Topics in Derivative Pricing Coursework
Micro RV

Waner Chen
CID: 01854827

Zhenyi Chen
CID: 06011402

Chengdong Song
CID: 02030540

April 21, 2025

Abstract

This project investigates Micro Relative Value (Micro RV) opportunities by analyzing localized dislocations in dealer-constructed interest rate curves. Using historical swap data across multiple currencies, we construct daily yield curves via natural cubic spline interpolation and identify curvature-based anomalies through butterfly spreads. We evaluate each structure using rolling-window z-scores and carry metrics, and select the most statistically significant configurations for numerical experiments. The profitability of these trades is assessed through DV01-adjusted PnL simulations under dynamically evolving maturities. Our empirical analysis shows strong profitability in GBP and USD markets, where dislocations are more stable and persistent. By contrast, EUR results highlight the importance of market-specific calibration and risk control, as performance is more fragile and sensitive to outliers.

1 Introduction

Interest rate term structures exhibit complex dynamics driven by macroeconomic fundamentals, central bank policy expectations, and market liquidity. Traditional curve trades—including level, slope, and curvature strategies—are often designed to express directional views on the shape of the yield curve, and have been studied extensively both theoretically and empirically [2][3].

Among these, butterfly trades serve as canonical instruments for isolating curvature risk. Typically constructed via a weighted combination of fixed-rate instruments or swaps with three distinct maturities, butterfly structures are level-neutral, allowing traders to express views on localized convexity in the term structure. The mathematical underpinnings of butterfly positioning as a discrete second derivative approximation have been widely discussed in the literature [4].

This project mainly investigates the curvature-based perspective into the domain of Micro relative value (Micro RV) trading [1], a class of strategies aimed at exploiting small, localized misalignments in the interpolated yield curve. Such distortions often arise from non-economic factors—such as sparse quoting, interpolation artifacts, or asynchronous market updates—and are not always justified by macroeconomic data.

In this project, we identify Micro RV opportunities using historical swap rate data across various currencies. We construct daily interest rate curves from Bloomberg data using C^2 natural cubic splines on quoted market tenors. Each day's curve is used to compute forward swap rates and derive a large number of butterfly trade structures of the form (a, b, c) representing the fly spread $2r_b - r_a - r_c$. These structures serve as the primary instruments for expressing local curvature views on the yield curve.

To quantify the attractiveness of a given fly trade, we compute rolling-window standardized scores (z-scores) for each structure over time, measuring their deviation from historical norms. we also incorporate carry—the expected PnL assuming curve persistence—to further refine trade selection. The top N structures with the most statistically significant dislocations are selected each day for further testing. For each selected structure, we simulate its mark-to-market evolution and compute the associated PnL using DV01-scaled sensitivities to ensure fair comparison across trades with different maturities and volatilities. The PnL is computed dynamically, with maturities rolling down through time to ensure the trade consistently tracks the same terminal maturity rather than fixed tenor.

The remainder of this report is organized as follows: Section 2 provides a comprehensive review of the theoretical foundations underpinning butterfly trades, including detailed discussions of Micro Relative Value, carry, z-scores metrics, DV01, and profit and loss calculations. Section 3 outlines the methodological framework adopted for conducting our analyses and strategies. In Section 4, we present numerical results derived from our empirical experiments, accompanied by insightful analyses of these findings. Finally, Section 5 summarizes the primary conclusions drawn from this study.

2 Theoretical Background

Here we present key theoretical concepts essential to understanding the study's framework.

2.1 Flies

In fixed income markets, butterfly trades—commonly referred to as "flies"—represent interest rate strategies designed primarily to capture changes in the curvature of the yield curve, often measured as its second derivative. Unlike outright duration positions or slope trades (steepeners/flatteners), butterflies aim to be relatively neutral to parallel shifts in the curve and express a view on its second derivative—the rate of change in slope.

Let $T_a < T_b < T_c$ be three selected maturities on the yield curve, with corresponding interest rates denoted by $r_a(t), r_b(t), r_c(t)$ at time t, respectively. A butterfly structure is defined as a linear combination of these rates:

$$\text{Fly}_{a,b,c}(t) = w_a \cdot r_a(t) + w_b \cdot r_b(t) + w_c \cdot r_c(t).$$

The weights $(w_a, w_b, w_c) \in \mathbb{R}^3$ are chosen to satisfy:

$$w_a + w_b + w_c = 0 \quad \text{(level neutrality)}.$$

A common choice is $(-1, 2, -1)$, yielding:

$$\text{Fly}(t) = -r_a(t) + 2r_b(t) - r_c(t)$$

This structure approximates a discrete second difference operator on the term structure and serves as a finite-difference proxy for the local curvature of the yield curve in the interval $[T_a, T_c]$. While some formulations additionally impose slope-neutrality (e.g., $w_aT_a + w_bT_b + w_cT_c = 0$), this is not necessary in general and is often relaxed in practice in favor of flexibility in notional sizing.

The curvature of the yield curve refers to the local convexity or concavity of the term structure and captures the rate at which slope itself changes across maturities. Intuitively, if we consider the yield curve as a smooth geometric object, curvature at a point reflects whether the curve is "humped" (concave) or "bent upward" (convex) in the local region.

In financial terms, curvature captures how intermediate maturities deviate from the average of surrounding tenors. Specifically, if

$$r_b(t) < \frac{1}{2}\left(r_a(t) + r_c(t)\right),$$

then the curve is concave downward at T_b, indicating richness in the belly. Conversely, if r_b is above the average, the curve is convex upward, implying cheapness in the belly.

Butterfly trades are constructed to monetize such local shape anomalies. A long butterfly (receiving the intermediate leg and paying the outer legs) profits when the intermediate rate declines relative to the wings—that is, when curvature increases downward. This can reflect either fundamental expectations (e.g., upcoming rate cuts) or flow/technical-driven mispricings in the belly of the curve.

2.2 Micro Relative Value (Micro RV)

In highly liquid fixed income markets, such as government bond or swap curves, the construction of the full yield curve typically involves interpolation between a discrete set of observable instruments. For example, market-makers or data vendors may have reliable quotes at anchor points like 1y, 5y, 10y, 20y, and 30y maturities. However, quotes for intermediate tenors such as 9y, 11y, or 14y may be either unavailable or derived through internal model-based interpolation.

As a result, the full curve may display localized irregularities due to: asynchronous updating of benchmark tenors, interpolation artifacts arising from algorithmic choices, misaligned quoting behavior between brokers and dealers, etc. These localized irregularities are not justified by macroeconomic or fundamental drivers and instead reflect microstructure frictions or technical imperfections in curve building.

Micro RV trading strategies seek to systematically identify and exploit such anomalies by targeting relative misalignments between nearby forward rates. These trades typically do not require a global curve view but instead rely on the assumption that the curve should be locally smooth and arbitrage-free.

Let $y(t,T)$ be the continuously-compounded zero rate at time t for maturity T. The instantaneous forward rate is defined as:

$$f(t,T) = \frac{\partial}{\partial T}\left[y(t,T)\cdot T\right].$$

Given this definition, a well-constructed curve should produce a smooth forward rate function $f(t,T)$. Deviations from smoothness in small neighborhoods of T are indicative of local interpolation errors.

A Micro RV signal is then constructed by considering second-order forward differences across neighboring tenors. For example, define the discrete forward butterfly signal around a center maturity T_b with wings $T_a < T_b < T_c$ as:

$$\mathrm{MicroRV}(t) = -f(t,T_a) + 2f(t,T_b) - f(t,T_c).$$

Under ideal conditions, this expression should be approximately zero. A large deviation is interpreted as an anomalous bump or kink in the curve, suggesting that T_b is mispriced relative to its neighbors.

Once a signal is identified, a Micro RV trade is executed. Profits arise as the curve "heals" over time—either due to update synchronization or as the position rolls into a more liquid part of the curve.

2.3 Carry

In the context of interest rate relative value strategies, particularly butterfly structures, the carry of a trade is defined as the present value of all known cashflows over a short horizon (typically 3 or 6 months), assuming the yield curve remains unchanged. It reflects the static income potential of the position under curve persistence, independent of mark-to-market fluctuations.

For a standard Libor swap, the carry can be approximated as:

$$\text{Carry} = H\tau_0 \left(L_0(0) - k\right),$$

where H is the notional, τ_0 is the year fraction, $L_0(0)$ is the spot Libor rate known at trade inception, and k is the fixed coupon rate of the swap. This expression captures the expected cashflow difference between floating and fixed legs, scaled by notional.

