Volatility Jumps and Pricing Kernel Features

Wolfgang Karl Härdle
Levin Schach

Ladislaus von Bortkiewicz Professor of Statistics
Humboldt-Universität zu Berlin
theIDA.net
ASE, UEdinburgh, NYCU 玉山學者

Pricing Kernel

- Under the risk-neutral measure (RNM) Q, the current price p_t of an option with a payoff $\psi(S_T)$ at maturity T is given by

$$p_t = \mathsf{E}_Q\left\{\psi(S_T)\right\} = \int_0^\infty \psi(x)q(x)\,dx$$

- Relating the risk-neutral density (RND) to the historical density (HD) introduces the Pricing Kernel (PK) K, which adjusts the HD to reflect the RND

$$p_t = \int_0^\infty \psi(x)\frac{q(x)}{p(x)}p(x)dx = \int_0^\infty \psi(x)K(x)p(x)dx$$

- Therefore, we also express the price as an expectation under the physical measure (PM) P :

$$p_t = \mathsf{E}_P\left[\psi(S_T)K(S_T)\right]$$

Classical PK Estimator: Backward-looking return

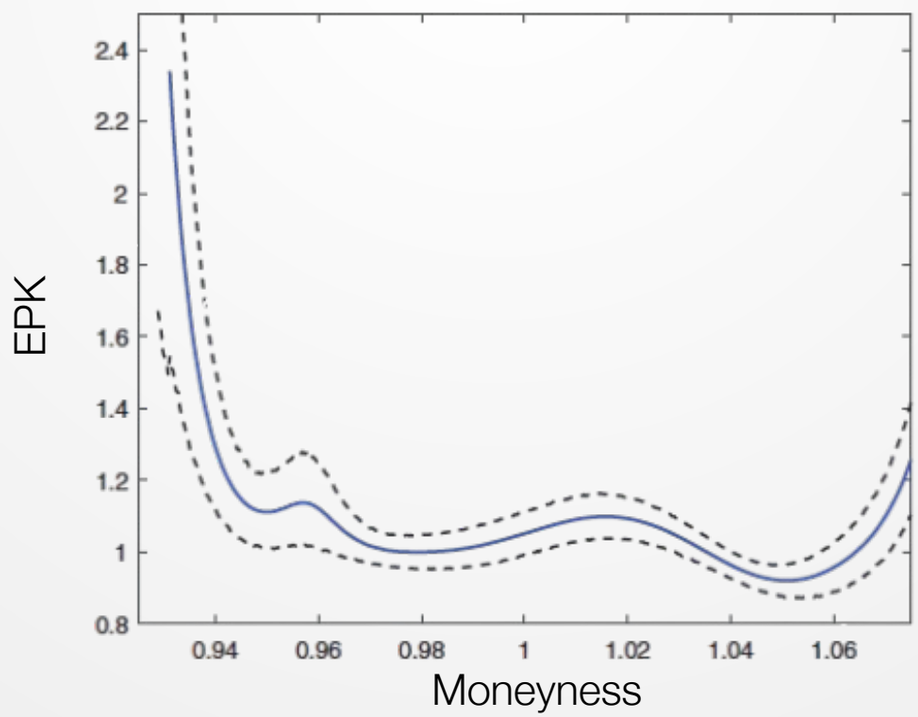
- RND and HD are estimated separately

$$\widehat{K}(S_T) = \frac{\widehat{q}(S_T)}{\widehat{p}(S_T)}$$

- $\widehat{q}$ estimated using option prices, while $\widehat{p}$ is estimated non parametrically (e.g. kernel density estimation)
- This PK represents the projection of any admissible stochastic discount factor (SDF) onto the space spanned by S_T
- P is estimated using backward-looking returns.

Pricing Kernel Puzzle

- According to economic theory the PK should be non-increasing in returns, reflecting the assumption that investors are risk averse.
- Empirically, it is often observed to be locally increasing in returns.
- This creates a "Pricing Puzzle", as it contradicts economic intuition:
  - Local "humps" in the PK suggest that investors prefer higher risk in certain regions of the return distribution.



Linn et al. Critique

- The Radon-Nikodym Theorem states that the SDF is the derivative that transforms the PM P into the RNM Q, both conditioned on the information set $\mathscr{F}_t$

$$m_t(x_{t+s}) = e^{-rs}\frac{dQ(x_{t+s} \mid \mathscr{F}_t)}{dP(x_{t+s} \mid \mathscr{F}_t)}$$

- Breeden and Litzenberger's results enables the forward-looking estimation of RND, conditional on all information used by investors.
- There is no method to estimate the forward-looking HD of returns from *t* to *t*+*s*.

$\succ$ The SDF is inconsistent with microeconomic theory.

Test the robustness of CDI in different volatility regimes

- The proposed Conditional Density Integration addresses inconsistencies by matching forward-looking option prices with realized returns.

- Claim: CDI estimates remain stable under conditions of stochastic volatility and sudden volatility jumps as it correctly matches conditional information.
  - Resolves PK-Puzzle

- We simulate option series and introduce abrupt volatility jumps to assess the behavior of the proposed pricing kernel.

Outline

1. Motivation✓
2. Conditional Density Integration (CDI) Approach
3. Challenges of averaging volatility regimes in PK estimation
4. Black-Scholes Pricing Model
5. Risk Neutral Density & Historical Density Calculation
6. Options simulation
7. Risk-free-density Data overview
   - Methodology
   - Experimental results

Conditional Density Integration (CDI) by Linn et al. (2018)

- Linn et al. integrate the RND, adjusted by the Radon-Nikodym derivative relative to the HD, up to the realized returns, ensuring alignment under the assumption that both densities should ideally yield a uniform distribution when correctly modeled.

$$\int_{-\infty}^{r_{t+s}} dP(r_{t+s}|\mathcal{F}_t) = \int_{-\infty}^{r_{t+s}} \frac{dP(r_{t+s}|\mathcal{F}_t)}{dQ(r_{t+s}|\mathcal{F}_t)} dQ(r_{t+s}|\mathcal{F}_t)$$

$$= \int_{-\infty}^{r_{t+s}} \left\{ \frac{dQ(r_{t+s}|\mathcal{F}_t)}{dP(r_{t+s}|\mathcal{F}_t)} \right\}^{-1} dQ(r_{t+s}|\mathcal{F}_t)$$

- $dP(r_{t+s}|\mathcal{F}_t)$ is the physical, $dQ(r_{t+s}|\mathcal{F}_t)$ the risk-neutral density and r_t is the return at time point t

Probability Integral Transformation

- Probability Integral Transformation (PIT)
  - $F(X)$, where F is the CDF of X, transforms any variable X into a uniformly distributed variable over [0,1].
- PIT validates the uniqueness

$$\int_{-\infty}^{r_{t+s}} \left\{ \frac{dQ(r_{t+s} \mid \mathscr{F}_t)}{dP(r_{t+s} \mid \mathscr{F}_t)} \right\}^{-1} dQ(r_{t+s} \mid \mathscr{F}_t) = \int_{-\infty}^{r_{t+s}} g_{t,s}(y) dQ(r_{t+s} \mid \mathscr{F}_t) \sim U(0,1)$$

Estimation

- Assume cubic B-spline

$$g(y) = \sum_{j=1}^{J} \theta_j B_j(y)$$

- Hence,

$$\int_{-\infty}^{r_{t+s}} g(y) dQ(y|\mathcal{F}_t) = \sum_{j=1}^{J} \theta_j \int_{-\infty}^{r_{t+s}} B_j(y) dQ(y|\mathcal{F}_t)$$

- Use GMM to get $\hat{\theta}$ (moments of the uniform distribution)

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^J}{\operatorname{argmin}} \sum_{i=1}^{m} \left[\frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{J} \theta_j \int_{-\infty}^{r_{t+s}} B_j(y) dQ(y|\mathcal{F}_t) \right\}^i - \frac{1}{i+1} \right]^2$$

