no. 8, pp. 1271–1291, 2019.

[26] P. Carr and D. Madan, "Option valuation using the fast fourier transform," *Journal of computational finance*, vol. 2, no. 4, pp. 61–73, 1999.

[27] D. Doyle and C. Groendyke, "Using neural networks to price and hedge variable annuity guarantees," *Risks*, vol. 7, no. 1, p. 1, 2018.

[28] R. Storn and K. Price, "Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces," *Journal of global optimization*, vol. 11, pp. 341–359, 1997.

[29] D. S. Bates, "Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options," *The Review of Financial Studies*, vol. 9, no. 1, pp. 69–107, 1996.

Appendix A

A.1. Displaced Heston Model Simulation Scheme

The following Python code implements the displaced Heston model simulation scheme:

```
def heston(scheme, negvar, numPaths, rho, S_0, V_0, T, dt, kappa,
    theta, sigma, r, q, phi, sim='RN', lambda_=0, mu=0):
    """
    Simulates paths using the Heston stochastic volatility model.

    Parameters:
    -----------
    scheme : str
        Numerical scheme ('Euler' or 'Milstein').
    negvar : str
        Method for handling negative variance ('Reflect' or '
           Trunca').
    numPaths : int
        Number of simulation paths.
    rho : float
        Correlation between Brownian motions.
    S_0 : float
        Initial stock price.
    V_0 : float
        Initial variance.
    T : float
        Maturity time (in years).
    dt : float
        Discretization interval for simulation.
    kappa : float
        Mean reversion rate for variance.
    theta : float
        Long-term mean variance.
    sigma : float
        Volatility of variance (vol-of-vol).
    r : float
        Risk-free interest rate.
    q : float
        Dividend yield.
```

```
    phi : float
        Displacement factor for variance.
    sim : str
        Simulation type ('RN' for risk-neutral, 'RW' for real-
           world).
    lambda_ : float, optional
        Market price of volatility risk (default is 0).
    mu : float, optional
        Drift for real-world simulations (default is 0).

    Returns:
    --------
    S : ndarray
        Simulated stock price paths.
    V : ndarray
        Simulated variance paths.
    rho_adj : ndarray
        Adjusted correlation for each time step.
    """
    np.random.seed(123456)
    num_time = int(T/dt)
    sqrt_dt = np.sqrt(dt)
    S = np.zeros((num_time+1, numPaths))  # Stock price paths
       initialization
    S[0,:] = S_0
    V = np.zeros((num_time+1, numPaths))  # Variance paths
       initialization
    V[0,:] = V_0
    Vcount0 = 0
    Zs = np.zeros((num_time+1, numPaths))  # Correlated random
       variables
    rho_adj = np.zeros((num_time+1, numPaths))  # Adjusted
       correlation
    rho_adj[0,:] = rho
    phi_t = 0  # Variance adjustment factor

    for t_step in range(1, num_time+1):
        Zv = np.random.randn(numPaths)  # Random noise for
           variance
```

```
        Zs_temp = np.random.randn(numPaths)  # Random noise for
           stock prices

        # Update variance process based on chosen scheme
        if scheme == 'Euler':
            V[t_step] = V[t_step-1] + kappa*(theta-V[t_step-1])*dt
                + sigma*np.sqrt(V[t_step-1])*(sqrt_dt*Zv)
        elif scheme == 'Milstein':
            V[t_step] = V[t_step-1] + kappa*(theta-V[t_step-1])*dt
                + sigma*np.sqrt(V[t_step-1])*(sqrt_dt*Zv) + 1/4*
               sigma**2*dt*((Zv)**2 -1)

        # Handle negative variance
        if negvar == 'Reflect':
            V[t_step] = np.abs(V[t_step])
        elif negvar == 'Trunca':
            V[t_step] = np.maximum(V[t_step], 0.00001)

        Vcount0 += np.sum(V[t_step] <= 0)  # Track occurrences of
           non-positive variance

        phi_t = phi  # Adjusted variance term

        # Adjust correlation dynamically if needed
        if phi == 0:
            rho_adj[t_step] = rho
        else:
            rho_adj[t_step] = rho*np.sqrt(V[t_step]/(V[t_step]+
               phi_t))

        # Update stock price paths
        Zs[t_step] = (rho_adj[t_step-1])*Zv + np.sqrt(1-(rho_adj[
           t_step-1])**2)*Zs_temp
        if sim == 'RN':
            S[t_step] = S[t_step-1] * np.exp((r-q-(V[t_step-1]+
               phi_t)/2)*dt + np.sqrt(V[t_step-1] + phi_t)*sqrt_dt
               *Zs[t_step])
        elif sim == 'RW':
            S[t_step] = S[t_step-1] * np.exp((mu-q-(V[t_step-1]+
```

```
                phi_t)/2)*dt + np.sqrt(V[t_step-1] + phi_t)*(
                sqrt_dt*Zs[t_step]))

    return np.transpose(S), np.transpose(np.sqrt(V)), np.transpose
        (rho_adj)
```