In a butterfly trade, which consists of a weighted portfolio of multiple swaps, the overall carry is computed as the notional-weighted sum of the individual swap carries. This approach isolates structural roll-down potential and curve shape asymmetry.

Roll-down, in contrast, quantifies the change in a trade's value as it shifts along the forward curve due to time passing. For vanilla swaps or bonds, this is well-defined. However, for short-dated or synthetic fly structures involving offsetting positions (e.g., long 10y1y, short 9y1y and 11y1y), the forward leg at center maturity becomes the new spot-starting leg as the position rolls. As a result, roll-down is not well-defined for fly portfolios.

2.4 Z-scores

To identify and filter for statistically significant curve distortions, we define the normalized z-scores, $z(t)$, for any signal $X(t)$ (such as the forward fly value):

$$z(t) = \frac{X(t) - \mu(t)}{\sigma(t)}$$

where $\mu(t)$ and $\sigma(t)$ are the rolling mean and standard deviation of X computed over a historical lookback window $[t - T, t]$. A high magnitude, such as $|z(t)| \geq 2$ indicates that the signal deviates significantly from its historical equilibrium and may represent a trading opportunity. The z-scores provides a standardized way to quantify how far a signal is from its historical center, assuming approximate normality.

2.5 DV01

The DV01 (Dollar Value of One Basis Point) is a fundamental measure of interest rate risk, quantifying the change in the present value of a financial instrument or portfolio for a parallel 1 basis point (0.01%) shift in the yield curve. For linear interest rate derivatives such as swaps, DV01 reflects the first-order sensitivity to the level of rates.

Mathematically, for a position with value V, DV01 can be expressed as:

$$\text{DV01} = \frac{\partial V}{\partial r} \cdot 0.0001$$

where r represents the relevant curve rate. The DV01 is typically quoted per unit notional, and the sign of the DV01 indicates directional exposure: a positive DV01 implies gains if rates fall (long duration), while a negative DV01 implies gains if rates rise (short duration).

In the context of multi-leg trades such as butterflies or swaps portfolios, the DV01 is often computed leg-by-leg and aggregated. For relative value trades, the goal is often to construct DV01-neutral positions (net DV01 ≈ 0), so that the strategy is immunized against parallel level shifts and more directly expresses views on curvature or slope.

2.6 Profit and Loss

Changes in the value of a fixed income portfolio over time are captured by the mark-to-market Profit and Loss (PnL). Assuming the portfolio exhibits linear sensitivity to rates, the first-order approximation of daily PnL is given by:

$$\Delta \mathrm{PnL}(t) \approx \sum_i \mathrm{DV01}_i \cdot \left(r_i(t) - r_i(t_0)\right),$$

where $r_i(t)$ is the market rate at curve point i on date t, $\mathrm{DV01}_i$ is the dollar value of a 1 basis point move in rate r_i, measured at trade initiation t_0. The sum is taken over all relevant curve tenors or buckets to which the portfolio has exposure.

This formulation treats the PnL as a dot product between the portfolio's fixed risk exposure vector (DV01s) and the observed rate changes since inception. It assumes that DV01s remain approximately constant over the holding period, which is a valid approximation for small rate moves and relatively short horizons.

Such an approach allows the simulation of PnL trajectories for arbitrary portfolios using only the initial DV01 exposure profile across curve segments and the time series of observed or simulated market rates. This PnL attribution framework provides a clean and interpretable decomposition of returns that is especially useful for backtesting systematic strategies, understanding risk-adjusted performance, and isolating contributions from specific curve movements.

Higher-order effects such as convexity (gamma), carry, or roll-down may be incorporated in more refined models, but the linear DV01-based framework remains the primary tool for fast and robust PnL simulation and risk diagnostics in practice.

3 Methodology

3.1 Data and Curve Construction

To accurately identify "Micro RV" opportunities within the chosen currency and interest rate market, it is essential to construct a precise and smooth yield curve from the provided benchmark market quotes. In this project, the given Bloomberg historical dataset consists of market quotes across various maturities, ranging from short-term tenors such as one-week and two-week instruments to longer-term instruments extending up to 50 years. The fitting of yield curves from these discrete market data points is crucial because it facilitates the identification of arbitrage opportunities and ensures consistent valuation across different maturities.

An interest rate curve or discount curve observed at time 0 provides a mapping from maturity times T to discount factors $P(T) = P(0, T)$. These discount factors represent the present value of a unit payment made at future time T.

3.1.1 Curve Calibration

The construction of a discount curve involves matching observed market prices V_i for N linear benchmark securities. Each of these securities can be represented in a general linear form as:

$$V_i = \sum_{j=1}^{M} c_{i,j} P(t_j), \quad i = 1, \ldots, N$$

where many coefficients $c_{i,j}$ may be zero.

A typical example is a standard fixed-for-floating interest rate swap. The present value V_{swap} of such a swap paying fixed coupon $c\tau$ at regular intervals $\tau, 2\tau, \ldots, n\tau$ and maturing at T_n is expressed as:

$$V_{\text{swap}} = 1 - P(T_n) - c\tau \sum_{j=1}^{n} P(j\tau)$$

At inception (par swap condition), the swap's market value V_{swap} is zero, thus:

$$1 = P(T_n) + c\tau \sum_{j=1}^{n} P(j\tau)$$

Solving explicitly for the par swap rate $S(T_n)$ yields:

$$S(T_n) = c = \frac{1 - P(T_n)}{\tau \sum_{j=1}^{n} P(j\tau)}$$

In a more generalized scenario where the swap market value is not necessarily zero, the relation becomes:

$$1 - V_{\text{swap}} = P(T_n) + c\tau \sum_{j=1}^{n} P(j\tau)$$

This formulation succinctly illustrates the dependence of benchmark securities pricing on discount factors, framing the core mathematical challenge of yield curve calibration as solving the system:

$$\mathbf{V} = \mathbf{cP}$$

where $\mathbf{V}$ is the vector of benchmark security prices, $\mathbf{P}$ is the vector of discount factors, and $\mathbf{c}$ is the coupon matrix composed of $c_{i,j}$ coefficients.

In our analysis, we began by carefully preprocessing the Bloomberg dataset. We then converted the date format into a consistent day-month-year (dd-mm-yyyy) structure and set it as our dataframe index for convenient chronological access. Following the preprocessing step, we standardized tenor labels into a uniform format, explicitly categorizing them into weekly ('W'), monthly ('M'), and yearly ('Y') maturities. Each tenor label was then systematically mapped to its corresponding numerical value expressed in years (e.g., 1M mapped to 1/12 year, 1Y mapped to 1 year).

In interpolating yield curves, it is crucial to select appropriate underlying values rather than directly interpolating discount factors $P(T)$. Direct interpolation of discount factors is typically avoided due to inherent constraints such as values bounded strictly between 0 and 1 and their typically monotonically decreasing nature, which makes direct interpolation challenging and impractical.

Instead, it is more suitable to interpolate zero rates (yield curves) $y(T)$, defined by a straightforward transformation:

$$P(T) = e^{-y(T)T}$$

Zero rates are commonly chosen because they yield smoother curves, reducing interpolation errors and providing a more realistic structure. These zero rates can be intuitively understood as an average of overnight forward rates from today up to maturity T:

$$y(T) = \frac{1}{T}\int_0^T f(u)\,du$$

We selected SONIA (Sterling Overnight Index Average) as the currency benchmark for this analysis due to its prevalent usage and liquidity in the UK money market.

3.1.2 Interpolation Methods

There are several interpolation methods applicable to yield curves, such as linear interpolation, Cubic Hermite spline, and Natural Cubic spline. Figure 1 illustrates these methods visually. However, following the project guidelines explicitly, we chose to apply a C^2 Natural Cubic spline interpolation on zero rates.