Estimation

- It should be

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^J}{\operatorname{argmin}} \sum_{i=1}^{m} \left[\frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{J} \theta_j \int_{-\infty}^{R_{t+s}} B_j(y) dQ(y|\mathcal{F}_t) \right\}^i - \frac{1}{i+1} \right]^2$$

because

$$\sum_{j=1}^{J} \theta_j \int_{-\infty}^{R_{t+s}} B_j(y) dQ(y|\mathcal{F}_t) \approx \int_{-\infty}^{R_{t+s}} g_{t,s}(y) dQ(y|\mathcal{F}_t) \sim U(0,1)$$

$$\frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{J} \theta_j \int_{-\infty}^{R_{t+s}} B_j(y) dQ(y|\mathcal{F}_t) \right\}^i = \hat{E} \left[\left\{ \sum_{j=1}^{J} \theta_j \int_{-\infty}^{R_{t+s}} B_j(y) dQ(y|\mathcal{F}_t) \right\}^i \right]$$

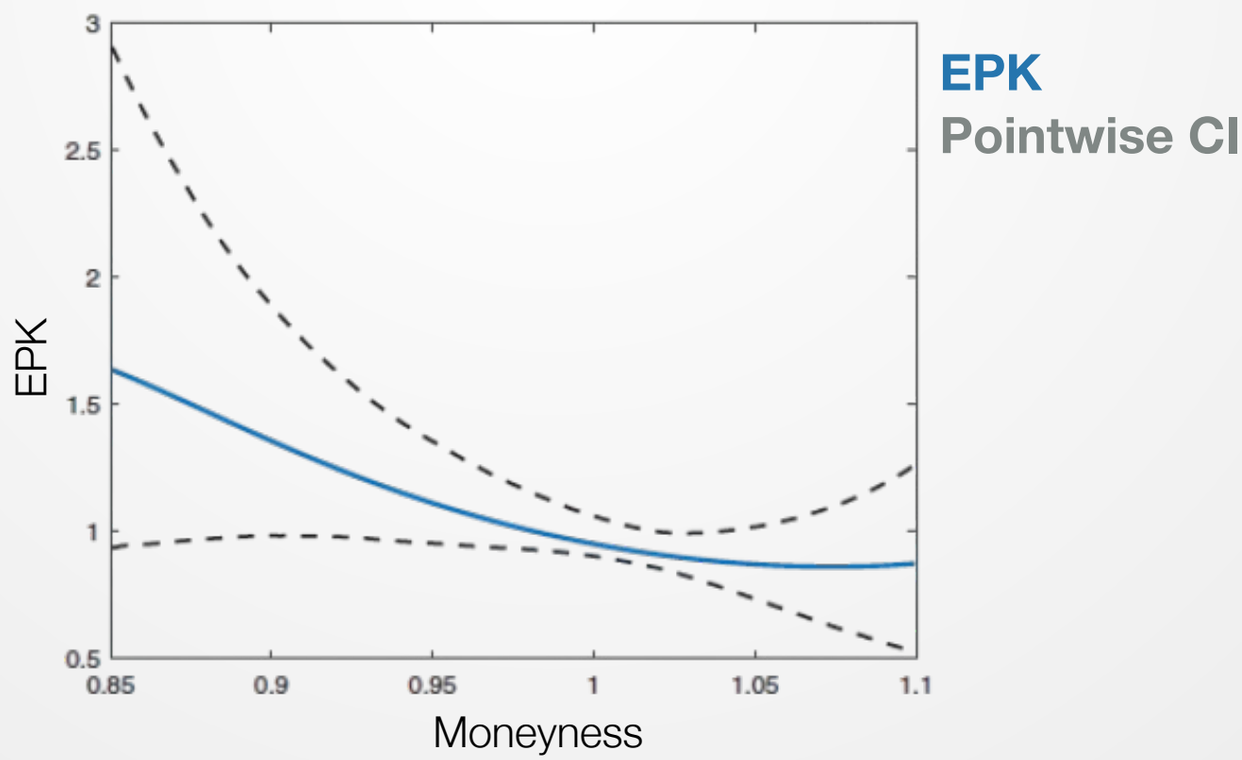
Should be close to $\dfrac{1}{i+1}$

Estimation II

- Plug-in estimate given by

$$\hat{g}(r_{t+s}) = \sum_{j=1}^{J} \hat{\theta}_j B_j(r_{t+s})$$

- EPK follows as $\widehat{K}(r_{t+s}) = \left\{\hat{g}(r_{t+s})\right\}^{-1}$

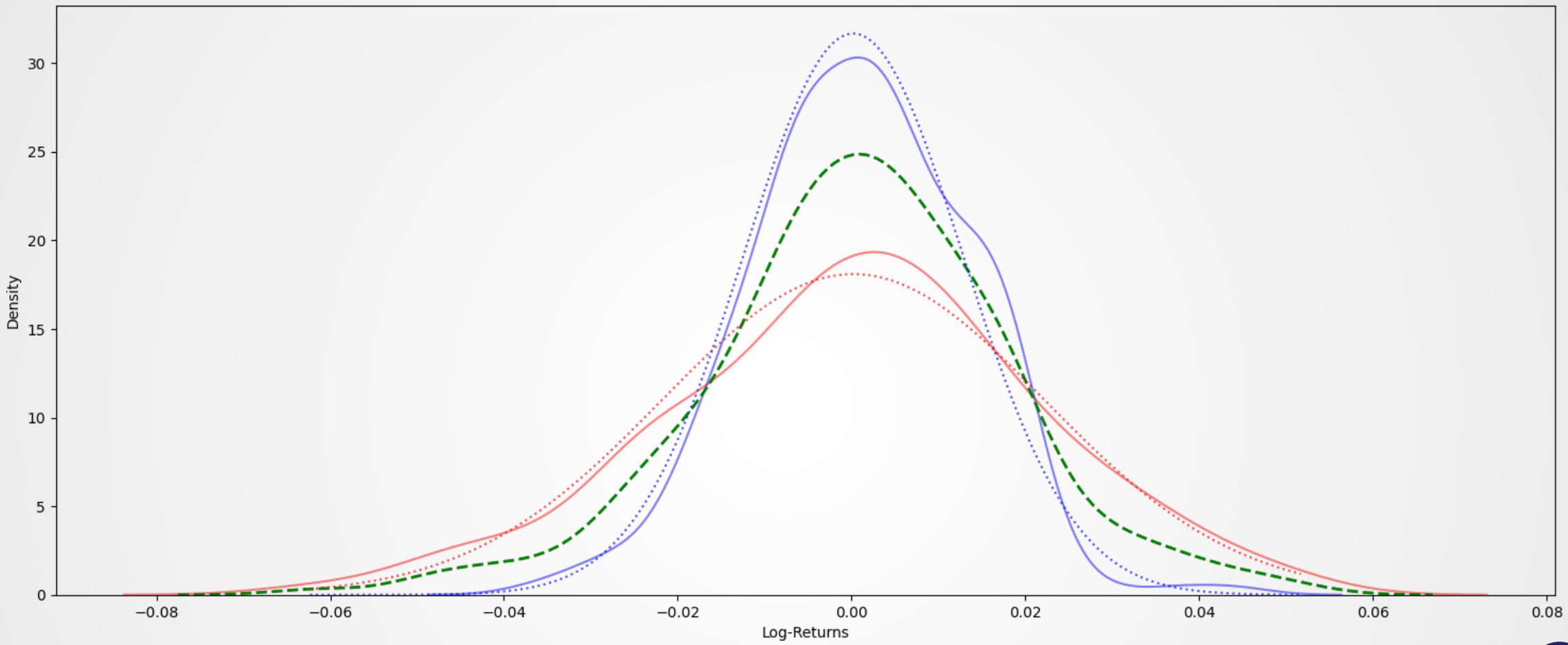


Physical density in the presence of volatility jumps

- The KDE ignores the volatility regimes the samples belong to, as it assumes a stationary distribution of returns.
  - In the presence of volatility jumps, this leads to a blending of characteristics from different regimes into a single density estimate.
- The averaged density under- or overestimates the probability of extreme values, as it doesn't account for different volatility levels separately.

