Functional Extreme Partial Least Squares: Unraveling the Intuition and Empirical Validation

Janis Aiad[1] Simon Elis[1]
[1]Master MVA - Statistical Learning with Extreme Values
ENS Paris Saclay, Paris, France
github.com/janisaiad/fctpls

Abstract

We present an in-depth study of the Functional Extreme Partial Least Squares (FEPLS) framework, a method designed to identify predictive features in high-dimensional or functional data associated with rare, extreme events. Our work applies the FEPLS approach to financial time series, with a particular focus on understanding how patterns in one asset's intraday returns anticipate large, infrequent moves in another. We empirically evaluate FEPLS in two settings: (1) cross-asset analysis using medium-frequency stock data, and (2) high-frequency financial data via a subsampling methodology. The results provide new insights into the strengths and limitations of FEPLS, practical guidelines for its calibration, and its applicability across a range of financial regimes.

1 Introduction

Extreme events—rare, significant deviations from normal behavior—play a crucial role in finance, signal processing, risk management, and many scientific fields. The way such extremes manifest in data, especially when those data are high-dimensional or functional (such as entire intraday price curves), presents substantial challenges both for modeling and for prediction. Conventional dimension reduction tools like Principal Component Analysis (PCA) or classical Partial Least Squares (PLS) are effective for understanding average behavior but may fail to capture the patterns associated specifically with rare, high-impact events.

Functional Extreme Partial Least Squares (FEPLS) is a methodology developed to address precisely this challenge. It extends the classical PLS approach into the domain of extremes, targeting the features of high-dimensional or infinite-dimensional covariates that are most informative for explaining or predicting the occurrence of extreme events in a response variable.

The central research question motivating this work is the following:

What is the most likely shape of the covariate X when the response Y is extreme?

Answering this question, for instance, can reveal which specific features of a price curve tend to occur in the lead-up to major market moves—a task of paramount importance in financial engineering, but also relevant in fields like signal processing and detection theory, where characterization of rare patterns is required.

In this report, we offer both a theoretical and comprehensive empirical study of the FEPLS framework. We begin by unpacking the intuition and mathematical underpinnings of FEPLS, clarifying how it focuses on conditional extremes, and describing the signal-to-noise assumptions required for consistent estimation. Our goal is to provide a clear narrative that connects the theory to practical considerations.

To bridge theory and practice, we conduct extensive experiments on open financial datasets. These analyses illuminate the strengths and limitations of FEPLS in realistic, noisy, heavy-tailed environments. We also move beyond the day-to-day context of classical financial studies, exploring high-frequency data using a subsampling-based methodology. This allows us to probe FEPLS performance and the validity of its foundational assumptions across a variety of time scales and predictive settings.

Our contributions in this work include:

- A concise evaluation of FEPLS's statistical assumptions and estimator stability in real-world data, along with practical

guidance on parameter calibration and computational considerations, culminating in the first empirical analysis of FEPLS applied to high-frequency financial data across milliseconds to minutes.

- A detailed empirical evaluation of FEPLS under challenging market noise, providing clear illustrations of when and where the method succeeds or faces limitations.

Through this combination of theoretical insight, methodological clarification, and practical experimentation, we aim to both advance the understanding of FEPLS and provide a resource for practitioners seeking to apply it in domains where extremes govern the risk landscape.

2 Intuitive Understanding of the FEPLS Framework

2.1 Difference Between PLS and PCA

The fundamental distinction between Partial Least Squares (PLS) and Principal Component Analysis (PCA) lies in their use of supervision. See Table 1 in the appendix for a detailed comparison.

FEPLS (PLS for extremes): FEPLS extends this supervised approach to the extreme regime. The fundamental idea is to find a **single direction** in the infinite-dimensional space of functional covariates that best explains extreme events. Think of it as finding the "recipe" for extreme market movements: if you could only look at one linear combination of all the intraday price movements, which combination would best predict tomorrow's crash?

2.2 The Mathematical Intuition: From Covariance to Direction

The optimization problem (1) asks: "Which direction w maximizes the covariance between the projection $\langle w, X\rangle$ and Y, *when we only look at extreme values of* Y?"

This is different from standard PLS because:

- Standard PLS: Maximizes covariance over all data points (average behavior).
- FEPLS: Maximizes covariance *conditionally* on $Y \geq y$ (tail behavior).

The key insight is that by conditioning on extremes, we focus the optimization on the regime that matters most for risk prediction.

3 Notation

3.1 Notation

A summary of the main notation used in the FEPLS framework is provided in Table 2 in the appendix.

4 Theoretical Framework

4.1 The FEPLS Problem

Suppose we observe data pairs (Y, X), where Y is a real-valued variable we care about (for example, the size of a market move), and X contains other observable information (for example, a time series or a function we record alongside Y). The central question is: **Given X and Y, can we find a summary of X that best explains the most extreme values of Y?**

FEPLS answers this by searching for a direction w in the space of possible summaries of X that is most strongly related to the largest values of Y. Formally, for a high threshold y, we solve:

$$w(y) := \underset{\|w\|=1}{\arg\max}\, \mathrm{Cov}\big(\langle w, X\rangle, Y \mid Y \geq y\big), \quad y \in \mathbb{R}, \tag{1}$$

where $\langle w, X\rangle$ simply means forming a linear combination of X using w (for instance, by integrating $w(t)X(t)$ if X is a function of t), and the Cov is computed only among the cases where Y exceeds y—that is, focusing only on the extremes.

4.2 Inverse Model: g (Model) vs. φ (Method)

The paper **suppose** the inverse model

$$X = g(Y)\beta + \varepsilon, \quad \beta \in H, \quad \|\beta\| = 1, \tag{2}$$

where $g \in \mathrm{RV}_\kappa(+\infty)$ ($\kappa > 0$) is an unknown link function and $\varepsilon : \Omega \to H$ is noise. When Y is extreme, $g(Y)$ tends to dominate X, making the signal $g(Y)\beta$ far larger than the noise.

What's fixed and what's chosen?

- g **(κ):** Describes the underlying, real-world relationship between Y and X (imposed by the data-generating process; not under our control). Example: $g(y) = y^{0.5}$.
- φ **(τ):** A user-chosen, "test" function (tuning parameter); not part of nature, but introduced in the estimator $\hat{\beta}_\varphi$ to optimize statistical performance for extremes.