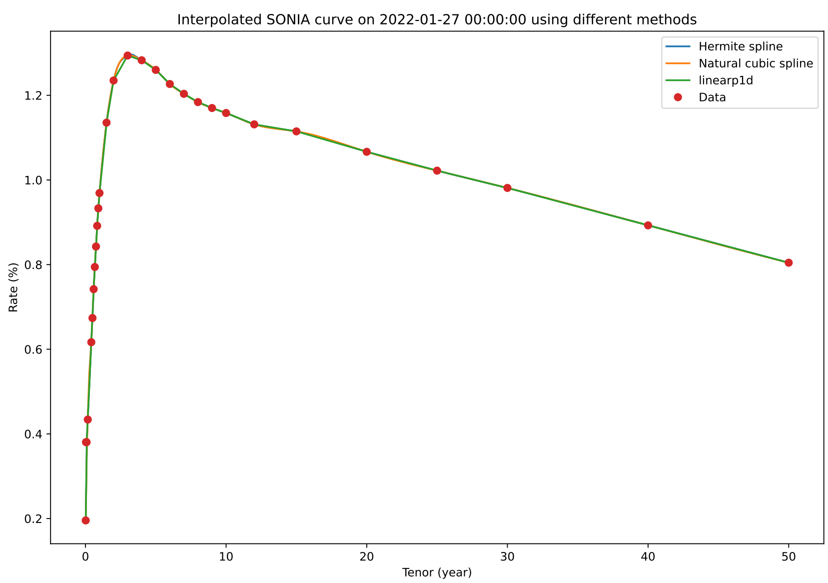


Figure 1: Interpolated SONIA Zero Curve on 2022-01-27 using Different Interpolated Methods

C^2 Natural Cubic spline interpolation is a smooth curve fitting technique characterized by its continuous first and second derivatives across the entire domain. This method inherently enforces boundary conditions that prevent unrealistic fluctuations, thus ensuring that the resulting curve remains stable and realistic for financial applications.

Figure 2 demonstrates the practical result of applying C^2 Natural Cubic spline interpolation to the SONIA zero rates on the first day of our dataset. This plot clearly shows the interpolated zero rates, overnight forward rates, and term forward rates derived from the constructed yield curve. Such curves serve as a foundational tool for subsequent pricing and trading strategy evaluations.

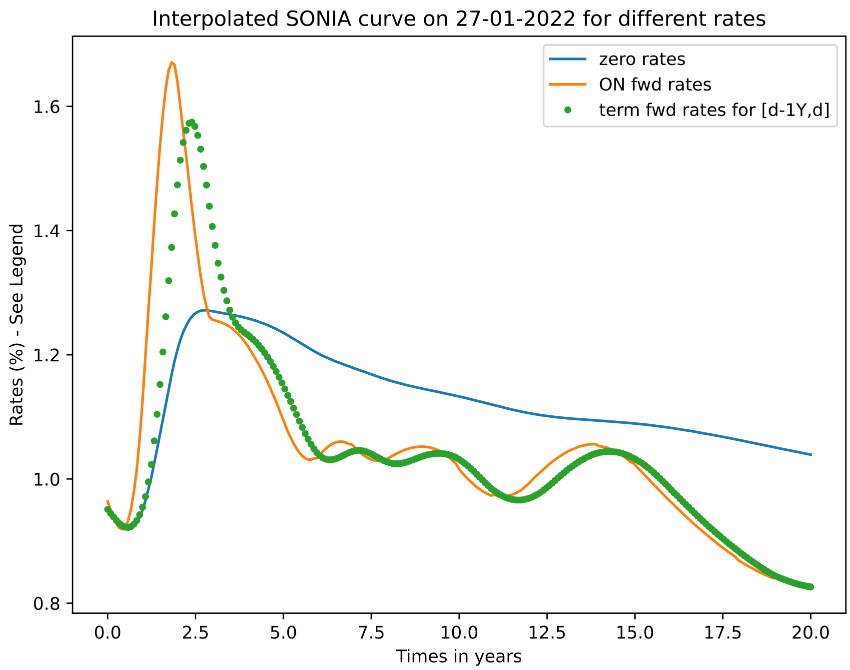


Figure 2: Interpolated SONIA Curve on 2022-01-27 for Different Curves

Using Natural Cubic spline interpolation, we further extended our yield curve analysis by calibrating SONIA swap curves. The calibration involved constructing a theoretical interest rate curve that precisely replicated observed market swap rates, ensuring internal consistency and arbitrage-free conditions. Specifically, we employed finance library functionalities to build these SONIA swap curves, accurately capturing annual coupon frequencies and the ACT/365 fixed day count convention prevalent in the UK market. Each day's curve calibration involved fitting the observed swap rates and subsequently extracting zero-coupon and forward rates for various maturities, enabling a precise and consistent valuation framework for evaluating subsequent trading strategies.

By systematically repeating the above calibration steps for each historical date in our dataset, we constructed a comprehensive set of yield curves, each accurately reflecting the prevailing market conditions. These curves serve as the foundational tool for the subsequent identification of potential "Micro RV" trading opportunities, facilitating detailed evaluations of historical trades and their resulting profit and loss profiles.

3.2 Best Trades Generation

To systematically identify the most promising Micro RV trading opportunities from historical data, we implemented a dynamic trade generation framework. This framework incorporates both statistical and economic indicators: z-scores and carry, and evaluates trade quality across multiple rolling windows to enhance the strategy's robustness and reliability.

3.2.1 Multi-Structure and Multi-Window Design

We begin by defining six butterfly structures—$(1y, 3y, 5y)$, $(1y, 4y, 7y)$, $(3y, 5y, 7y)$, $(2y, 5y, 10y)$, $(5y, 10y, 30y)$, and $(10y, 20y, 30y)$—to comprehensively capture curvature dynamics across the short, medium, and long segments of the yield curve. Each structure is denoted as a triplet (a, b, c), where a and c represent the "wings" and b the "belly". This selection ensures a broad coverage of the term structure, allowing for the analysis of curvature changes in various market conditions. The inclusion of both commonly used structures in market practice and additional configurations provides a robust framework for assessing relative value opportunities across different maturity horizons.

To evaluate signal stability over time, we analyze each structure across six different rolling windows—21, 42, 63, 84, 105, and 126 trading days—corresponding approximately to 1 to 6 calendar months. This multi-window approach enables cross-sectional comparison and helps identify the most effective historical lookback period for signal generation.

3.2.2 Trade Selection: Z-scores and Carry

For each structure and window, we computed two key indicators daily. Z-scores measures the deviation of the current fly spread from its historical mean over the rolling window. While carry estimates the expected static profit of holding a DV01-neutral butterfly structure, reflecting the economic attractiveness of the position.

To enhance the quality of selected trades, we implement a two-stage filtering procedure. In the first stage, we identify trades whose carry values rank within the top 80% across all observed carry metrics, ensuring economic attractiveness. In the second stage, we further refine this subset by identifying the top $N = 20$ trades exhibiting the largest absolute z-scores across all structures and observation dates. This sequential filtering ensures that the final selection of trades is both economically appealing and statistically significant, effectively balancing fundamental value with statistical rigor and reducing the likelihood of false positives.

3.2.3 DV01-Neutral Trade Simulation and PnL Calculation

For each selected trade, we simulate daily PnL based on DV01-neutral positioning. At the trade initiation date, we compute swap rates and corresponding DV01 values for each leg of the butterfly. The notional amounts of the two wing legs are initially set at a fixed value (e.g., \$1M). The notional of the belly leg is dynamically adjusted to ensure DV01 neutrality of the overall butterfly position. Formally, the DV01-neutral condition is represented as:

$$\mathrm{DV01}_{\mathrm{Fly}} = w_a \cdot \mathrm{DV01}_a + w_b \cdot \mathrm{DV01}_b + w_c \cdot \mathrm{DV01}_c = 0,$$

where $\mathrm{DV01}_a, \mathrm{DV01}_b, \mathrm{DV01}_c$ denote the DV01s of each leg, and w_a, w_b, w_c are their respective notional weights. Given fixed wing notionals, the belly notional w_b is computed as:

$$w_b = -\frac{w_a \cdot \mathrm{DV01}_a + w_c \cdot \mathrm{DV01}_c}{\mathrm{DV01}_b}.$$

Subsequently, we track daily swap rate evolutions for each leg. Daily PnL is calculated based on changes in swap rates, scaled by their respective DV01s and notionals. To accurately reflect real-world dynamics of interest rate swaps, maturity dates remain constant (fixed calendar dates rather than fixed tenors) throughout the simulation.

The cumulative PnL for each butterfly trade is directionally adjusted according to the trading signal—specifically, we establish short butterfly positions (selling the belly, buying the wings) for positive z-scores and long butterfly positions (buying the belly, selling the wings) for negative z-scores. For each trade, we record final realized profit, maximum drawdown, and trade duration. The resulting performance metrics and dynamics of each butterfly trade are further analyzed and visualized in the subsequent section.

4 Numerical Results and Analysis

In this section, we present the empirical results of our Micro RV strategy. The analysis is based on historical interest rate swap data across multiple currencies and evaluates the profitability, robustness, and signal quality of butterfly-based trades. The detailed source code can be found in the project Github repository (`https://github.com/cs821/IC_MF_Topics-in-Derivative-Pricing_CW.git`).