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right)$$

Physical density in the presence of volatility jumps



KDE regime 1 (σ = 20%), KDE regime 2 (σ = 35%), **Dotted**: Theoretical distribution of regime 1,Theoretical distribution of regime 2, Total KDE

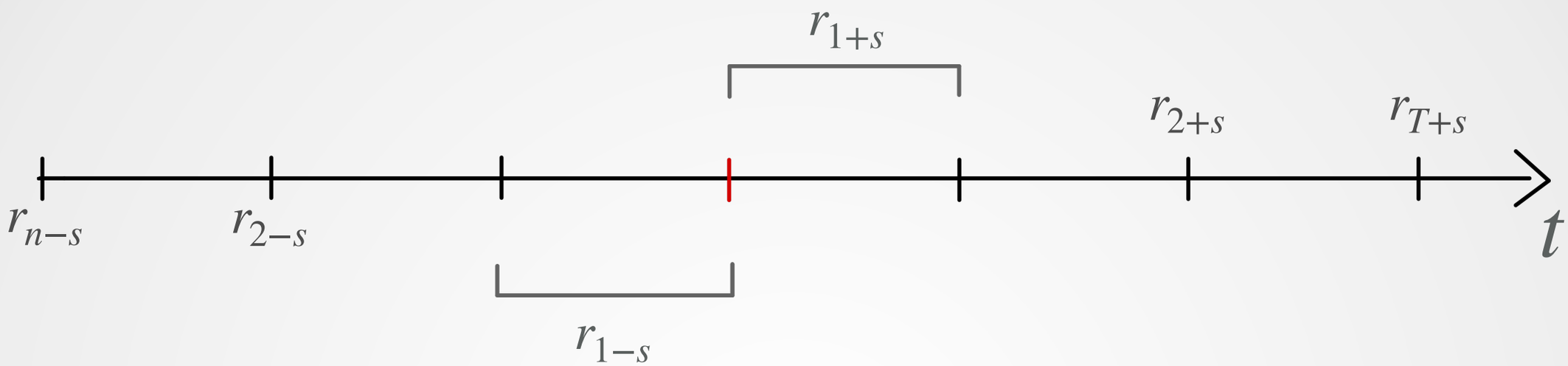
- KDE for HD estimation relies on temporal averaging, which performs well for stable distributions but fails when structural breaks in volatility occur.
- Here, the KDE produces a blended density that fails to accurately capture the characteristics of either individual regime.

CDI: Averaging across volatility regimes

- Temporal averaging in the GMM estimation yields a single θ vector that balances observations across T periods, masking regime-specific characteristics.
- Monotonicity may be an artifact of averaging across different market regimes rather than reflecting true investor preferences.
- The method's reliance on temporal averaging makes it susceptible to structural breaks in volatility.

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^J}{\operatorname{argmin}} \sum_{i=1}^{m} \left[\frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{J} \theta_j \int_{-\infty}^{r_{t+s}} B_j(y) dQ(y|\mathscr{F}_t) \right\}^i - \frac{1}{i+1} \right]^2$$

Temporal Averaging in PK Estimation Methods



Classical PK estimation

$$R^{\text{KDE}} = \{r_{1-s}, r_{2-s}, \ldots, r_{n-s}\}$$

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - R_i^{KDE}}{h}\right)$$

Conditional density integration

$$R^{\text{CDI}} = \{r_{1+s}, r_{2+s}, \ldots, r_{T+s}\}$$

$$\frac{1}{T} \sum_{t=1}^{T} \left\{ \sum_{j=1}^{J} \theta_j \int_{-\infty}^{R_i^{CDI}} B_j(y) dQ(y|\mathscr{F}_t) \right\}$$

$\succ$ Classical method averages the past,
CDI averages the future

Change of Measure

- The stock price S is a geometric Brownian motion, represented by

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}, \quad W_t \text{ a } P\text{-Brownian motion}$$

- To price derivatives, we switch from the PM P to the RNM Q, ensuring all investors are risk-indifferent and discount future cash flows at the risk-free rate.
- The No-arbitrage rule, using the Cameron-Martin-Girsanov theorem, changes the probability measure to Q by adjusting the process drift to r, ensuring risk-neutral pricing.

$$S_t = S_0 \exp\left\{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t^*\right\}, \quad W_t^* \text{ a } Q\text{-Brownian motion}$$

Black-Scholes Model - European Call

- The present value of the expected payoff under the RNM is given as

$$C_t = e^{-r\tau}\mathsf{E}_t^Q\left[(S_T - K)^+\right] = e^{-r\tau}\int_0^\infty (s-K)^+ q_t(s)\,ds,$$

- In the Black-Scholes world, the price is derived as

$$C_{BS} = S_t\Phi(d_1) - e^{-r\tau}K\Phi(d_2),$$

$$d_1 = \frac{\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},$$

$$d_2 = d_1 - \sigma\sqrt{T}\,.$$

Risk Neutral Density

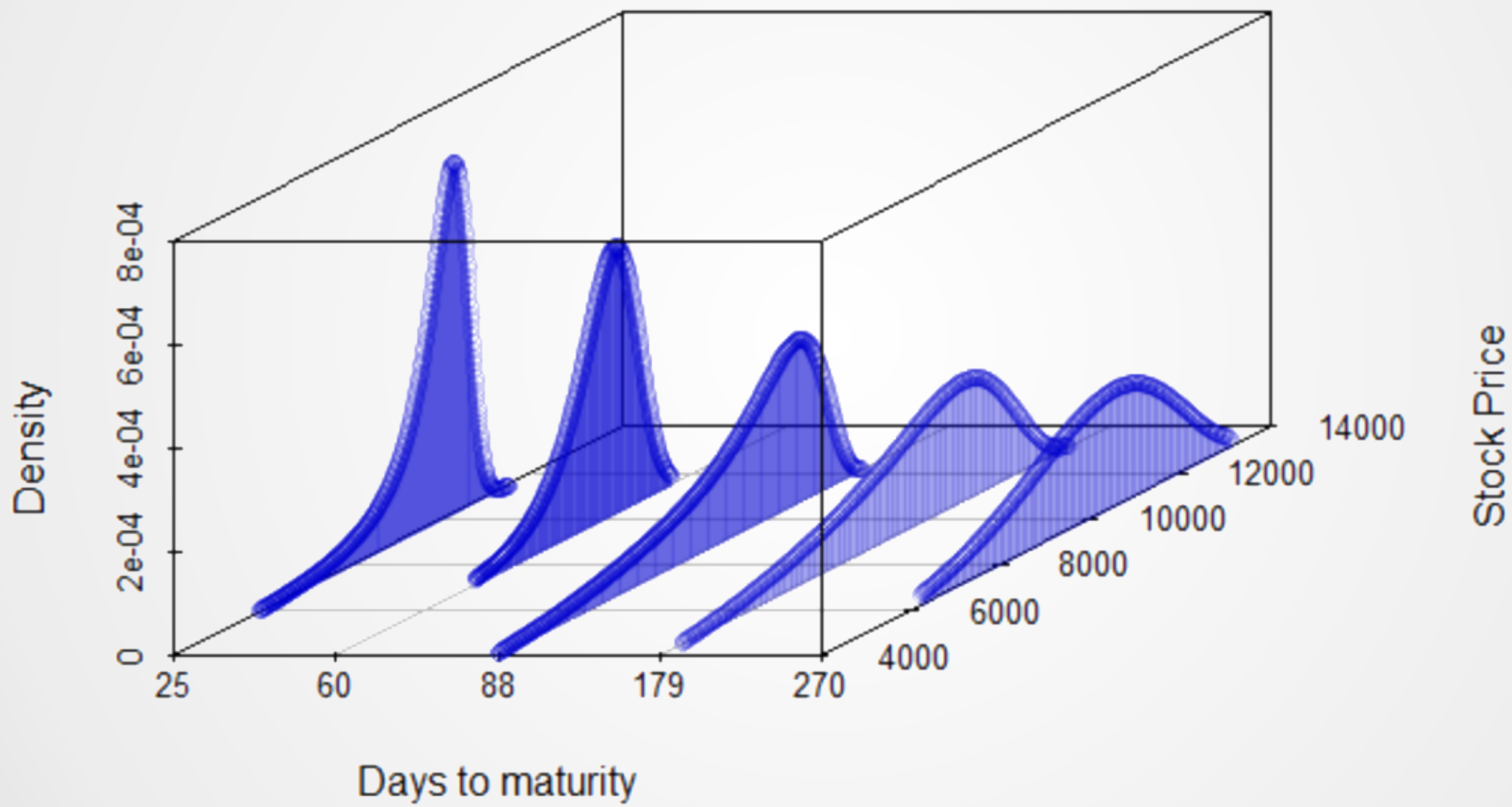
- Breeden and Litzenberger derived the RND as:

$$C_t = e^{-r\tau} \int_0^\infty (s - K)^+ q_t(s)\, ds,$$

$$q(\,\cdot\,) = e^{r\tau} \frac{\partial^2 C}{\partial K^2}$$

- To calculate the second derivative, we apply the Rookley-Method and estimate $\sigma_{IV}(M)$, $\sigma'_{IV}(M)$, $\sigma''_{IV}(M)$ by local polynomials, to smooth the implied volatility surface.
- We use the estimates to extract the risk neutral density based on the result of Breeden and Litzenberger.

Risk Neutral Density for the DAX



Physical Density

- Retrieve historical price data and calculate the return over the interval s for each day i

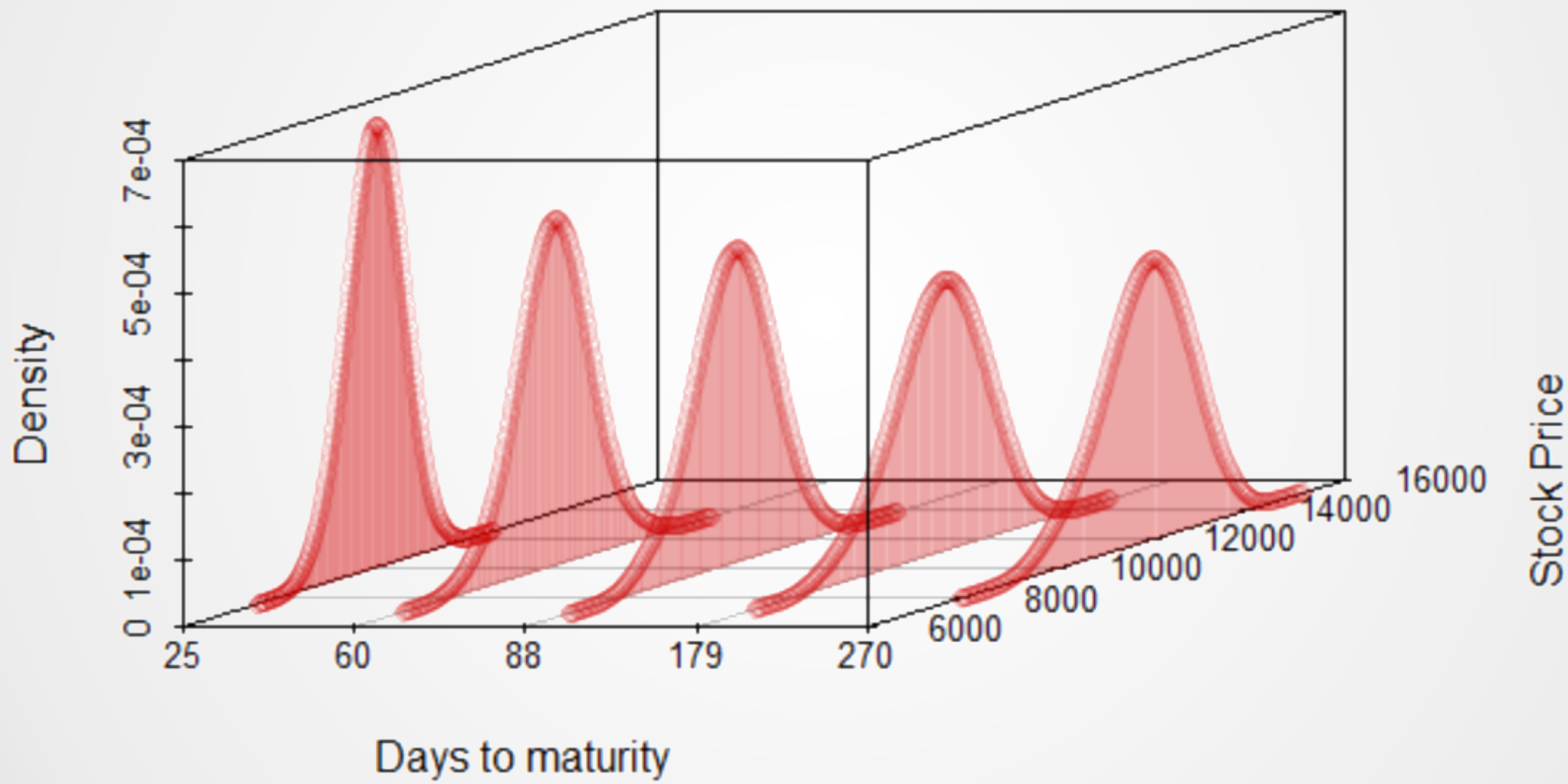
$$R_i = \frac{S_{t-i}}{S_{t-i-s}}, \quad \text{where } i = 1,2,\ldots,n$$

- Create a bootstrap sample on historical returns $\{R_i\}$. 10.000 random samples r_i are transformed into future price scenarios by multiplying it to current price S_t

$$S_T^i = S_t r_i^s, \; i = 1,2,\ldots,10.000$$

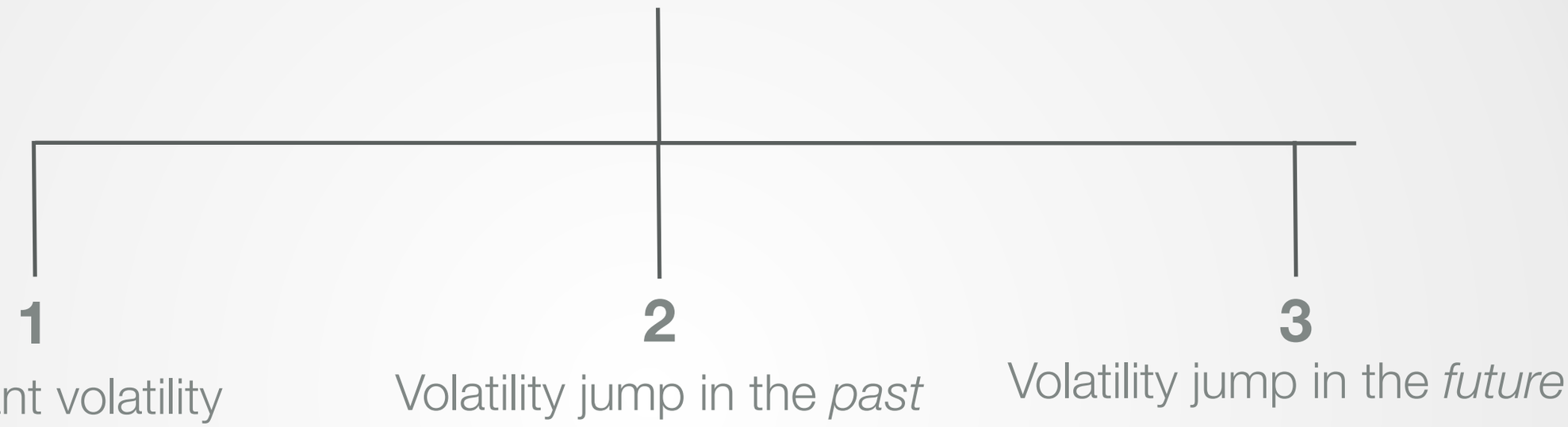
- Estimate physical density based on samples S_T^i using kernel density estimation.