Moment Condition and Tuning of τ

The first key condition, $0 < 2(\kappa+\tau)\gamma < 1$, ensures moments exist for reliable estimation. If g (large κ) grows too fast, sums like $\sum X_i$ may diverge; picking a decreasing φ (negative τ) counteracts this, keeping sums finite.
The estimator uses weights $\varphi(Y_i)$ for extremes, e.g.

$$\sum X_i \varphi(Y_i) \mathbf{1}_{\{Y_i \geq y\}}$$

so tuning τ adjusts influence: big τ (positive) focuses on the deepest extremes (low bias, high variance); small or negative τ spreads weight (better stability).

4.3 Second-Order Regular Variation

A new assumption is that the response variable Y is heavy-tailed to the second order. Specifically, the tail quantile function $U(t) := F^-(1-1/t)$ belongs to the class of second-order regularly-varying functions:

Definition 4.1 (Second-Order Regular Variation)**.** The function U belongs to $2\mathrm{RV}_{\gamma,\rho}(+\infty)$ if there exist $\gamma \in (0,1)$, $\rho \leq 0$ and an auxiliary function A ultimately of constant sign with $A(t) \to 0$ as $t \to +\infty$ such that:

$$\lim_{t\to+\infty} \frac{1}{A(t)} \left(\frac{U(ty)}{U(t)} - y^\gamma \right) = y^\gamma H_\rho(y) := y^\gamma \int_1^y u^{\rho-1} \tag{3}$$

A variation. Although the original paper does not state it explicitly, this definition immediately yields that A is of the form $A(t) = t^\rho$ which is important to derive the convergence scaling law in (7).

Interpretation of ρ 2nd-order RV

Why is Second-Order Regular Variation (2RV) necessary?
The first-order regular variation (RV_γ) tells us that the distribution "resembles" a Pareto distribution asymptotically, but in finite samples, we are never truly "at infinity." The distribution is:

$$\text{True Distribution} = \text{Pareto} + \text{Error}$$

The 2RV quantifies this error through the auxiliary function $A(t) = t^\rho$. This becomes crucial when choosing k:

- If k is very small (deep in the tail), the approximation error is small (close to the limit), but the variance is huge (few points).
- If k is larger (to stabilize variance), we move away from the extreme tail where the distribution deviates from pure Pareto. The approximation error grows.

The 2RV tells us *how fast* this error grows. The condition $\sqrt{k}A(n/k) = O(1)$ balances:

- $\sqrt{k}$: The statistical variance (noise)
- $A(n/k)$: The model bias (error quantified by 2RV)

This condition says: "You may increase k as long as your model error (A) remains smaller than your statistical noise ($\sqrt{k}$)."
In summary: 1RV (γ) gives the *direction* of the tail (the slope), while 2RV (ρ) gives the *straightness* of the tail (is it a perfect straight line in log-log scale, or is it curved?). If ρ is very negative, the tail is almost straight, making estimation easier. If ρ is close to 0, convergence is very slow.

4.4 Signal Dominance over Noise

Signal Dominance over Noise

The tail of $g(Y)$ (signal channel) must be heavier than the tail of the noise ε. In practical terms, this ensures that among the largest observed Y values, the associated signal $g(Y)\beta$ is much more extreme than the noise, so the extreme behavior is informative about β.
Why necessary? If the noise's tail is too fat, its fluctuations swamp the regression signal in the extremes, making recovery impossible even with infinite data. *If κ increases (sharper g growth), you can choose larger k and variance decreases.*
That is why the model is only identifiable if the signal is dominant over the noise, under the inequality

$$q\kappa\gamma > 1. \tag{4}$$

We call this condition **signal dominance** for **identifiability**.

4.5 What do we estimate to ensure consistency ?

To ensure that we are into the consistency inequality taht allows signal recovery, we need to estimate both γ and ρ. κ cannot be estimated from the data.

so there is no best τ choice a priori, we need to construct estimator for a lot of different τ values and find an interval that contains some stationarity in the beta estimate. We call this procedure τ **tuning** to satisfy the **integrability condition**.

4.5.1 Choice of k

When testing a particular τ, we need to choose a number of extreme observations k to minimize the bias-variance tradeoff. We call this procedure **k tuning** for **convergence**.

Optimal Choice of k

Following the classical extreme-value theory framework (see [5], Equation 3.2.10), the optimal choice of k balances bias (governed by 2RV, i.e., $A(n/k)$) and variance (proportional to $k^{-1/2}$):

$$\text{error}(n,k) \approx C_1 A(n/k) + C_2 k^{-1/2}$$

where $A(n/k)$ controls bias and $k^{-1/2}$ controls the noise variance.
Solving for the optimal tradeoff yields:

$$k_n \sim c\, n^{-2\rho/(1-2\rho)}$$

for some $c > 0$ (constant) and $\rho < 0$.

Once you have chosen proper τ, estimated ρ and γ, you have chosen the number of observations to use, you have chosen k.

It remains that you have performed your FEPLS estimation, which convergence guarantee do you have ?

4.6 Threshold choice and growth of y

Growth of the Threshold and Intermediate Sequences

For any threshold $y \geq 0$, the FEPLS estimator is defined as

$$\hat{\beta}_\varphi(y) := \frac{\hat{v}_\varphi(y)}{\|\hat{v}_\varphi(y)\|}, \qquad \hat{v}_\varphi(y) = \frac{1}{n}\sum_{i=1}^{n} X_i\, \varphi(Y_i)\, \mathbf{1}_{\{Y_i \geq y\}}.$$

Thus, $\hat{\beta}_\varphi(y)$ is a weighted combination of the X_i whose Y_i are in the tail, with weights set by the test function φ.
For consistency results, we consider a sequence of (deterministic) thresholds $(y_{n,k})$ such that

$$y_{n,k} \sim U\left(\frac{n}{k}\right), \qquad n \to \infty,$$

where U is the tail quantile function of Y. In practice, $y_{n,k}$ is approximated by the $(n-k+1)$-th order statistic $Y_{n-k+1,n}$:

$$Y_{n-k+1,n} \approx y_{n,k} \sim U\left(\frac{n}{k}\right) \quad \text{as } n \to \infty.$$