4.1 Parameter Analysis of Micro RV

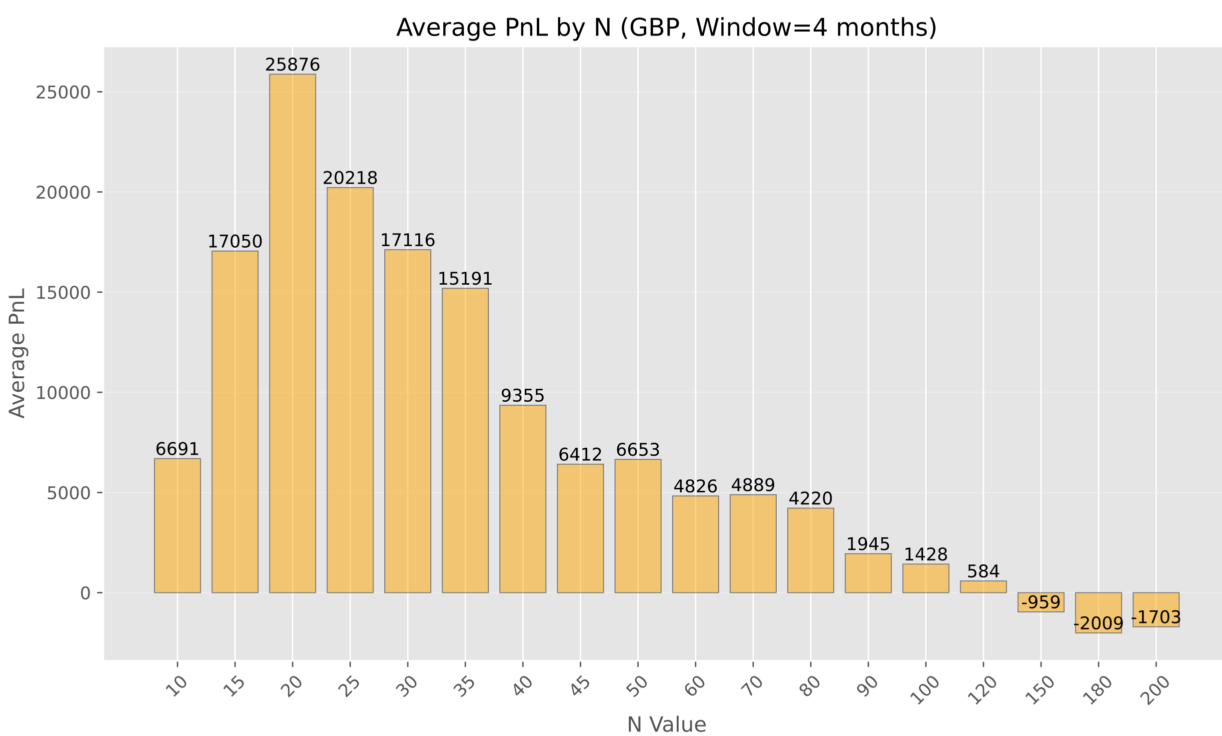


Figure 3: Average PnL by N (GBP, Window=4 months)

Figure 3 analyzes the average PnL when holding the rolling window fixed at 4 months and varying the number of best-selected trades (N). The analysis reveals that selecting too large an N

results in a decreased PnL. This decline is likely due to an influx of trades diluting the accuracy of the Micro RV signals, leading to the inclusion of less profitable or irrelevant trading opportunities. Particularly in more volatile market conditions, selecting a larger number of trades increases the likelihood of including signals that are primarily driven by short-term volatility fluctuations, rather than by genuine mean-reversion opportunities. Consequently, the optimal trade selection appears at $N = 20$, which provides a balance between trade selectivity and sufficient diversification.

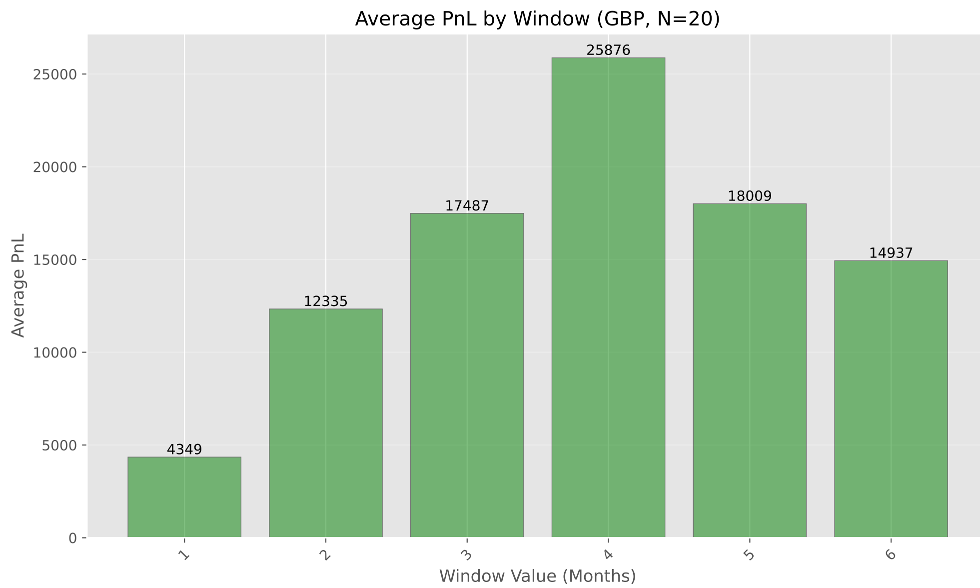


Figure 4: Average PnL by Window (GBP, N=20)

Figure 4 examines the impact of varying rolling window sizes (months) while keeping the number of best-selected trades fixed at 20. Result indicates that excessively long windows diminish trading effectiveness as historical data become overly smooth, losing responsiveness to current market conditions and thereby reducing the signal's practical relevance. Conversely, excessively short windows introduce excess noise from insufficient historical context. Consequently, a rolling window of approximately 4 months appears to offer a reasonable balance between responsiveness and noise reduction, as suggested by the observed peak in average PnL.

Importantly, the average PnL remains positive across most of parameters, highlighting the strong robustness and persistent profitability of the Micro RV framework. This consistent performance across a wide range of settings suggests that Micro RV is not only statistically effective but also economically meaningful in capturing exploitable inefficiencies in the yield curve.

Based on these observations, all subsequent analyses in this report utilize this parameter configuration: a rolling window of 4 months and 20 best-selected trades.

4.2 Signal Dynamics: Z-scores, Carry and PnL Evolution

To better understand the behavior and reliability of the trading signals over time, we visualize the historical evolution of both z-scores and carry across all butterfly structures. These two metrics form the foundation of our selection criteria and provide insights into the underlying market dynamics that drive relative value opportunities. To evaluate realized performance, we also track the PnL evolution of selected trades, providing empirical validation of the strategy's effectiveness under different market conditions.

4.2.1 Z-scores Evolution

Figure 5 shows the time series of z-scores for all six butterfly structures across the full sample period. Several patterns are evident from the visual inspection of z-score trajectories. Most of the time, z-scores fluctuate within the ± 2 range, which serves as a common threshold for signal activation. However, repeated breaches of this range are observed, suggesting potential curve dislocations or abrupt shifts in curvature. Notably, these extreme deviations are concentrated around mid-2022 and again in the second half of 2024.

These two periods likely correspond to distinct phases of macroeconomic stress and policy realignment. The mid-2022 spike in z-score volatility aligns with a period of aggressive monetary tightening by the Bank of England in response to persistent inflation. This phase saw a series of rapid policy rate hikes, which induced significant re-pricing along the SONIA swap curve—a key reference in sterling fixed-income markets.

In addition to policy shifts, such extreme dislocations in z-scores may also reflect broader systemic stress. During episodes of sharp risk-off shocks, market participants often flee to the safest ends of the curve (ultra-short and long maturities), while intermediate segments become illiquid. This "liquidity barbell" behavior leads to extreme local distortions in the curve, which are identifiable by Micro RV signals.

Similarly, the resurgence of z-score volatility in late 2024 coincides with a shift in the macro-policy narrative. The Bank of England began to cut interest rates in response to signs of economic slowdown, following a prolonged period of restrictive monetary policy. Simultaneously, the UK government unveiled an expansive fiscal budget, triggering market concerns around inflation and debt sustainability. These domestic developments, combined with heightened global uncertainty—such as geopolitical tensions and policy shifts surrounding the U.S. election—likely contributed to a re-pricing of expectations along the sterling yield curve. As a result, elevated z-score activity in this period may reflect market participants adjusting to a more volatile and less predictable rate environment.