Physical Density for the DAX



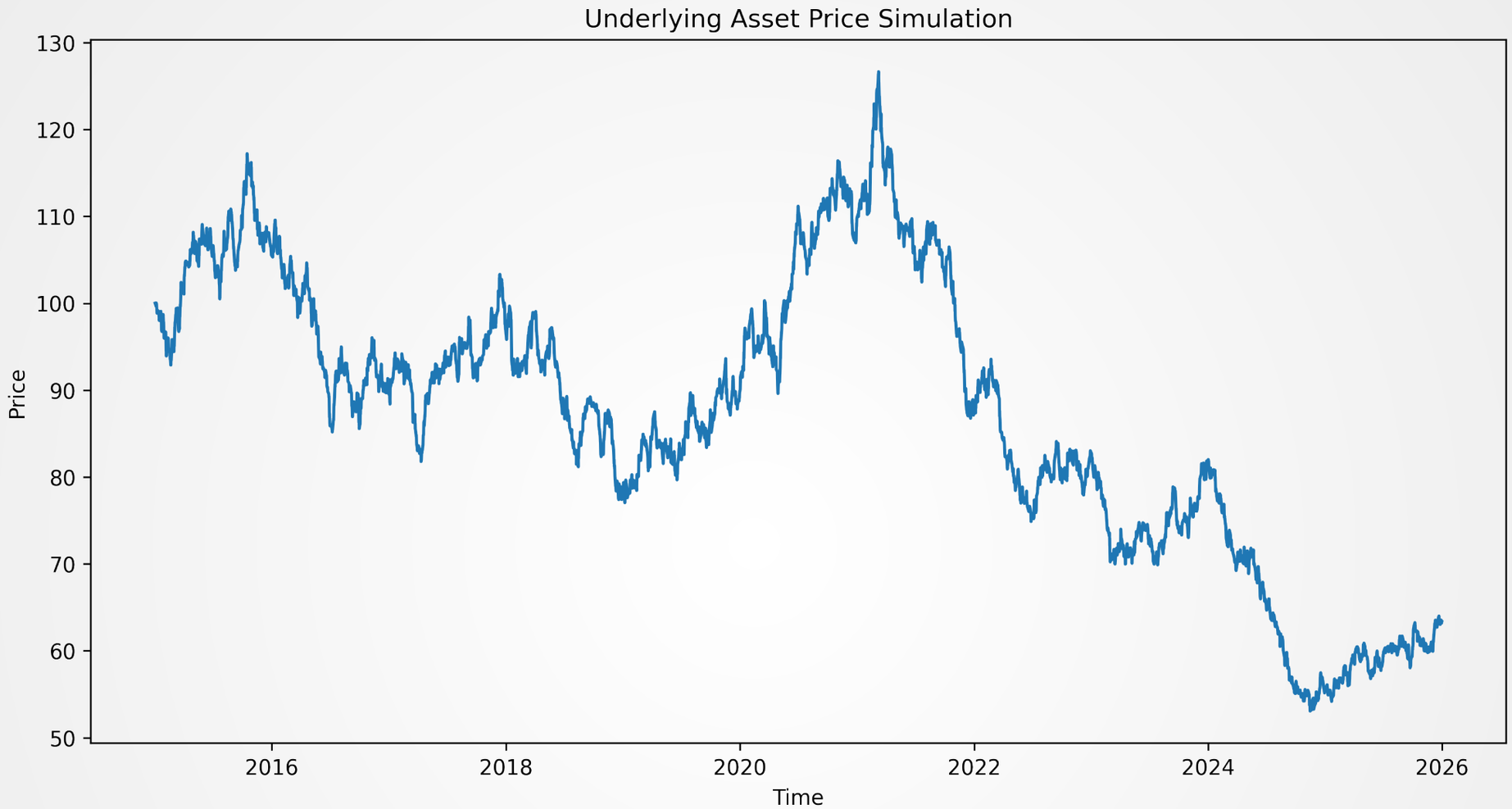
Volatility jumps

Simulated volatility regimes



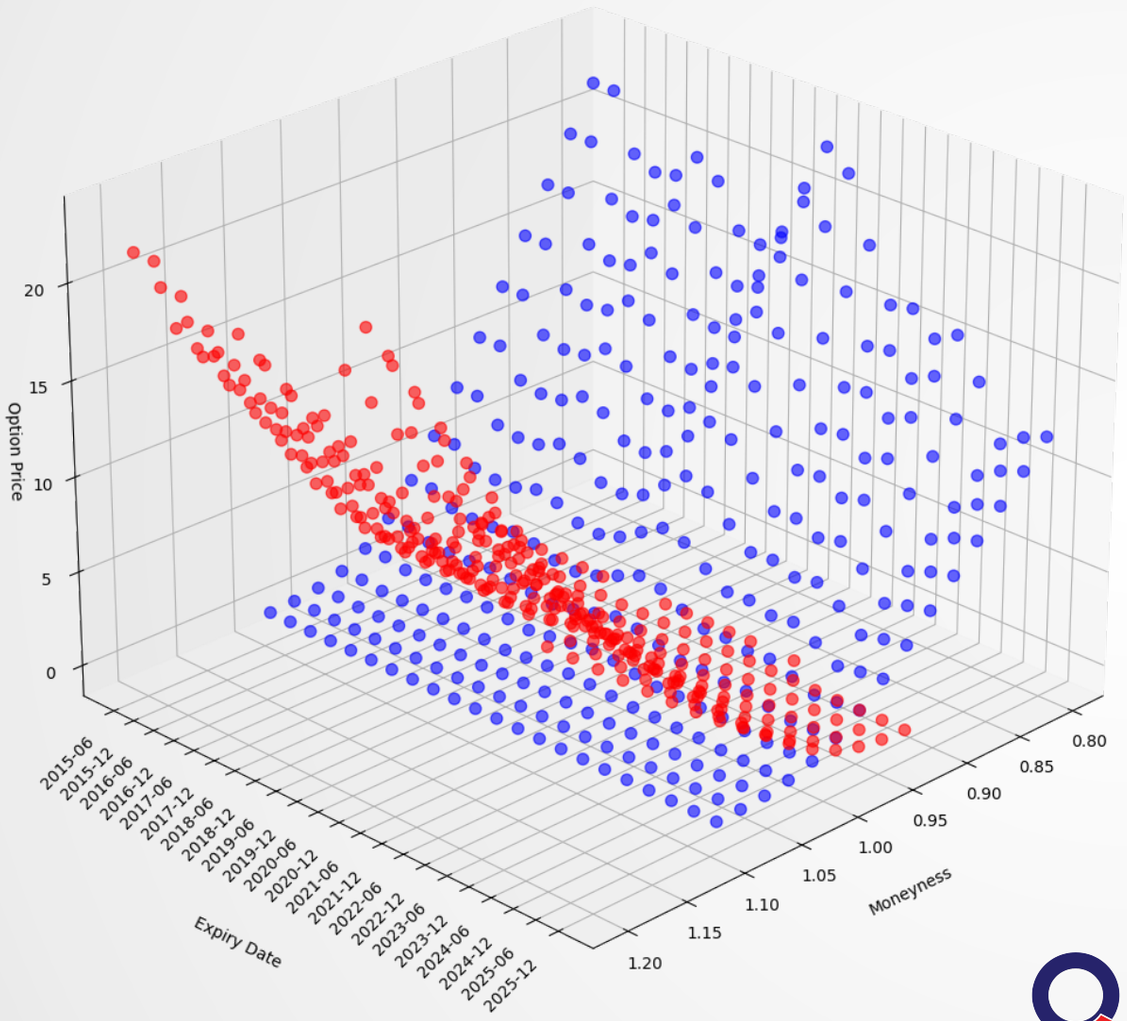
- To analyze the impact of volatility regime switches on PK estimation methods, we examine the PKs at the same point in time under different overall volatility regime scenarios.
- The classical PK estimation relies on historical returns. We assumed that a volatility jump in the past influences the kernel, while a volatility jump in the future does not.
- In contrast, the CDI PK estimation uses forward-looking returns. In theory past regime switches have no effect on the pricing kernel compared to future regime switches.