If the tail of Y is regularly varying, i.e., $\bar{F}(y) = \mathbb{P}(Y > y) \in \mathrm{RV}_{-1/\gamma}(+\infty)$, then

$$U(t) \sim C t^{\gamma} \quad (t \to \infty)$$

for some $C > 0$; for example, in the standard Pareto, $U(t) \sim t^{\gamma}$. For $t = n/k$,

$$y_{n,k} \sim C \left(\frac{n}{k}\right)^{\gamma}.$$

We consider intermediate sequences $k = k_n \to \infty$ with $k_n/n \to 0$ so $n/k_n \to \infty$, hence

$$y_{n,k_n} \sim C \left(\frac{n}{k_n}\right)^{\gamma} \to +\infty \quad \text{as } n \to \infty.$$

So, the threshold y grows with n and k_n:

$$y = Y_{n-k+1,n} \approx U\left(\frac{n}{k}\right) \to +\infty$$

and more and more extreme observations are used as n increases.

4.7 Consistency Results

The FEPLS estimator is

$$\hat{\beta}_{\varphi}(y) := \frac{\hat{v}_{\varphi}(y)}{\|\hat{v}_{\varphi}(y)\|}, \quad \text{where} \quad \hat{v}_{\varphi}(y) = \frac{1}{n}\sum_{i=1}^{n} X_i \varphi(Y_i) \mathbf{1}_{\{Y_i \geq y\}}. \tag{5}$$

Consistency of FEPLS Estimator (Theorem ??)

Theorem 4.2 (Consistency of FEPLS Estimator)**.** *Under the previous assumptions, the FEPLS estimator is consistent:*

$$\|\hat{\beta}_{\varphi}(Y_{n-k+1,n}) - \beta\| = O_{\mathbb{P}}(\delta_{n,k}) \xrightarrow[n\to+\infty]{} 0, \tag{6}$$

where the convergence rate is

$$\delta_{n,k} := \left(g(y_{n,k})\left(\frac{k}{n}\right)^{1/q}\right)^{-1}$$

with g and q as defined in the model assumptions. When $k \sim c\, n^{-2\rho/(1-2\rho)}$ for $\rho < 0$ and $c > 0$, and the bias condition $\sqrt{k}A(n/k) = O(1)$ holds, the rate satisfies

$$\delta_{n,k} = O\left(n^{(1/q-\gamma\kappa)/(1-2\rho)}\right). \tag{7}$$

Consistency requires (i) signal dominance $q\kappa\gamma > 1$ (signal tail heavier than noise) and (ii) convergence stability $2(\kappa+\tau)\gamma < 1$ (finite variance).

The rate $\delta_{n,k}$ depends on the tail of Y (via γ and ρ), noise integrability q, and the link function g (via κ). See [1] for detailed proof and results.

From now we are all set to perform extensive empirical validation of the FEPLS estimator.

Summary of Main Assumptions and Scaling Laws

5 Empirical Validation

All our reproducible experiments are available in this repository.

5.1 Data Description

OHLC data: We use one year of mid-frequency (5-minute) open-high-low-close (OHLC) market data for Hungary and Poland, sourced from the open and widely-used financial data aggregator Stooq. The dataset covers both the Hungarian (Budapest Stock Exchange) and Polish (Warsaw Stock Exchange) stock indices.

Name	Condition / Regime	Role	Main Scaling / Consequence
Tail heaviness	$\bar{F}(y) \in \mathrm{RV}_{-1/\gamma}(+\infty)$, $\gamma \in (0,1)$	Heavy-tailed response Y (first-order RV)	Tail quantile $U(t) \sim Ct^{\gamma}$, threshold $y_{n,k} \sim C(n/k)^{\gamma}$
Second-order RV	$U \in 2\mathrm{RV}_{\gamma,\rho}(+\infty)$, $\rho \leq 0$, $A(t) = t^{\rho}$	Controls bias of Pareto approximation (curvature of the tail)	Bias term $A(n/k)$, choice of k via $\sqrt{k}A(n/k) = O(1)$
Moment / integrability	$0 < 2(\kappa+\tau)\gamma < 1$	Ensures existence of moments for FEPLS statistics; governs valid τ range	Admissible (κ, τ, γ) region for stable estimation
Signal dominance	$q\kappa\gamma > 1$	Signal tail heavier than noise; guarantees identifiability	FEPLS direction β recoverable from extremes; noise cannot dominate
Intermediate sequence	$k_n \to \infty$, $k_n/n \to 0$	Balances number of extremes and sample size	Threshold $y_{n,k_n} \sim C(n/k_n)^{\gamma} \to +\infty$; more and more extreme Y's used
Optimal k_n	$k_n \sim c\, n^{-2\rho/(1-2\rho)}$ for $\rho < 0$	Explicit bias-variance tradeoff using 2RV	Minimizes asymptotic error $\mathrm{error}(n,k)$; feeds into optimal convergence rate
Convergence rate	$\delta_{n,k} = \left(g(y_{n,k})(k/n)^{1/q}\right)^{-1}$	Speed of convergence of $\hat{\beta}_{\varphi}$ to β	For optimal k_n: $\delta_{n,k} = O\left(n^{(1/q-\gamma\kappa)/(1-2\rho)}\right)$

Table 1: Summary of the main assumptions, regimes, and resulting scaling laws underlying the FEPLS consistency and convergence analysis.

Data Characteristics:

- **Frequency:** 5-minute bars (OHLC)
- **Markets:** Hungary (`BUX`) 50 stocks, Poland (`WIG20`) 100 stocks
- **Size:** Approximately 50 megabytes compressed for 1 year (per market)
- **Coverage:** Full calendar year (about 50,000–60,000 bars per market)
- **Format:** CSV, standardized with columns for open, high, low, close, and volume
- **Source:** Freely available and reproducible from https://stooq.com/db/h/

The analysis uses X as the daily return curve from a stock A and Y as the next day's max return from a stock B. So we reproduce the analysis from the paper using open-source data.