4.2.2 Carry Evolution

Figure 6 displays the carry profiles for the same structures, with selected trades highlighted in blue. Firstly, the carry profile is highly structure-dependent. For instance, structures such as $(2y, 5y, 10y)$ and $(10y, 20y, 30y)$ generally exhibit persistently positive carry throughout the sample, while others like $(5y, 10y, 30y)$ tend to remain negative. Also, carry values are relatively stable for some structures (e.g., $(1y, 3y, 5y)$), whereas others (e.g., $(5y, 10y, 30y)$) show higher volatility, occasionally experiencing sharp drops or spikes. This volatility likely reflects shifting curve dynamics and changing expectations about forward rates.

Lastly, the selected trades (highlighted as blue dots) tend to align with periods where carry is not only positive but also elevated relative to the structure's own history. This confirms that the carry filter effectively prioritizes trades that are not only statistically significant but also supported by favorable static PnL expectations.

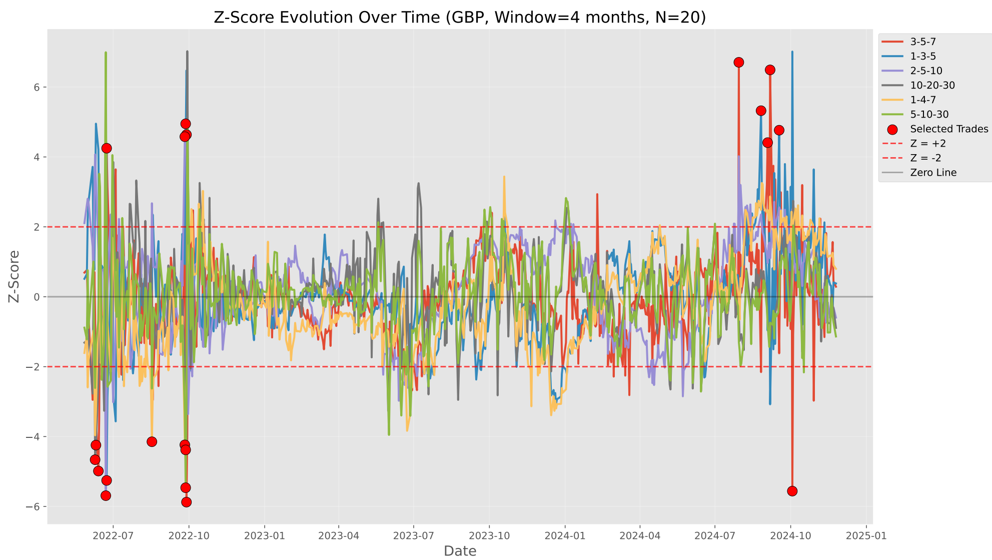


Figure 5: Z-score Evolution Over Time (GBP, Window=4 months, N=20)

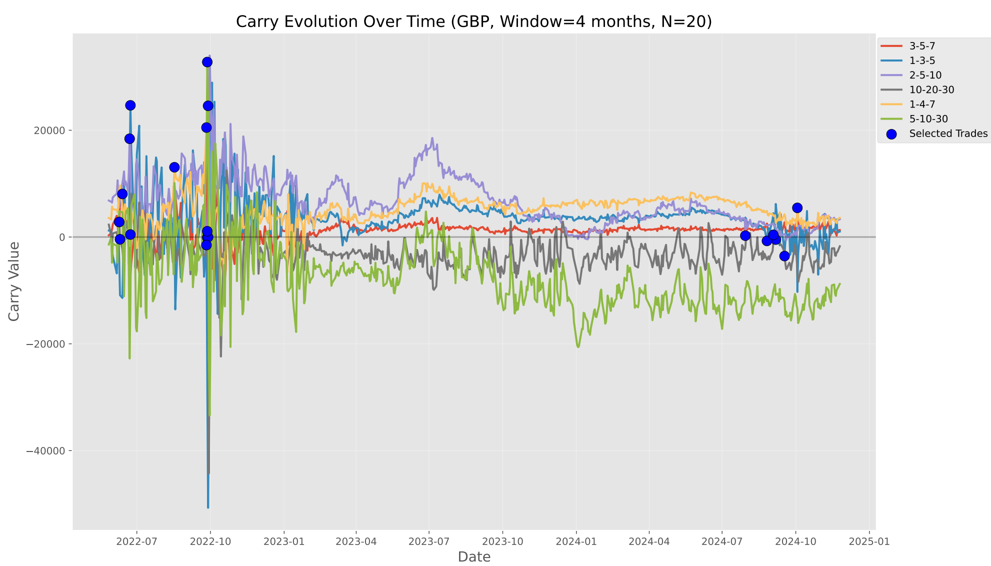


Figure 6: Carry Evolution Over Time (GBP, Window=4 months, N=20)

4.2.3 PnL Evolution of Top Trades

After selecting trades based on a combination of z-scores extremity and favorable carry, we simulate the performance of each trade using a DV01-neutral approach. This method ensures that the exposure across legs is properly balanced in terms of interest rate sensitivity, and that profits and losses are primarily driven by curvature changes rather than directional shifts in the level of rates.

Figure 7 illustrates the cumulative performance trajectories of all selected trades across the sample period. Each line represents a single trade, initiated at the time of signal generation, and tracked until the maturity of the longest leg. The direction of the trade—long or short fly—is determined by the sign of the z-scores at entry, with positive z-scores corresponding to short curvature positions and negative z-scores to long curvature positions.

We observe substantial variation in trade performance. A subset of trades generated steady, positive PnL with low drawdowns, while others exhibited high volatility or ended in loss. This dispersion highlights the importance of filtering trades not only by statistical significance but also by forward-looking economic attractiveness, as captured by the carry signal. Trades initiated in early- to mid-2022 achieved strong profits, likely benefiting from rapid shifts in the shape of the curve driven by central bank tightening cycles.

Overall, the PnL evolution analysis provides an empirical view into how the strategy performs when applied in a real-time, forward-looking manner. The ability of certain structures to generate consistent profits validates the predictive power of our combined signal framework. It also highlights the importance of selecting the right structure and entry timing to maximize the effectiveness of Micro RV strategies.

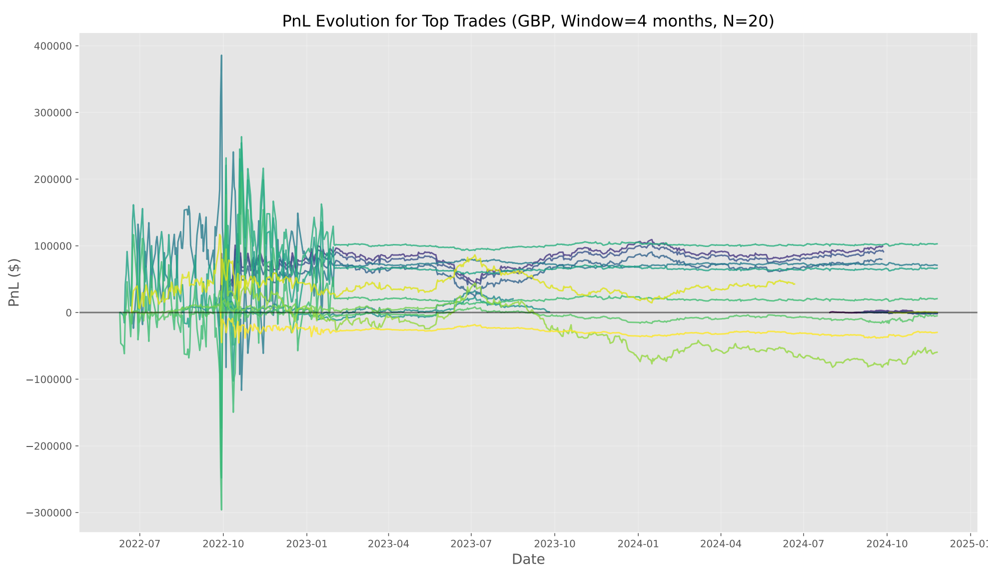
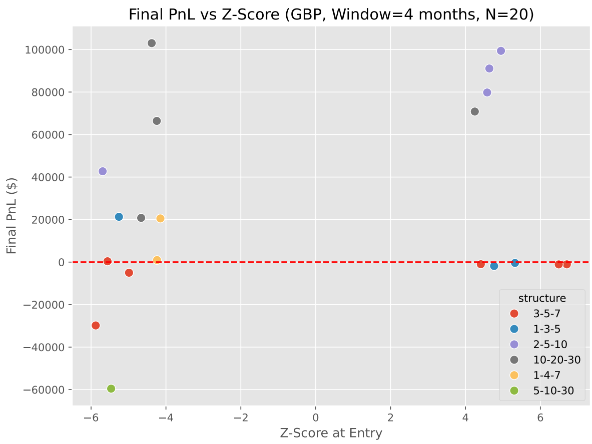