A.2. Displaced Heston Model Pricing Through Characteristic Function

The following Python code implements the displaced Heston model theoretical pricing of European plain vanilla options using the Characteristic Function with both the Gauss-Laguerre and Newton-Coates quadrature rules.

A.2.1. Heston Price with Gauss-Laguerre

```
def HestonPriceGaussLaguerre(PutCall, S, K, T, r, q, kappa, theta,
    sigma, lambd, v0, rho, trap, x, w, phi_disp=0.):
    """
    Heston (1993) call or put price by Gauss-Laguerre Quadrature
    Uses the original Heston formulation of the characteristic
        function,
    or the "Little Heston Trap" formulation of Albrecher et al.
    """
    # Ensure all inputs are arrays
    S = np.asarray(S)
    K = np.asarray(K)
    T = np.asarray(T)
    v0 = np.asarray(v0)

    # Initialize integration results
    int1 = np.zeros_like(S)
    int2 = np.zeros_like(S)

    # Gauss-Laguerre integration
    for k in range(len(x)):
        int1 += w[k] * HestonProb(x[k], kappa, theta, lambd, rho,
            sigma, T, K, S, r, q, v0, 1, trap, phi_disp)
```

```
        int2 += w[k] * HestonProb(x[k], kappa, theta, lambd, rho, sigma, T, K, S, r, q, v0, 2, trap, phi_disp)

    # Risk-neutral probabilities
    P1 = 0.5 + 1/np.pi * int1
    P2 = 0.5 + 1/np.pi * int2

    # Option prices
    HestonC = S * np.exp(-q * T) * P1 - K * np.exp(-r * T) * P2
    HestonP = np.maximum(HestonC - S * np.exp(-q * T) + K * np.exp(-r * T), 0.001)

    # Return call or put price
    if PutCall == 'C':
        return HestonC
    else:
        return HestonP
```