Tick-by-tick data: For the most granular analysis, we purchased 3 months of nanosecond-resolution tick-by-tick data from Databento, covering all limit order book events (including every limit order, cancelation, and execution) for p rominent NASDAQ stocks: Google, Apple, American Airlines (AAL), Amazon, and Microsoft. This dataset, considered extremely high quality, cost approximately $1000 and represents a gold standard for empirical financial microstructure research.

- **Frequency:** Nanosecond-scale tick-by-tick (full limit order book)
- **Markets:** NASDAQ (Google, Apple, American Airlines (AAL), Amazon, Microsoft)
- **Duration:** 3 months
- **Events:** All limit order submissions, cancellations, executions (Level 3 order book data)
- **Size/Cost:** Approximately $1,000 for the full period via Databento
- **Quality:** Highest market microstructure fidelity available
- **Source:** Commercially obtained from Databento from previous work.

5.2 Statistical Workflow for Empirical Validation

We adhere to a transparent, reproducible workflow for FEPLS model validation using both Stooq (5-minute OHLC) and Databento (tick) data. The procedure is as follows:

1. **Data Preparation:** Download raw data (Stooq or Databento). Select target stocks and extract daily return curves X (covariate) and the scalar Y (next-day maximum return), possibly from different stocks to capture cross-dependence.
2. **Cleaning and Log Return Construction:** Compute 5-minute (or tick) log-returns, handle missing values (e.g., using linear interpolation or LOCF), and align data to uniform time grids.
3. **Dependence Check:** Assess temporal dependence (volatility clustering) in extremes via autocorrelation plots; optionally thin the data by taking every nth day to reduce dependence.
4. **Heavy-Tail Verification:** Plot the Hill estimator for Y to confirm a heavy-tailed regime appropriate for FEPLS. Proceed only if the estimated tail index γ is positive. If applicable for sufficient data, we estimate ρ as well.
5. **Correlation Check (Optional):** Compute canonical or ordinary correlation between projections $\langle X, \beta \rangle$ and Y to confirm a tractable signal for dimension reduction.
6. **Train/Test Chronological Split:** Split the data into training and testing periods. τ **tuning** is carried out on the training set.
7. **Parameter Calibration:**
   (a) Select a conservative test function parameter τ (e.g., $\tau = -1$).
   (b) Optimize the number of extremes k via a grid search: for each k, estimate $\hat{\beta}_{\tau,k}$ on the training set and maximize the in-extremes correlation between projections and Y.
8. **Validation:**
   - Apply the fitted $\hat{\beta}$ to the test set.
   - Project new curves X^{test} onto $\hat{\beta}$ to obtain scalar scores.
   - Estimate conditional Value-at-Risk (VaR) via quantile regression or local smoothing.
   - Assess visual coverage (exceedance) and independence of VaR violations.

This workflow ensures both statistical rigor and full reproducibility for all empirical FEPLS validation results.

6 Empirical Results

We present a comprehensive empirical validation of FEPLS on Hungarian stock market data. The analysis focuses on cross-asset relationships, examining how intraday return curves from one stock predict extreme returns in another. We analyze multiple stock pairs, with detailed results for the 4IG $\rightarrow$ AKKO pair shown in Figure 1. Additional results for 4IG paired with other stocks (Appeninn, Autowallis, BIF, CIG Pannonia, Dunahouse, GSPark, Richter, Alteo) are provided in the appendix.

6.1 Comments

Comments on the low frequency data: In contrast to the high-frequency setting, the low-frequency analysis (daily data) shows excellent Hill estimator performance, indicating proper heavy-tail behavior. However, we work with a much smaller sample size, with $k < 10$ extreme observations. In this regime of very noisy data with limited samples, we observe that the τ dependency becomes crucial: the choice of test function parameter τ significantly impacts the FEPLS estimates and their stability.

Our empirical findings reveal an important structural pattern: we can identify distinct stock groups based on their sensitivity to end-of-day volatility. Some stock pairs exhibit a clear relationship where end-of-day volatility drives the maximum return of the next day's stock movement, while others remain insensitive to this effect. This grouping provides valuable insights into cross-asset dependencies and volatility spillover mechanisms in financial markets.

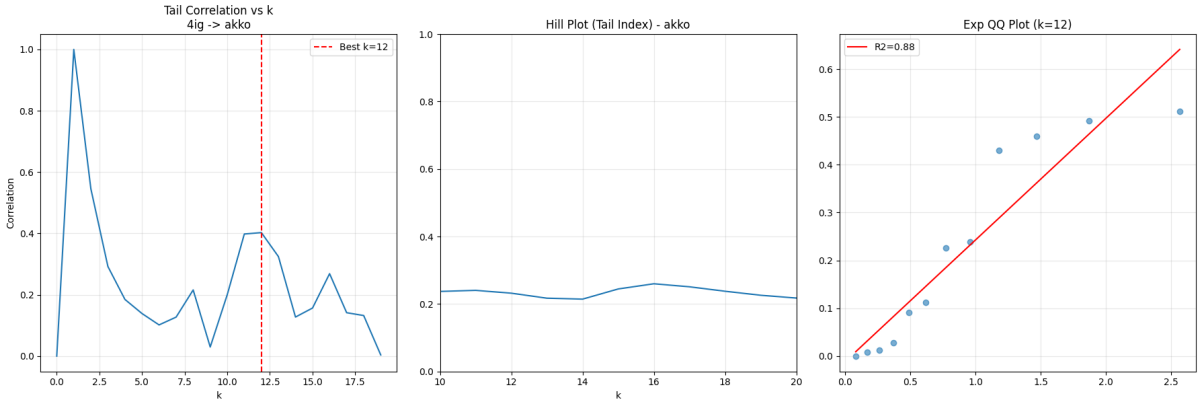


Figure 1: Empirical validation of FEPLS consistency: diagnostic plots for the 4IG → AKKO pair.

Comments on the high frequency data: We present results from high-frequency analysis using 5-minute intraday data for AAPL. In this setup, X represents the 5-minute window of returns, and Y is the maximum return of the next 5-minute window for the same stock. With k ranging from 0 to 800, we have access to a large amount of data.

However, the high-frequency nature of the data (5-minute intervals) presents challenges: the Hill estimator is not well estimated due to the tick-by-tick microstructure effects, and the Q-Q plot shows poor fit. Despite these limitations, we observe a clear increasing relationship between the FEPLS projections and extreme responses, as shown in Figure 2. This increasing curve provides strong evidence that mid-frequency price movements are primarily trend-driven, where current 5-minute patterns predict subsequent extreme moves in the same direction.