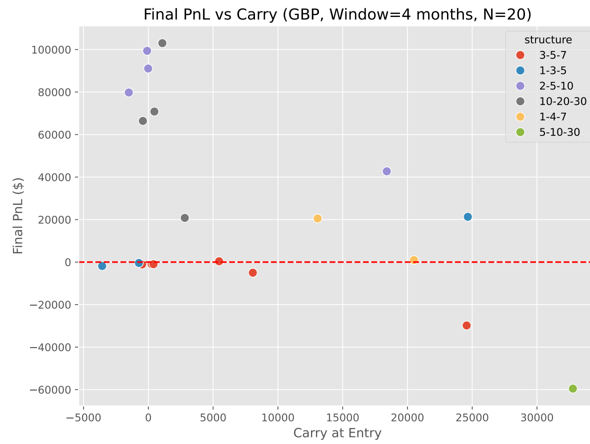
Figure 7: PnL Evolution for Top Trades (GBP, Window=4 months, N=20)

4.3 Relationship Between Final PnL and Entry Signals

Figure 8 illustrates the relationship between final PnL and two key entry signals—Z-scores and carry—for the top 20 selected trades. The left panel plots final PnL against the Z-scores at trade



Final PnL vs Z-score



Final PnL vs Carry

Figure 8: Final PnL vs Z-Score and Carry (GBP, Window=4 months, N=20)

entry. It is important to note that Z-scores is used primarily in terms of its absolute value, as it measures the degree of statistical dislocation relative to historical norms, regardless of direction. High absolute Z-scores suggest extreme deviations and are used as a trigger for trade entry.

As shown, since there trades are with large-magnitude Z-scores, they tend to cluster toward positive PnLs, confirming that more extreme statistical signals are generally more reliable indicators of mispricing. However, there is still variation across structures, with some PnLs near 0 or even negative.

The right panel plots final PnL against carry at entry. While no strong functional relationship is observed, this aligns with our methodology in which carry is not intended to act as a predictive variable but rather as a filtering criterion. Specifically, we include carry to exclude trades with low or negative expected static returns. By retaining only the top 80% of trades by carry, we ensure that selected opportunities are not only statistically extreme but also economically viable.

4.4 PnL Result Analysis

This section presents the performance analysis of the trading strategy over a 4-month rolling window with a sample size of 20 best trades.

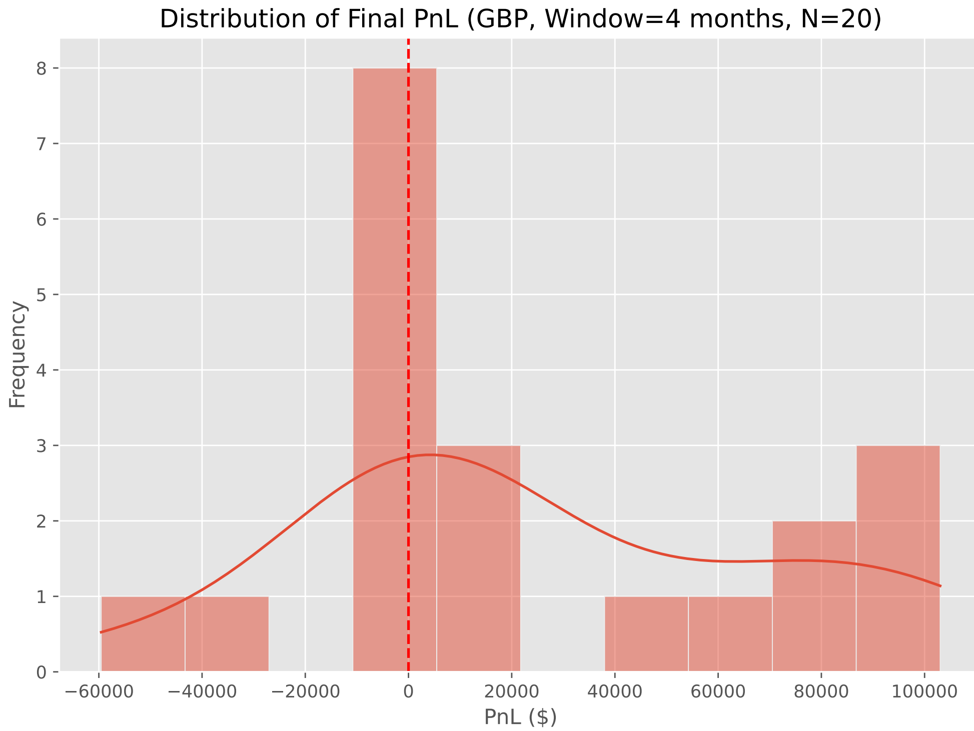


Figure 9: Distribution of Final PnL (GBP, Window=4 months, N=20)

Table 1: Trade Performance Summary (GBP, Window=4 months, N=20)

Metric	**Value**
Number of Trades	20
Win Rate	60.00%
Average PnL	25,875.82
Average Win[a]	51,437.79
Average Loss[b]	-12,467.14
Max PnL	103,016.71
Min PnL	-59,544.26

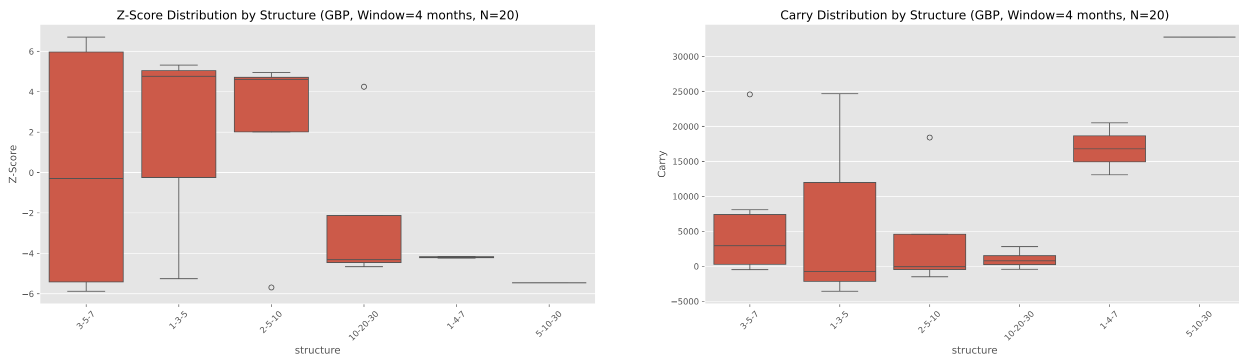
[a] Average of all positive-return (winning) trades. Same for Table 2
[b] Average of all negative-return (losing) trades. Same for Table 2

According to Figure 9, the distribution of final PnL across all trades reveals a positively skewed shape, with the majority of trades clustered around modest gains and a few outliers generating significantly higher returns. This asymmetry suggests that while the strategy may not win consistently, it captures occasional large opportunities that compensate for smaller losses. The presence of a meaningful right tail also supports the use of z-scores and carry as filters for extreme but profitable dislocations.

As shown in Table 1, the strategy demonstrates solid profitability with an average PnL of $25,875.82$ per trade and a win rate of 60.00%, indicating a reasonable edge in directional forecasting.

Notably, the average profit per winning trade (51,437.79) is significantly higher than the average loss (12,467.14), reflecting a favorable asymmetry in the payoff structure. Overall, the strategy appears structurally sound, with an attractive risk-reward profile. Nonetheless, the presence of volatility in outcomes and occasional large losses underscores the importance of effective position sizing and risk management.