Price simulation



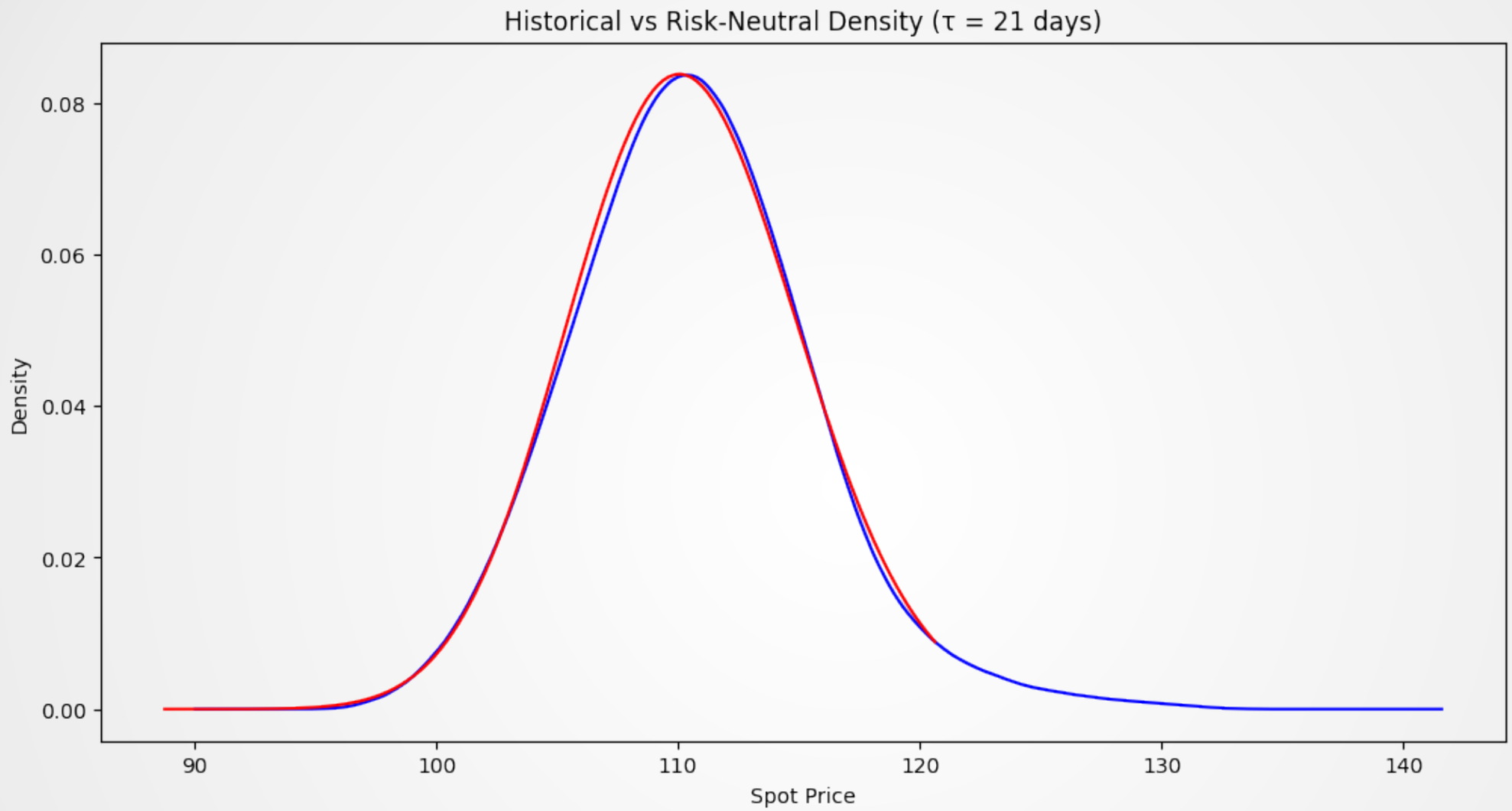
- The underlying price is simulated as a Q-Brownian motion, with
  - $S_0 = 100$
  - $r = 0$
  - $\sigma = 0.15$

Options simulation



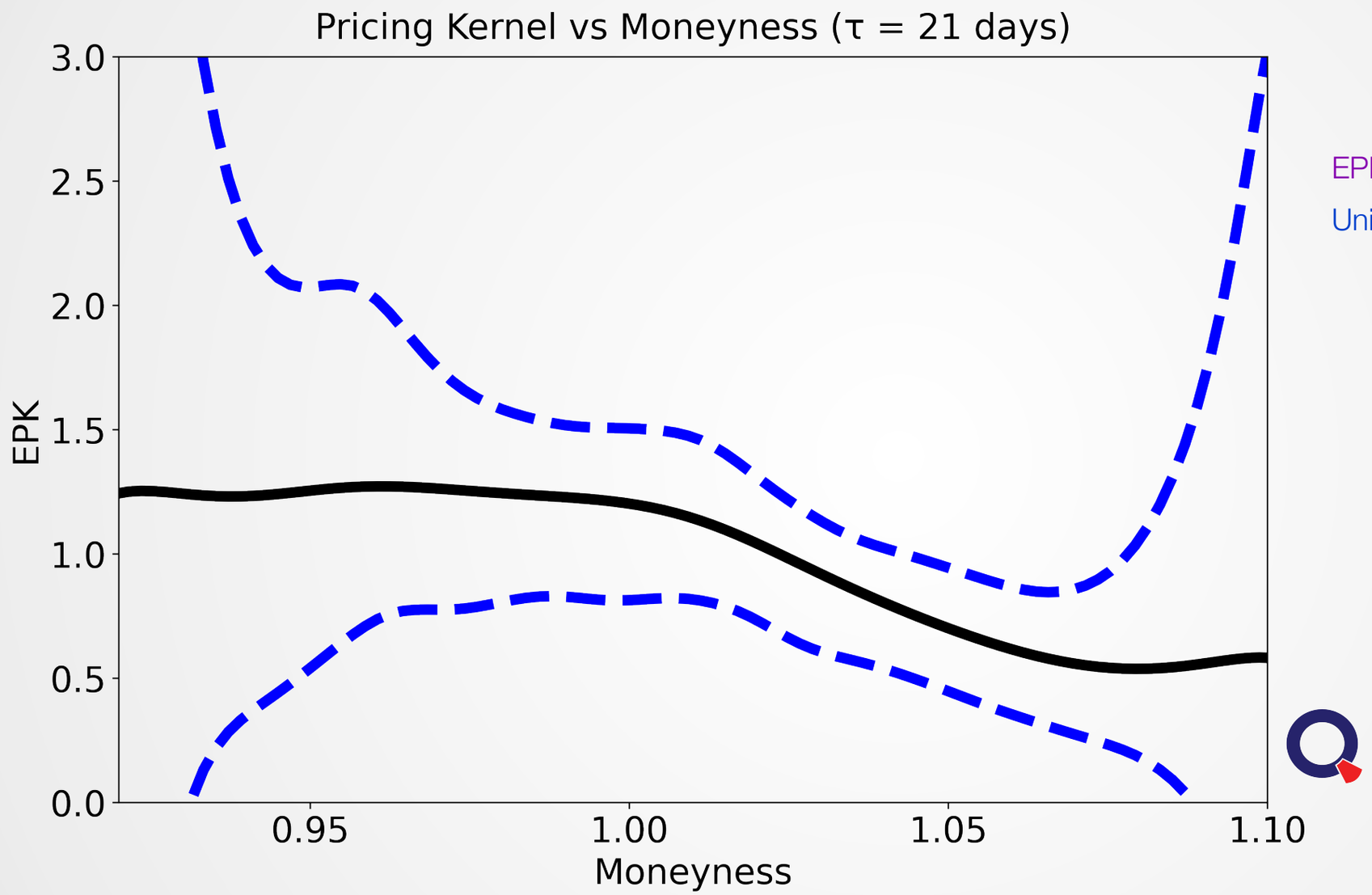
- On the first trading day of each expiry, 20 puts and calls are simulated within a moneyness range of 0.8 to 1.2.
- They are valued using the Black-Scholes model and filtered for outliers.