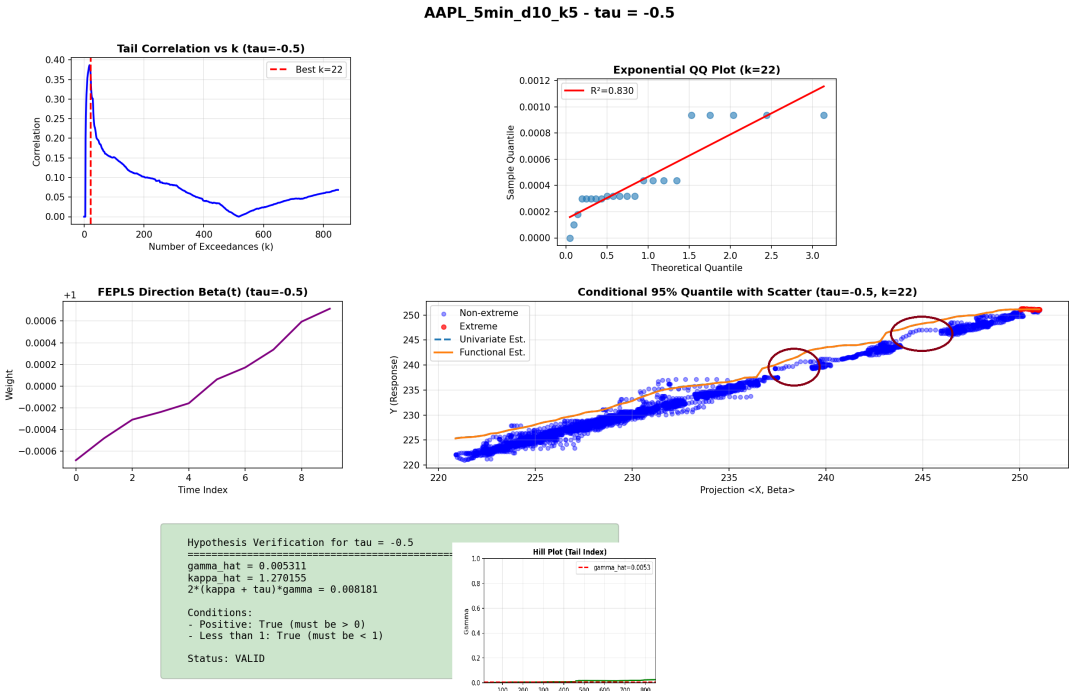


Figure 2: FEPLS analysis for AAPL using 5-minute high-frequency data. The figure shows the relationship between 5-minute return windows (X) and the maximum return of the next 5-minute window (Y). Despite challenges with Hill estimation and Q-Q plot fit due to microstructure effects, the increasing relationship clearly indicates trend-driven behavior at mid-frequencies.

Acknowledgments

We thank Charles-Albert Lehalle for providing the expensive high-frequency data.

References

[1] Girard, S., and Pakzad, A. 2023. *Functional Extreme Partial Least Squares*. Journal of Multivariate Analysis.

[2] de Haan, L., and Ferreira, A. 2007. *Extreme value theory: an introduction*. Springer Science & Business Media.

7 Appendix

7.1 Comparison between PCA and PLS

	PCA (Unsupervised)	**PLS (Supervised)**
Uses response variable?	No: PCA operates solely on the covariates X, ignoring any relationship with a response variable Y.	Yes: PLS explicitly incorporates the response Y in the optimization.
Objective	Find directions that maximize the *variance* of the projected data, i.e., $\arg\max_{\|w\|=1} \mathrm{Var}(\langle w, X\rangle)$.	Find directions that maximize the *covariance* between the projection and response, i.e., $\arg\max_{\|w\|=1} \mathrm{Cov}(\langle w, X\rangle, Y)$.
Interpretation	The first principal component captures the direction of maximum variance in X, regardless of its relevance to predicting Y.	The PLS direction captures the direction in X most correlated with Y, making it directly relevant to prediction.
Strengths/Limitations	Limitation: In regression contexts, high-variance directions can be orthogonal to the relationship with Y, so PCA may prioritize irrelevant directions.	Advantage: By focusing on covariance with Y, PLS finds directions that are both informative about X and predictive of Y.

Table 2: Comparison between PCA (Principal Component Analysis) and PLS (Partial Least Squares).

Column 1		**Column 2**	
Symbol	**Explanation**	**Symbol**	**Explanation**
Regular Variation & Auxiliary Functions			
$\mathrm{RV}_\tau(+\infty)$	Regularly-varying (index τ)	$2\mathrm{RV}_{\gamma,\rho}(+\infty)$	Second-order RV
$H_\rho(y)$	$y^\gamma \int_1^y u^{\rho-1}\mathrm{d}u$	$A(t) = t^\rho$	Auxiliary function, $A(t) \to 0$
Model and FEPLS Notation			
$X = g(Y)\beta + \varepsilon$	Inverse regression model	$\beta \in H$, $\|\beta\| = 1$	Index vector (unit norm)
$g \in \mathrm{RV}_\kappa(+\infty)$	Link function (index κ)	$\varepsilon : \Omega \to H$	Noise term
$\varphi \in \mathrm{RV}_\tau(+\infty)$	Test function (index τ)	$w_\varphi(y)$	Theoretical FEPLS direction
$\hat{\beta}_\varphi(y)$	Empirical FEPLS estimator		
Key Parameters			
$\gamma \in (0,1)$	Tail index of Y	$\rho \le 0$	Second-order parameter
$\kappa > 0$	Link/model function index	$\tau \in \mathbb{R}$	Test function index
$k = k_n$	Intermediate sequence	$y_{n,k} \sim U(n/k)$	High threshold
$\delta_{n,k}$	$(g(y_{n,k})(k/n)^{1/q})^{-1}$	$q > 2$	Noise integrability order

Table 3: Summary of main notation used in the FEPLS framework.