4.5 Structure-Level Effects on Signal Quality and PnL



Z-Score Distribution by Structure   Carry Distribution by Structure

Figure 10: Z-Score and Carry Distribution by Structure (GBP, Window=4 months, N=20)

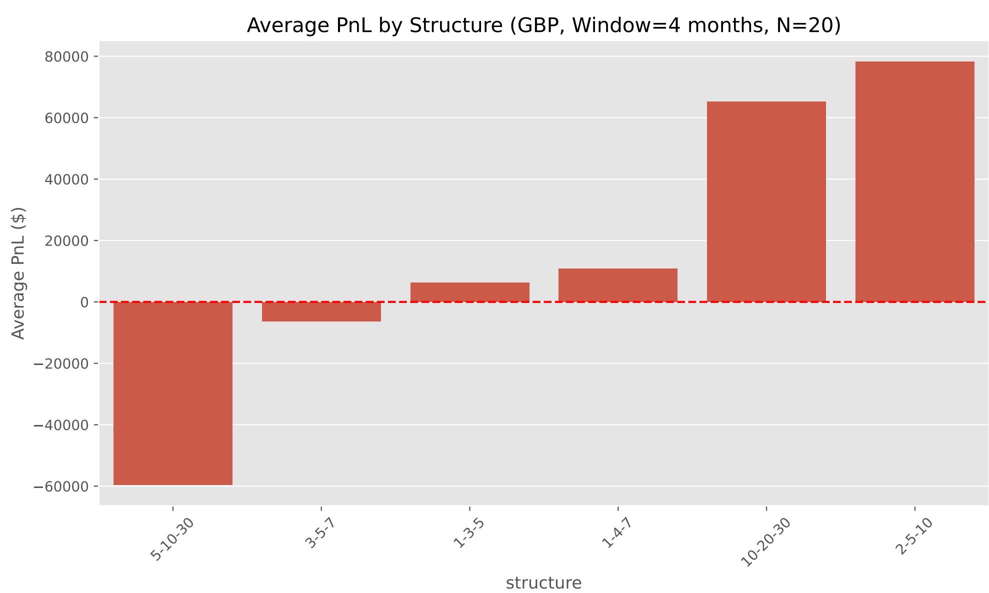


Figure 11: Average PnL by Structure (GBP, Window=4 months, N=20)

Figure 10 presents the distribution of z-scores and carry values across different butterfly structures, while Figure 11 reports the corresponding average final PnL per structure. These plots reveal that signal characteristics and profitability can vary meaningfully across butterfly configurations.

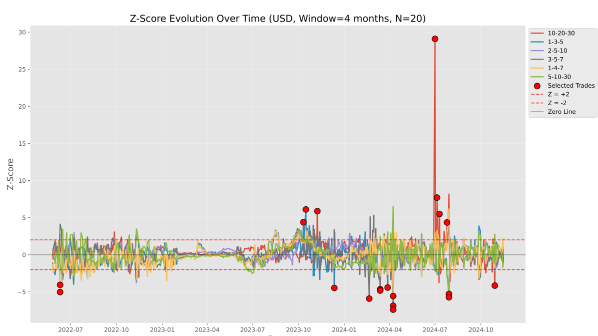
One notable example is the $(5y, 10y, 30y)$ structure, which delivers a significantly negative average PnL. This result stems from a single trade in the top 20 that generated an outsized loss, suggesting that the statistical extremeness in this configuration may not be aligned with economically viable dislocations. The combination of long-dated wings and convexity exposure could contribute to greater PnL volatility and may lead to a huge loss.

In contrast, the $(2y, 5y, 10y)$ and $(10y, 20y, 30y)$ structures demonstrate consistently high profitability, supported by tight z-scores clustering and relatively attractive carry distributions. Their performance suggests that dislocations in these segments of the curve—typically involving more liquid and well-arbitraged tenors—may be more reliably captured and monetized using a Micro RV framework. These structures appear to align better with curve dynamics and carry mechanisms, leading to more robust outcomes.

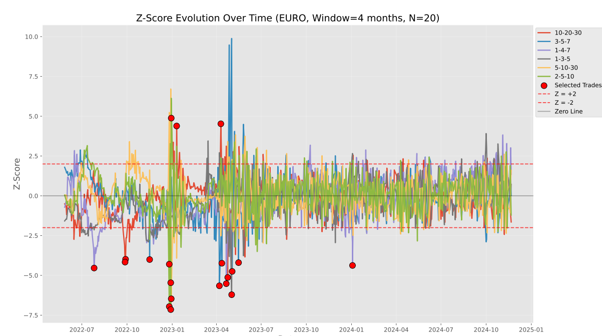
Shorter-dated structures such as $(1y, 3y, 5y)$ and $(1y, 4y, 7y)$ show moderate profitability and narrower dispersion in both signal dimensions, indicating stable but less pronounced opportunities. Meanwhile, the $(3y, 5y, 7y)$ structure shows negative average PnL, reinforcing that not all butterfly configurations are equally effective and that structural design materially impacts strategy effectiveness.

4.6 Other currencies

We apply the same Micro RV methodology to both USD and EUR swap curves, using a four-month rolling window and selecting the top $N = 20$ trades based on Z-scores and carry filtering. This allows for a cross-currency comparison of strategy robustness and effectiveness.



Z-score Evolution Over Time for USD

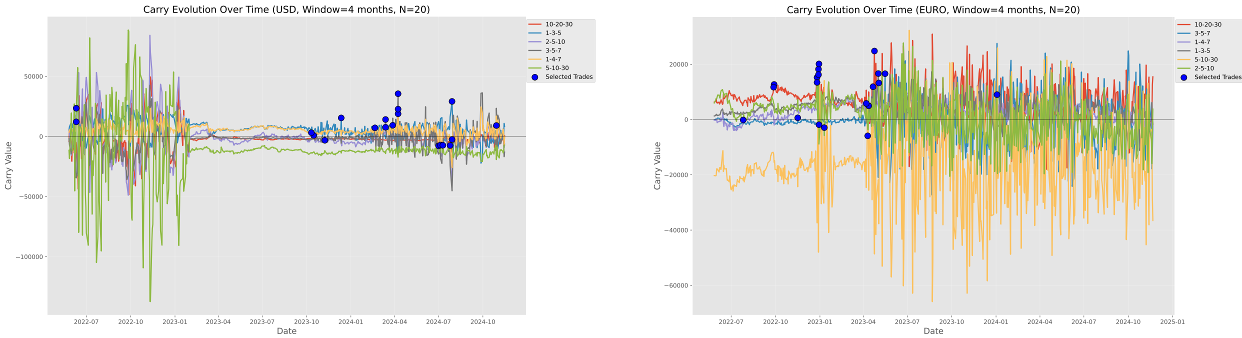


Z-score Evolution Over Time for EURO

Figure 12: Z-Score Evolution Over Time for USD and EURO (Window=4 months, N=20)

Figures 12 and 13 present the time-series evolution of Z-scores and carry for USD and EUR, respectively. While both currencies exhibit variation over time, the EUR curve shows considerably more erratic patterns, with frequent and pronounced spikes in both indicators. This suggests that signals derived from the EUR curve are less stable, potentially reflecting noisier curve dynamics, uneven maturity coverage, or episodic distortions in specific tenors.

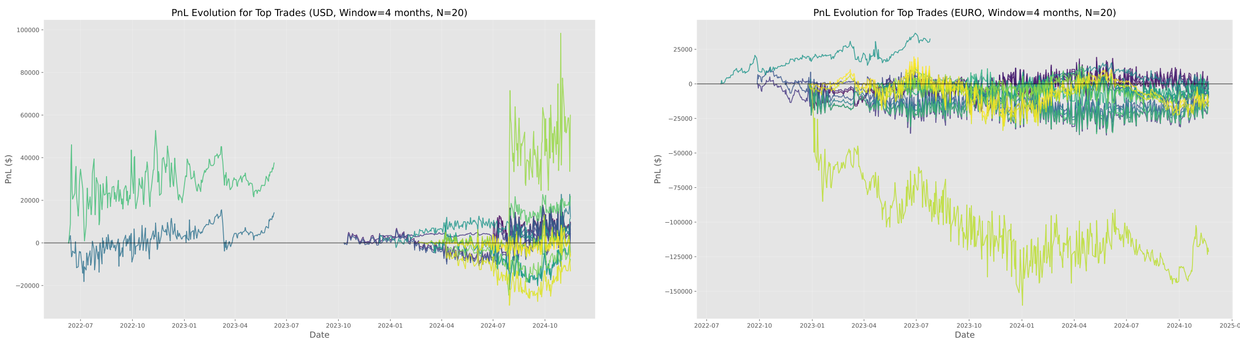
In particular, the EUR signals display sharp transitions in both Z-score and carry metrics, concentrated in certain periods and curve segments. These dynamics may indicate a higher sensitivity to localized structural effects or transient market conditions, as opposed to more systematic deviations. In contrast, the USD signals evolve more gradually, and Z-score clusters tend to persist over time, allowing for more consistent signal extraction and trade identification.