Case 1: RND and HD at constant volatility



- RND and HD at 2020-01-10 with
  - $\sigma = 0.15$
  - $r = 0.00$
  - $\tau = 21$

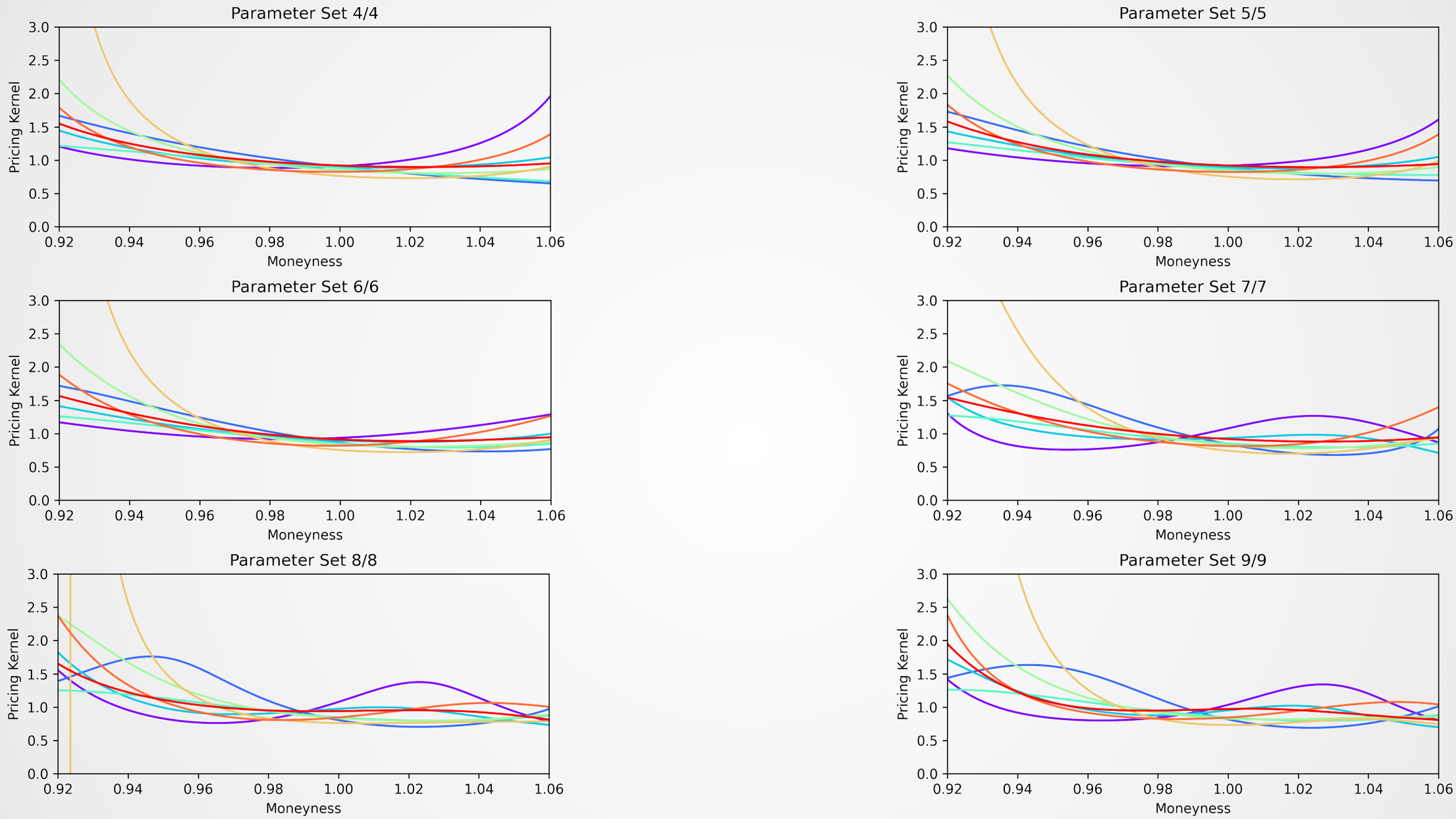
Case 1: Classical Pricing Kernel



EPK,

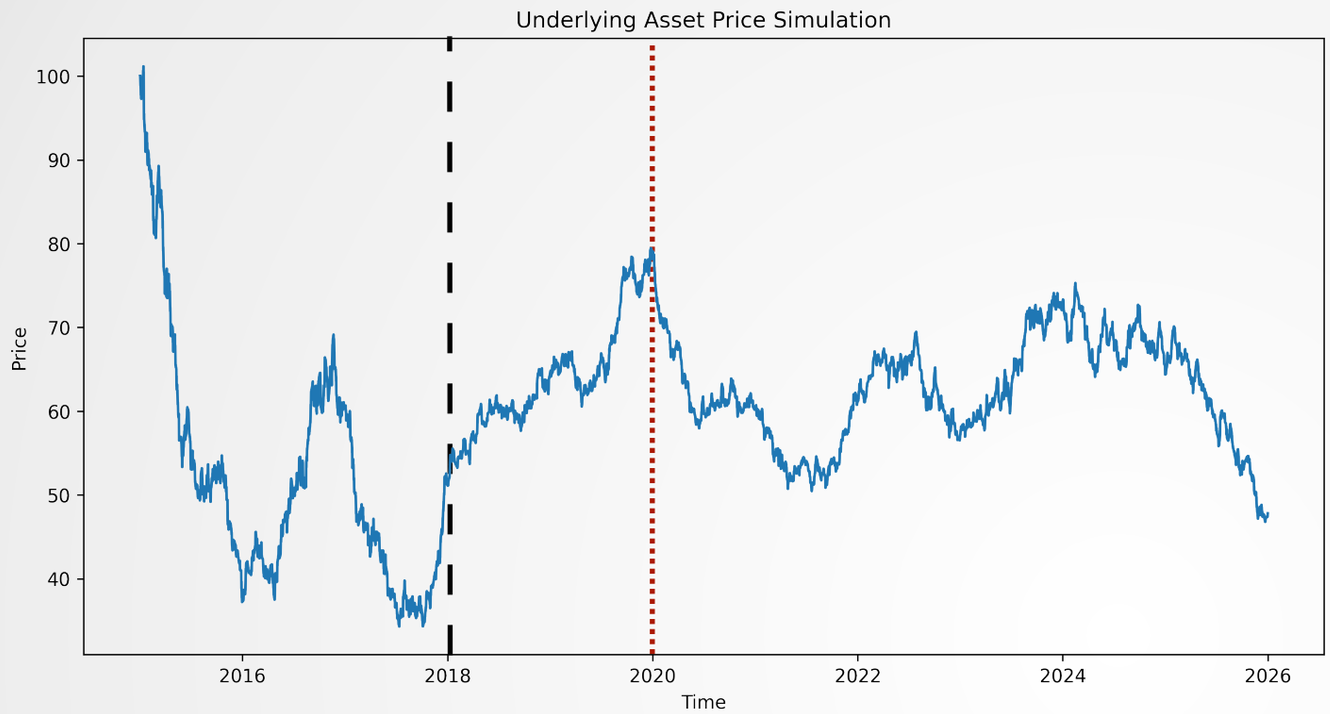
Uniform CI Bands

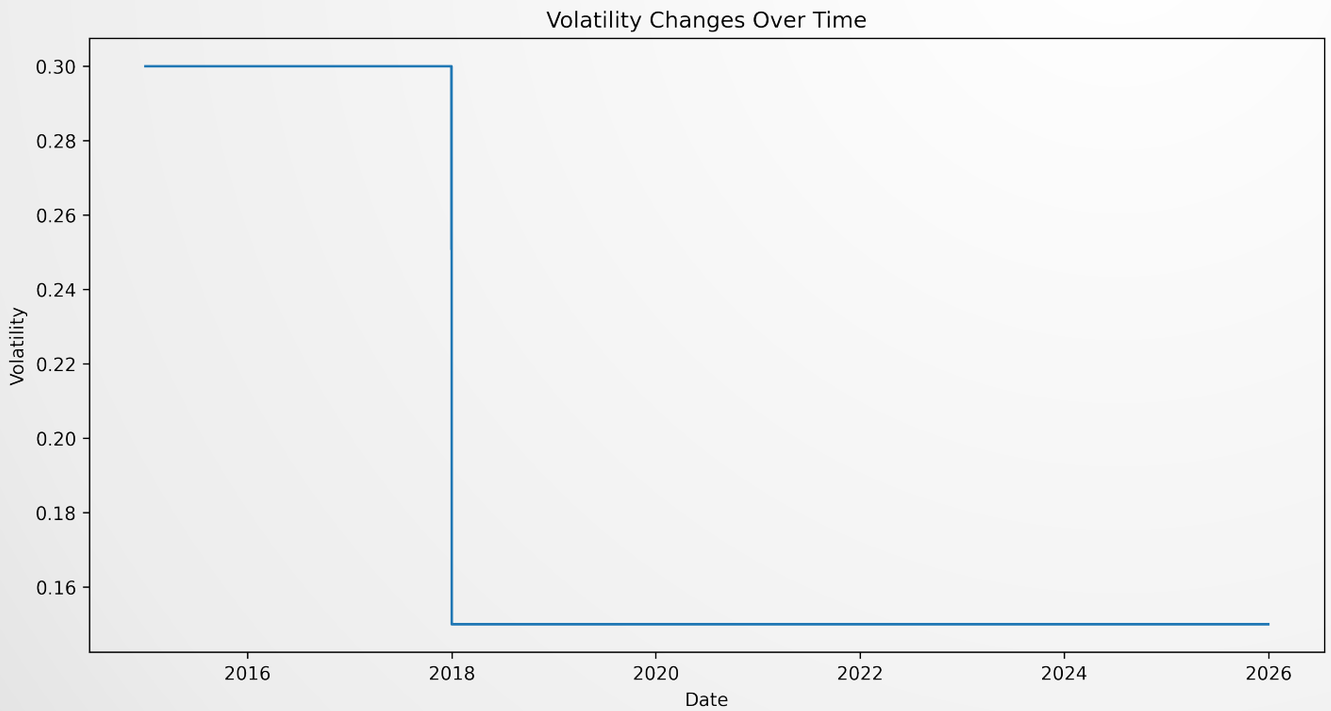
Case 1: CDI kernel variations by parameter set