7.2 Additional Empirical Results: 4IG Stock Pairings

This section presents detailed FEPLS analysis results for 4IG paired with several other stocks from the Hungarian market. We selected this collection of stocks to illustrate both consistently estimated and non-consistently estimated cases, providing a realistic picture of what can be expected when applying the FEPLS framework to real financial data without curated experimental design.

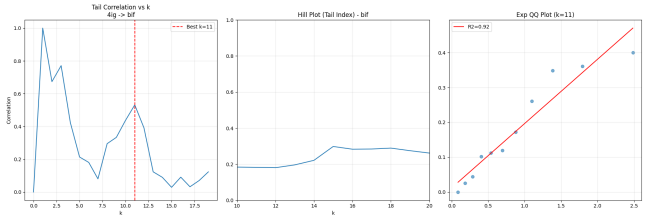
Additional figures for all stock pairs are available in the figures directory:

- Hill, Q-Q, and tail plots (files with `hill_qq_tail` in the name) for diagnostic assessment
- Conditional quantile plots (files with `conditional_quantile` in the name) showing the relationship between projections and extreme responses
- Beta comparison plots (4 digits in filename, e.g., `4ig_akko_tau_-2.0_-0.5.png`) for $\hat{\beta}$ across different τ pairs

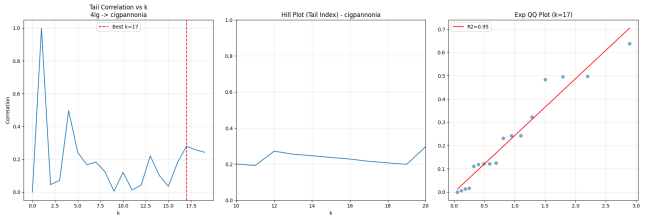
7.3 Conditional Quantile Plots

7.4 Beta Comparison Plots

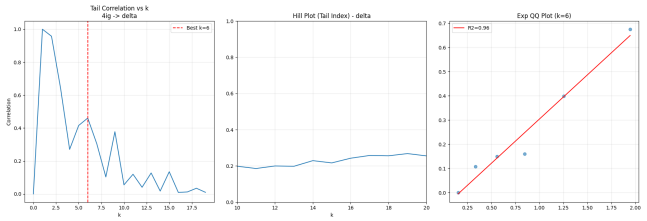
To visualize how the estimated FEPLS direction $\hat{\beta}$ changes across different pairs of τ values, we provide beta comparison plots below for 4IG $\rightarrow$ AKKO. Each plot corresponds to a specific pair of (τ_1, τ_2) values. This enables us to assess the stability and sensitivity of the FEPLS estimator under varying test function parameters.