A.2.2. Heston Integrand

```
def HestonProb(phi, kappa, theta, lambd, rho, sigma, tau, K, S, r, q, v, Pnum, Trap, phi_disp):
    """
    Returns the integrand for the risk-neutral probabilities P1 and P2.
    """
    # Ensure all inputs are arrays
    phi = np.asarray(phi)
    K = np.asarray(K)
    S = np.asarray(S)
    v = np.asarray(v)

    # Log of the stock price
    x = np.log(S)

    # Parameter "a" is the same for P1 and P2
    a = kappa * theta

    # Parameters "u" and "b" are different for P1 and P2
```

```
    u = 0.5 if Pnum == 1 else -0.5
    b = kappa + lambd - rho * sigma if Pnum == 1 else kappa +
       lambd

    # Characteristic function
    d = np.sqrt((rho * sigma * 1j * phi - b)**2 - sigma**2 * (2 *
       u * 1j * phi - phi**2))
    g_denominator = b - rho * sigma * 1j * phi - d
    g = (b - rho * sigma * 1j * phi + d) / g_denominator

    if Trap == 1:
        # "Little Heston Trap" formulation
        c = 1 / g
        D = (b - rho * sigma * 1j * phi - d) / sigma**2 * ((1 - np
           .exp(-d * tau)) / (1 - c * np.exp(-d * tau)))
        C = (r - q) * 1j * phi * tau + a / sigma**2 * ((b - rho *
           sigma * 1j * phi - d) * tau - 2 * np.log(1 - c * np.exp
           (-d * tau)))
    else:
        # Original Heston formulation
        G = (1 - g * np.exp(d * tau)) / (1 - g)
        C = (r - q) * 1j * phi * tau + a / sigma**2 * ((b - rho *
           sigma * 1j * phi + d) * tau - 2 * np.log(G))
        D = (b - rho * sigma * 1j * phi + d) / sigma**2 * ((1 - np
           .exp(d * tau)) / (1 - g * np.exp(d * tau)))

    # The characteristic function
    f = np.exp(C + D * v + 1j * phi * x)

    # Avoid division by zero in the integrand
    integrand = np.exp(-1j * phi * np.log(K)) * f / (1j * phi)
    return np.real(integrand)
```

A.2.3. Heston Price with Newton-Coates

```
def HestonPriceNewtonCoates(PutCall, S, K, T, r, q, kappa, theta,
      sigma, lambd, v0, rho, trap, method, phi_disp, a, b, N):
    """
    Heston (1993) Call price by Newton-Coates Formulas.
```

```
    """
    # Ensure inputs are arrays
    S = np.asarray(S)
    K = np.asarray(K)
    T = np.asarray(T)
    v0 = np.asarray(v0)

    # Build the integration grid
    h = (b - a) / (N - 1)
    phi = np.linspace(a, b, N)

    # Initialize weights
    if method == 1:
        wt = h * np.ones(N)
    elif method == 2:
        wt = h * np.concatenate(([1/2], np.ones(N - 2), [1/2]))
    elif method == 3:
        wt = (h / 3) * np.concatenate(([1], np.where(np.arange(2, N) % 2 == 0, 4, 2), [1]))
    elif method == 4:
        N = N - N % 3 + 1  # Ensure N-1 is divisible by 3
        h = (b - a) / (N - 1)
        wt = (3 * h / 8) * np.concatenate(([1, 3, 3], np.tile([2, 3, 3], (N - 1) // 3 - 1), [1]))
        phi = np.linspace(a, b, N)
    else:
        raise ValueError("Invalid method. Choose from 1, 2, 3, or 4.")

    # Compute the integrals
    int1 = np.zeros_like(S)
    int2 = np.zeros_like(S)
    for k in range(N):
        int1 += wt[k] * HestonProb(phi[k], kappa, theta, lambd, rho, sigma, T, K, S, r, q, v0, 1, trap, phi_disp)
        int2 += wt[k] * HestonProb(phi[k], kappa, theta, lambd, rho, sigma, T, K, S, r, q, v0, 2, trap, phi_disp)

    # The probabilities P1 and P2
```

```
    P1 = 1/2 + 1/np.pi * int1
    P2 = 1/2 + 1/np.pi * int2

    # The call price
    HestonC = S * np.exp(-q * T) * P1 - K * np.exp(-r * T) * P2

    # The put price by put-call parity
    HestonP = np.maximum(HestonC - S * np.exp(-q * T) + K * np.exp
       (-r * T), 0.01)

    # Output the option price
    return HestonC if PutCall == 'C' else HestonP
```