Carry Evolution Over Time for USD

Carry Evolution Over Time for EURO

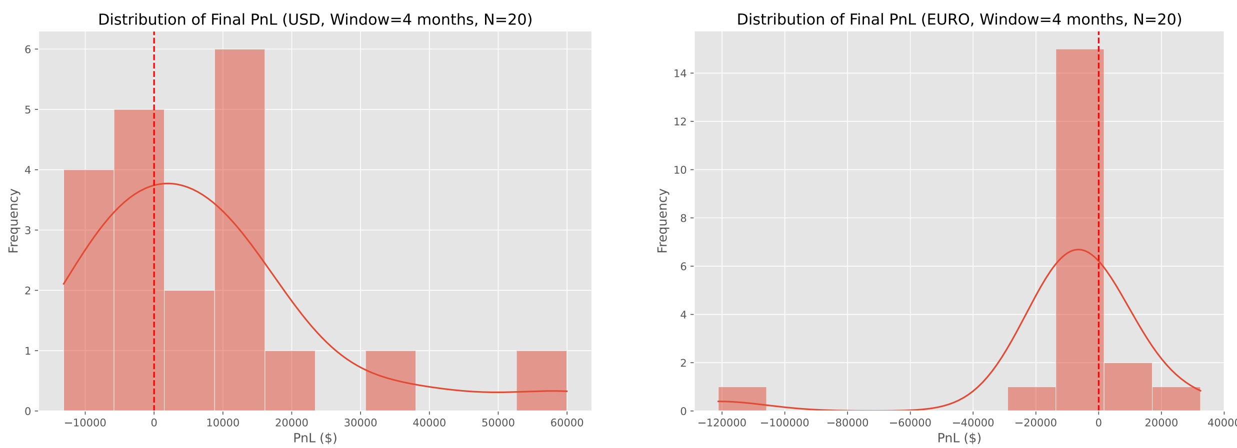
Figure 13: Carry Evolution Over Time for USD and EURO (Window=4 months, N=20)



PnL Evolution Over Time for USD

PnL Evolution Over Time for EURO

Figure 14: PnL Evolution Over Time for USD and EURO (Window=4 months, N=20)



Distribution of Final PnL for USD

Distribution of Final PnL for EURO

Figure 15: Distribution of Final PnL for USD and EURO (Window=4 months, N=20)

Figures 14 and 15 show the cumulative PnL evolution and final outcome distributions for the top-ranked trades. In the USD market, PnL paths are relatively smooth and exhibit a mild right-skew, with the majority of trades generating positive returns. The outcome distribution shows a compact range, with no extreme losses and moderate positive outliers. In contrast, the EUR performance is dominated by one exceptionally poor trade, which resulted in a loss of over 120,000. While several other EUR trades cluster near breakeven, this single extreme outcome significantly skews the average return downward, as reflected in the summary statistics. This asymmetry highlights the impact of tail events on aggregate performance and introduces greater variability into the trade distribution.

Table 2: Comparison of Trade Performance between USD and EURO (Window=4 months, N=20)

Metric	USD	EURO
Win Rate	60.00%	20.00%
Average PnL	7,331.90	-10,105.59
Average Win	16,421.42	11,248.06
Average Loss	-6,302.38	-15,444.00
Max PnL	59,962.71	32,451.73
Min PnL	-13,129.41	-121,182.43

Table 2 summarizes the performance metrics. For USD, the strategy achieves a 60% win rate and an average PnL of $7,331.90$, with a relatively well-behaved distribution of wins and losses. In the EUR case, however, the win rate drops to 20% and the average PnL is $-10,105.59$. The presence of an extreme negative outlier also increases the average loss per losing trade to -$15,444.00$, compared to -$6,302.38$ in USD.

This contrast suggests that while the Micro RV framework performs well in certain markets like USD, it may not generalize equally across all currency environments. The EUR curve, in particular, appears less suited to this approach, potentially due to market segmentation, differences in tenor liquidity, or reduced efficiency in arbitrage across the term structure. Additionally, European rates may be more sensitive to discrete policy shocks and flow imbalances, making statistically extreme observations less reliable as trading signals.

5 Discussions and Conclusions

While the numerical results validate the effectiveness of Micro RV strategies under historical data and theoretical signals, their real-world implementation is subject to significant practical frictions. These include liquidity constraints, execution costs, position sizing limitations, and the necessity of robust risk management.

Micro RV strategies inherently rely on the assumption that the identified curve dislocations are tradable. However, in practice, curve segments—particularly those involving intermediate maturities or off-the-run swap tenors—may suffer from low market depth. For instance, if a 7-year point is interpolated due to infrequent trades, the apparent dislocation in a (5y, 7y, 10y) butterfly may not be actionable. The observed z-score might reflect a stale or noisy quote rather than true mispricing. Strategies should therefore implement liquidity filters—for example, limiting signal generation to actively traded maturities or confirming dislocations via dealer quotes or bid-ask spreads.

The strategy assumes frictionless execution at daily closing mid prices. In reality, each leg of a butterfly structure incurs bid-ask costs and brokerage fees. A typical 3-leg trade might involve a round-trip cost of 2–4 basis points, which could consume a substantial portion of the expected return. High-turnover variants of Micro RV strategies may thus suffer from negative net profitability after costs unless signals are sufficiently strong or trades are executed with favorable timing and size. This underscores the importance of signal robustness and execution-aware design.

Because Micro RV opportunities are usually narrow in scale and short-lived, they impose natural capacity constraints. Attempting to allocate large notional size to a small curve dislocation can lead to market impact and slippage, or cause the signal to decay before full execution. Moreover, if many participants chase similar trades (e.g., the same butterfly structures in liquid currencies), dislocations may self-correct quickly, leaving little room for profit. So, this strategy might work best in a low-AUM, nimble setup, with well-chosen trades and tight position sizing.

Even though butterfly structures are constructed to be DV01-neutral, they remain exposed to curve convexity changes, idiosyncratic flow shocks, and liquidity squeezes. During episodes of macro instability or crisis, curve distortions may deepen rather than revert, leading to persistent unrealized losses.To mitigate these risks, real-world implementation should incorporate several safeguards. These include applying stop-loss thresholds to individual trades to cap downside exposure, and limiting the capital allocated to each trade or maturity bucket to avoid concentration risk. Additionally, performing portfolio-level stress tests based on hypothetical curve shifts or volatility shocks can help anticipate potential drawdowns. Finally, ensuring diversification across different currencies and segments of the yield curve provides a further layer of protection, helping reduce the impact of localized dislocations.

This project demonstrates the effectiveness of a Micro Relative Value (Micro RV) framework in identifying and monetizing curvature-based dislocations in interest rate swap markets. Through systematic signal construction using Z-score and carry, combined with DV01-neutral butterfly positioning, the strategy shows robust performance under the GBP curves, with consistent profitability across reasonable parameter configurations.

In the GBP market, Micro RV signals exhibit stable statistical properties and meaningful economic backing. Structures such as $(2y, 5y, 10y)$ and $(10y, 20y, 30y)$ deliver particularly strong returns, benefiting from favorable carry dynamics and liquidity. This can be easily extended to other currencies. In the USD curve, the framework also performs well, with a 60% win rate and positive average PnL, driven by a more symmetric distribution of trade outcomes and manageable downside risk.

However, when applied to the EUR curve, the strategy underperforms. The EUR results are skewed by a single large loss, which reflects the greater risk of structural fragility or misaligned signals in that market. This suggests that while Micro RV is a powerful tool, its success may not translate directly across all currencies. Differences in liquidity, market depth, monetary policy regimes, and curve behavior all impact signal reliability.

Overall, Micro RV is a promising framework for capturing second-order term structure inefficiencies, but its application should be paired with thoughtful structural design, currency-specific calibration, and risk-aware implementation.

References

[1] Vladimir V. Piterbarg. (2025). Lecture Notes on Topics in Derivatives Pricing. Imperial College London, Department of Mathematics.

[2] Steeley, J. M. (1990). Modelling the dynamics of the term structure of interest rates. Economic and Social Review, 21(4), 337-361. http://www.tara.tcd.ie/handle/2262/64617

[3] Litterman, R.B. and Scheinkman, J. (1991) Common Factors Affecting Bond Returns. Journal of Fixed Income, 1, 54-61. http://dx.doi.org/10.3905/jfi.1991.692347

[4] Martellini, L., Priaulet, P. and Priaulet, S. (2003). Fixed-Income Securities: Valuation, Risk Management and Portfolio Strategies, John Wiley and Sons.