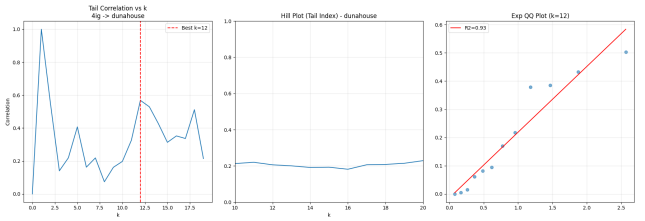
(a) 4IG → BIF



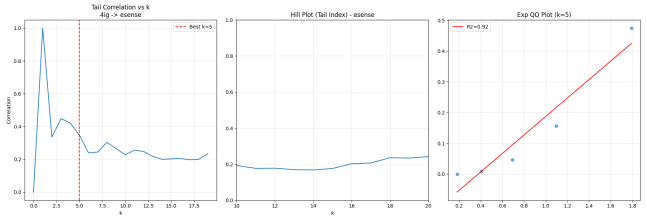
(b) 4IG → CIG Pannonia



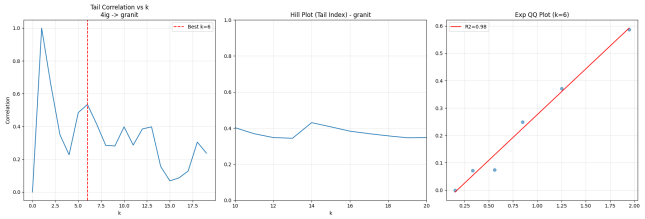
(c) 4IG → Delta



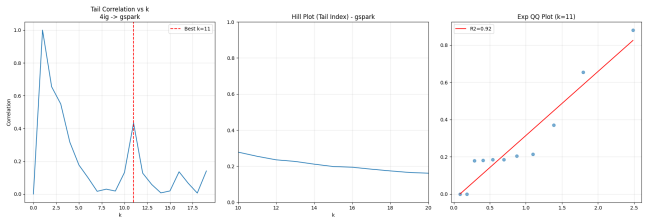
(d) 4IG → Dunahouse



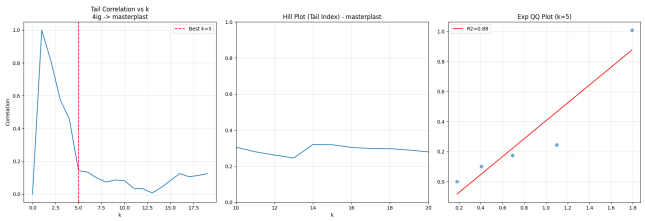
(e) 4IG → eSense



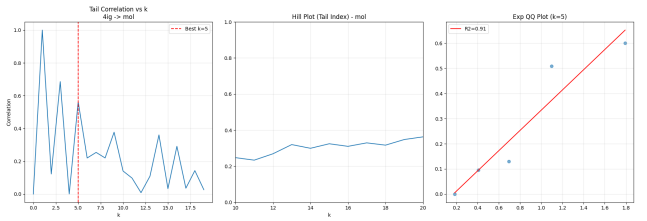
(f) 4IG → Granit



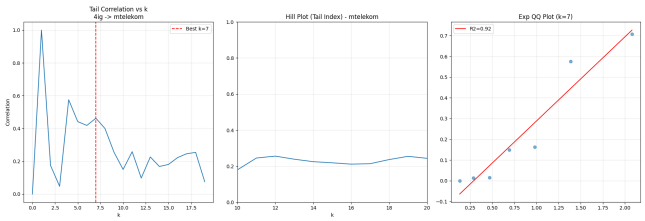
(g) 4IG → GSPark



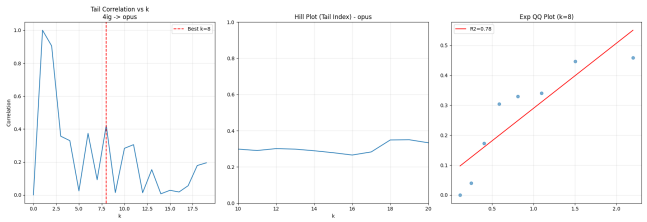
(h) 4IG → Masterplast



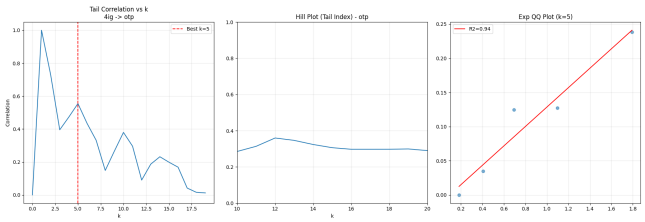
(i) 4IG → MOL



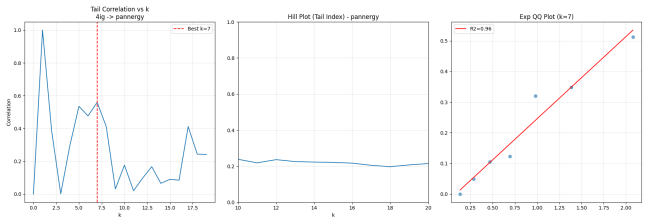
(j) 4IG → MTelekom

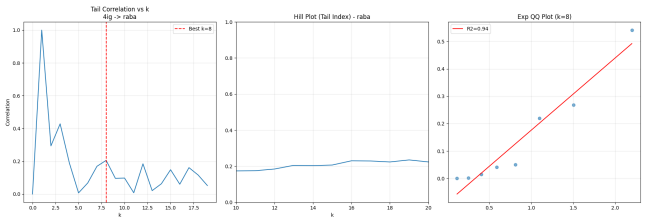


(k) 4IG → Opus



(l) 4IG → OTP





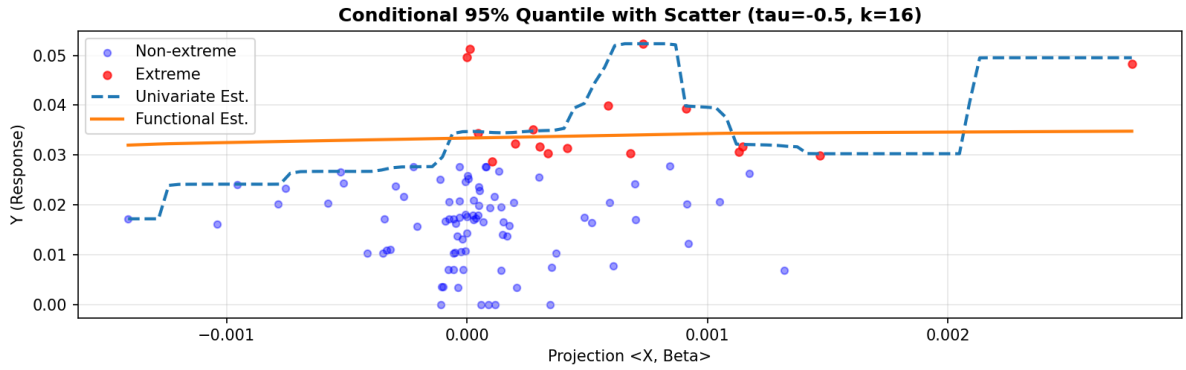


Figure 4: Conditional quantile plot for 4IG → AKKO.

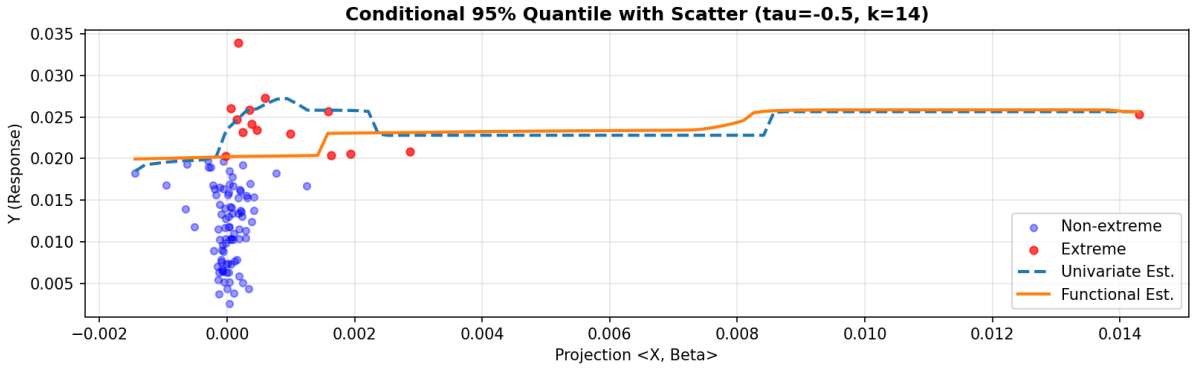


Figure 5: Conditional quantile plot for 4IG → Appeninn.