Appendix B

B.1. Complex LSTM Model for Derivatives Hedging with Global Risk Minimization

B.1.1. Model Definition

```
class ComplexLSTM(nn.Module):
    def __init__(self, input_size, hidden_size, num_layers,
       output_size):
        super(ComplexLSTM, self).__init__()
        self.hidden_size = hidden_size
        self.num_layers = num_layers

        # LSTM layer to learn temporal dependencies
        self.lstm = nn.LSTM(input_size, hidden_size, num_layers,
           batch_first=True)

        # Fully connected layer maps hidden state to output
        self.fc = nn.Linear(hidden_size, output_size)

    def forward(self, x):
        batch_size = x.size(0)  # Get batch size

        # Initialize hidden and cell states with zeros
        h0 = torch.zeros(self.num_layers, batch_size, self.
           hidden_size).to(x.device)
        c0 = torch.zeros(self.num_layers, batch_size, self.
           hidden_size).to(x.device)

        # Pass input through LSTM layer
        out, _ = self.lstm(x, (h0, c0))

        # Apply fully connected layer to LSTM output
        out = self.fc(out)

        # Flatten output to (batch_size, output_size * seq_length)
        out = out.reshape(out.size(0), -1)
        return out
```

B.1.2. Data Preparation

```
# Get number of samples and sequence length
num_samples = St_H_RW_sel.shape[0]
seq_length = St_H_RW_sel.shape[1]

# Ensure consistency across all input arrays
arrays = [
    St_H_RW_sel, T_Matrix_Puthedge, Puthedge_Price_t_NN,
    Puthedge_Price_t_NN_ITM, Puthedge_Price_t_NN_OTM,
    T_Matrix_all, K_target_Matrix, Puthedge_K_Matrix_NN,
    Puthedge_K_Matrix_ITM_NN, Puthedge_K_Matrix_OTM_NN
]

# Confirm that shapes match
for array in arrays:
    assert array.shape == (num_samples, seq_length), f"Shape
       mismatch: {array.shape}"

# Stack features into a single input tensor
X = np.stack(arrays, axis=-1)

# Create dataset indices
indices = np.arange(X.shape[0])

# Split into training+validation and test sets
train_val_indices, test_indices = train_test_split(indices,
   test_size=0.2, random_state=42)

# Further split train+val into train and val sets
train_indices, val_indices = train_test_split(train_val_indices,
   test_size=0.25, random_state=42)

# Create final splits
X_train, X_val, X_test = X[train_indices], X[val_indices], X[
   test_indices]
y_train = Put_Target_t[train_indices][:, -1]
```

```
y_val = Put_Target_t[val_indices][:, -1]
y_test = Put_Target_t[test_indices][:, -1]

# Convert to PyTorch tensors and move to device
device = torch.device('cuda' if torch.cuda.is_available() else '
    cpu')
X_train = torch.tensor(X_train, dtype=torch.float32).to(device)
X_val = torch.tensor(X_val, dtype=torch.float32).to(device)
X_test = torch.tensor(X_test, dtype=torch.float32).to(device)
y_train = torch.tensor(y_train, dtype=torch.float32).to(device)
y_val = torch.tensor(y_val, dtype=torch.float32).to(device)
y_test = torch.tensor(y_test, dtype=torch.float32).to(device)
```