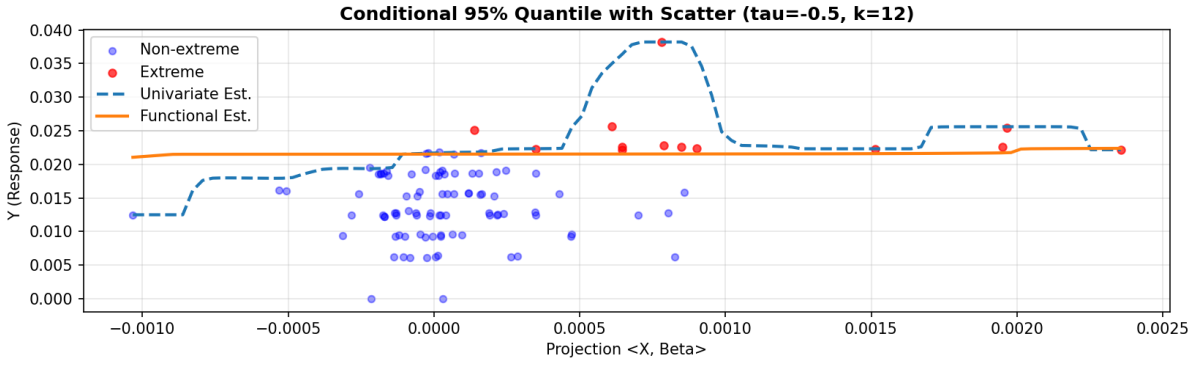
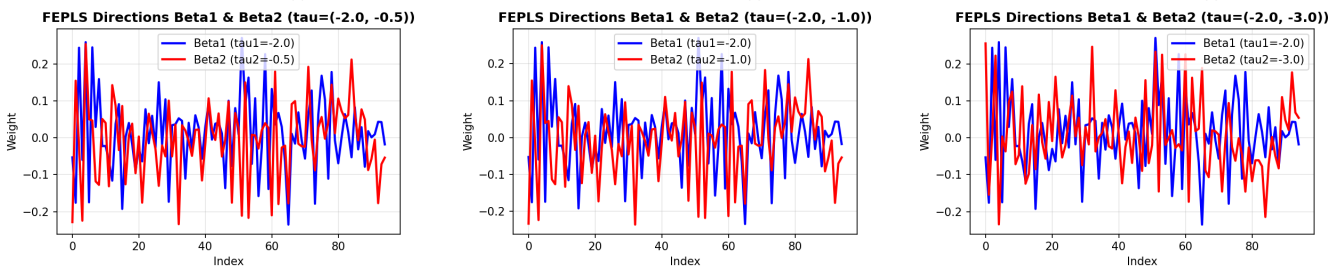


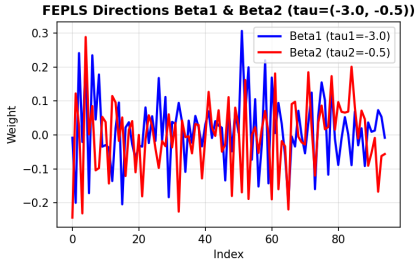
Figure 6: Conditional quantile plot for 4IG → Autowallis.



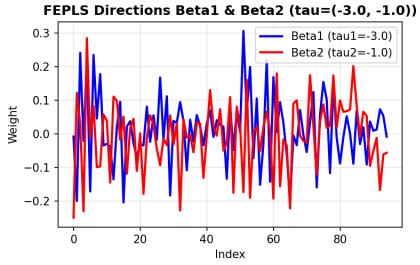
(a) $(\tau_1, \tau_2) = (-2.0, -0.5)$

(b) $(\tau_1, \tau_2) = (-2.0, -1.0)$

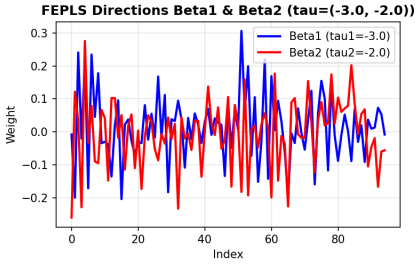
(c) $(\tau_1, \tau_2) = (-2.0, -3.0)$



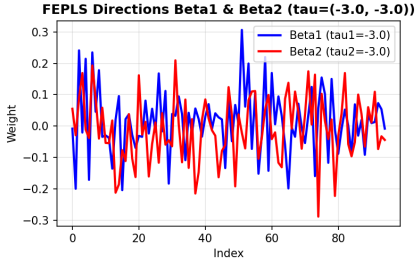
(d) $(\tau_1, \tau_2) = (-3.0, -0.5)$



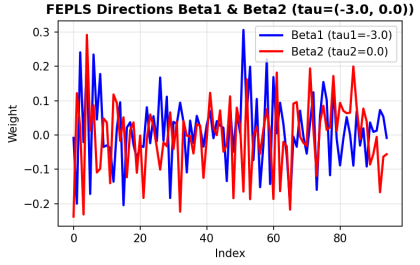
(e) $(\tau_1, \tau_2) = (-3.0, -1.0)$



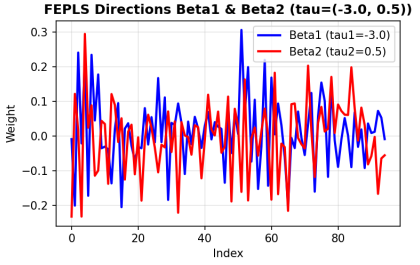
(f) $(\tau_1, \tau_2) = (-3.0, -2.0)$



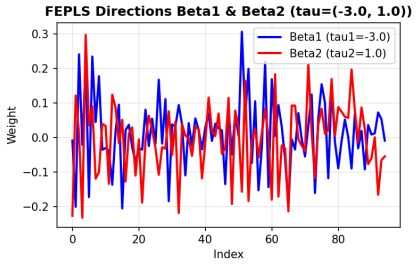
(g) $(\tau_1, \tau_2) = (-3.0, -3.0)$



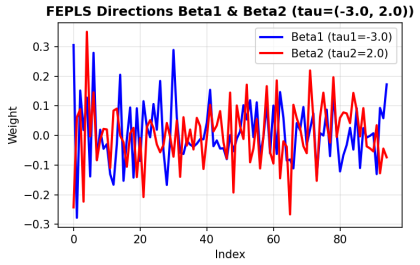
(h) $(\tau_1, \tau_2) = (-3.0, 0.0)$



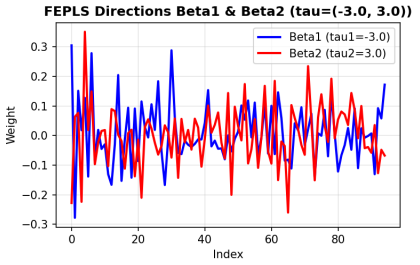
(i) $(\tau_1, \tau_2) = (-3.0, 0.5)$



(j) $(\tau_1, \tau_2) = (-3.0, 1.0)$



(k) $(\tau_1, \tau_2) = (-3.0, 2.0)$



(l) $(\tau_1, \tau_2) = (-3.0, 3.0)$

Figure 7: Beta comparison plots for 4IG $\rightarrow$ AKKO showing how the estimated FEPLS direction $\hat{\beta}$ varies across different pairs of τ values.