B.1.3. Custom Loss Function

```
def custom_loss_function(outputs, Put_Target_t, St_H_RW_sel,
    Puthedge_Price_t,
                         Puthedge_Price_t_ITM,
                            Puthedge_Price_t_OTM,
                            Puthedge_Price_t1,
                         Puthedge_Price_t1_ITM,
                            Puthedge_Price_t1_OTM, r, dt, q):
    """
    Custom loss: measures hedging portfolio performance.
    """
    # Reshape flat outputs to (batch, seq_length, output_size)
    num_samples, total_features = outputs.shape
    seq_length = St_H_RW_sel.shape[1]
    outputs = outputs.view(num_samples, seq_length, output_size)

    # Extract relevant predicted values
    Put_target_price_t = outputs[:, :, 0]
    S_units = outputs[:, :, 1]
    Puthedge_units_ATM = outputs[:, :, 2]
    Puthedge_units_ITM = outputs[:, :, 3]
    Puthedge_units_OTM = outputs[:, :, 4]

    # Initialize portfolio with option value
```

```
    HedgePtf_real = torch.zeros(St_H_RW_sel.shape).to(outputs.device)
    HedgePtf_real[:, 0] = Put_target_price_t[:, 0]

    # Update portfolio value at each time step
    for j in range(1, seq_length):
        HedgePtf_real[:, j] = (
            (HedgePtf_real[:, j-1]
             - Puthedge_units_ATM[:, j-1] * Puthedge_Price_t[:, j-1]
             - Puthedge_units_ITM[:, j-1] * Puthedge_Price_t_ITM[:, j-1]
             - Puthedge_units_OTM[:, j-1] * Puthedge_Price_t_OTM[:, j-1]
            ) * np.exp(r * dt)  # Risk-free return
            + S_units[:, j-1] * (St_H_RW_sel[:, j] - St_H_RW_sel[:, j-1])  # Delta exposure
            - S_units[:, j-1] * St_H_RW_sel[:, j-1] * (np.exp((r - q) * dt) - 1)  # Dividend yield
            + Puthedge_units_ATM[:, j-1] * Puthedge_Price_t1[:, j-1]
            + Puthedge_units_ITM[:, j-1] * Puthedge_Price_t1_ITM[:, j-1]
            + Puthedge_units_OTM[:, j-1] * Puthedge_Price_t1_OTM[:, j-1]
        )

    # Compute squared error at final time step
    error = torch.sum((HedgePtf_real[:, -1] - Put_Target_t) ** 2)
    return error
```

B.1.4. Training the Model

```
# Wrap data in TensorDatasets
train_dataset = TensorDataset(X_train, y_train)
val_dataset = TensorDataset(X_val, y_val)
test_dataset = TensorDataset(X_test, y_test)

# DataLoaders for batching
```

```python
batch_size = 32
train_loader = DataLoader(train_dataset, batch_size=batch_size,
    shuffle=True)
val_loader = DataLoader(val_dataset, batch_size=batch_size,
    shuffle=False)
test_loader = DataLoader(test_dataset, batch_size=batch_size,
    shuffle=False)

# Model hyperparameters
input_size = X_train.shape[-1]
hidden_size = 64
num_layers = 3
output_size = 5

# Initialize model and optimizer
model_LSTM = ComplexLSTM(input_size, hidden_size, num_layers,
    output_size).to(device)
optimizer = optim.Adam(model_LSTM.parameters(), lr=0.001)

num_epochs = 2000
patience = 400  # Early stopping patience

# Early stopping parameters
best_val_loss = float('inf')
patience_counter = 0

# Training loop
for epoch in range(num_epochs):
    model_LSTM.train()
    epoch_loss = 0.0

    for batch_X, batch_y in train_loader:
        optimizer.zero_grad()
        outputs = model_LSTM(batch_X)
        loss = custom_loss_function(outputs, batch_y, ...)
        loss.backward()
        optimizer.step()
        epoch_loss += loss.item()

```

```
    # Evaluate on validation set
    model_LSTM.eval()
    val_loss = 0.0
    with torch.no_grad():
        for batch_X_val, batch_y_val in val_loader:
            val_outputs = model_LSTM(batch_X_val)
            val_loss += custom_loss_function(val_outputs,
               batch_y_val, ...).item()

    # Save best model if validation loss improves
    if val_loss < best_val_loss:
        best_val_loss = val_loss
        patience_counter = 0
        torch.save(model_LSTM.state_dict(), 'best_model.pth')
    else:
        patience_counter += 1

    # Stop training if patience exhausted
    if patience_counter >= patience:
        print("Early stopping triggered")
        break
```