- Lower parameter sets (4–6 moments/basis functions) mostly produce monotonically decreasing kernels.
- In higher sets (7-9) boundary effect at high moneyness level become apparent.

Case 2: Volatility jump in the *past*



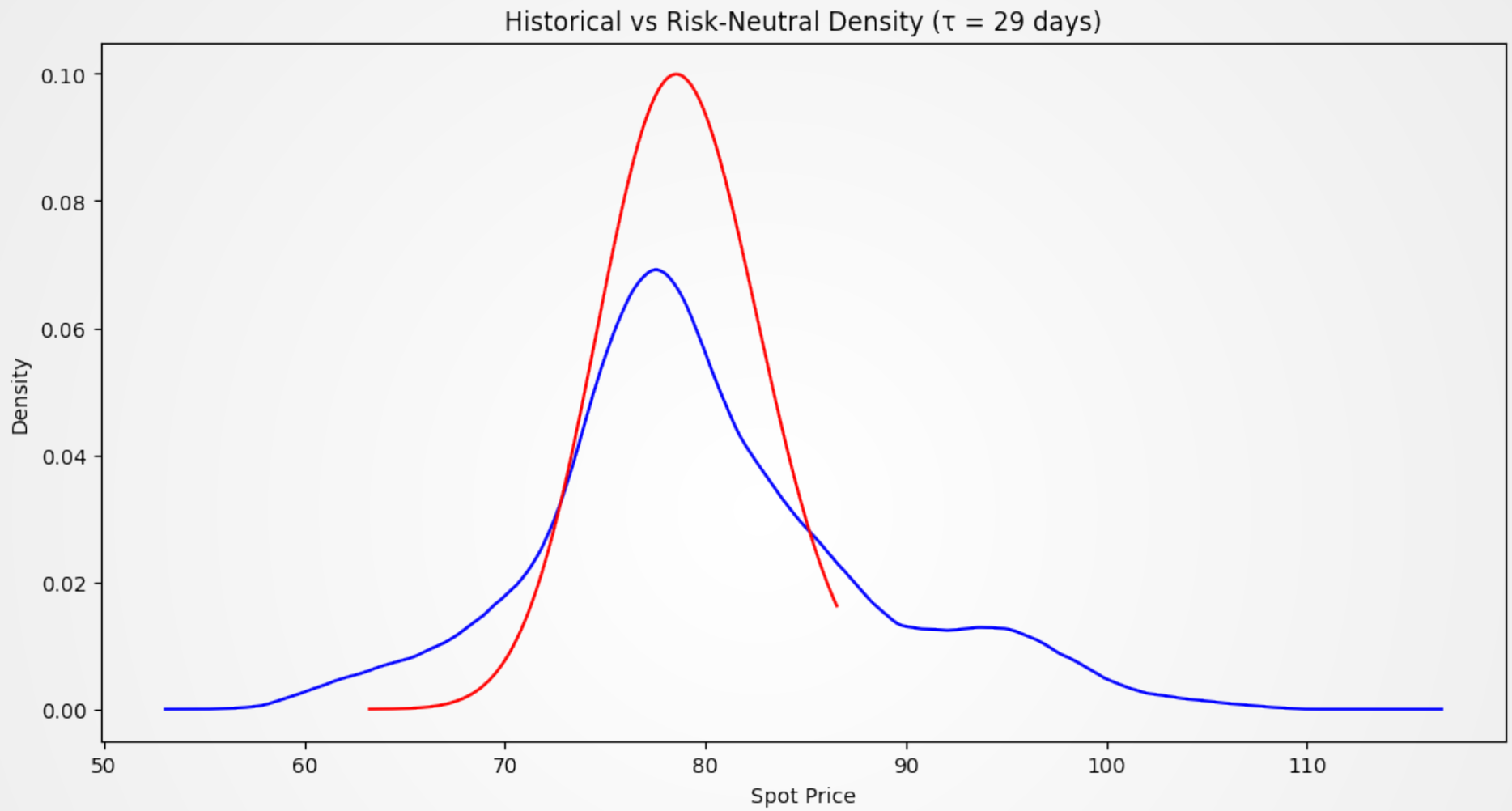


- Estimation of the PK
- Volatility jump

- Simulate a volatility regime shift at 2018-01-01

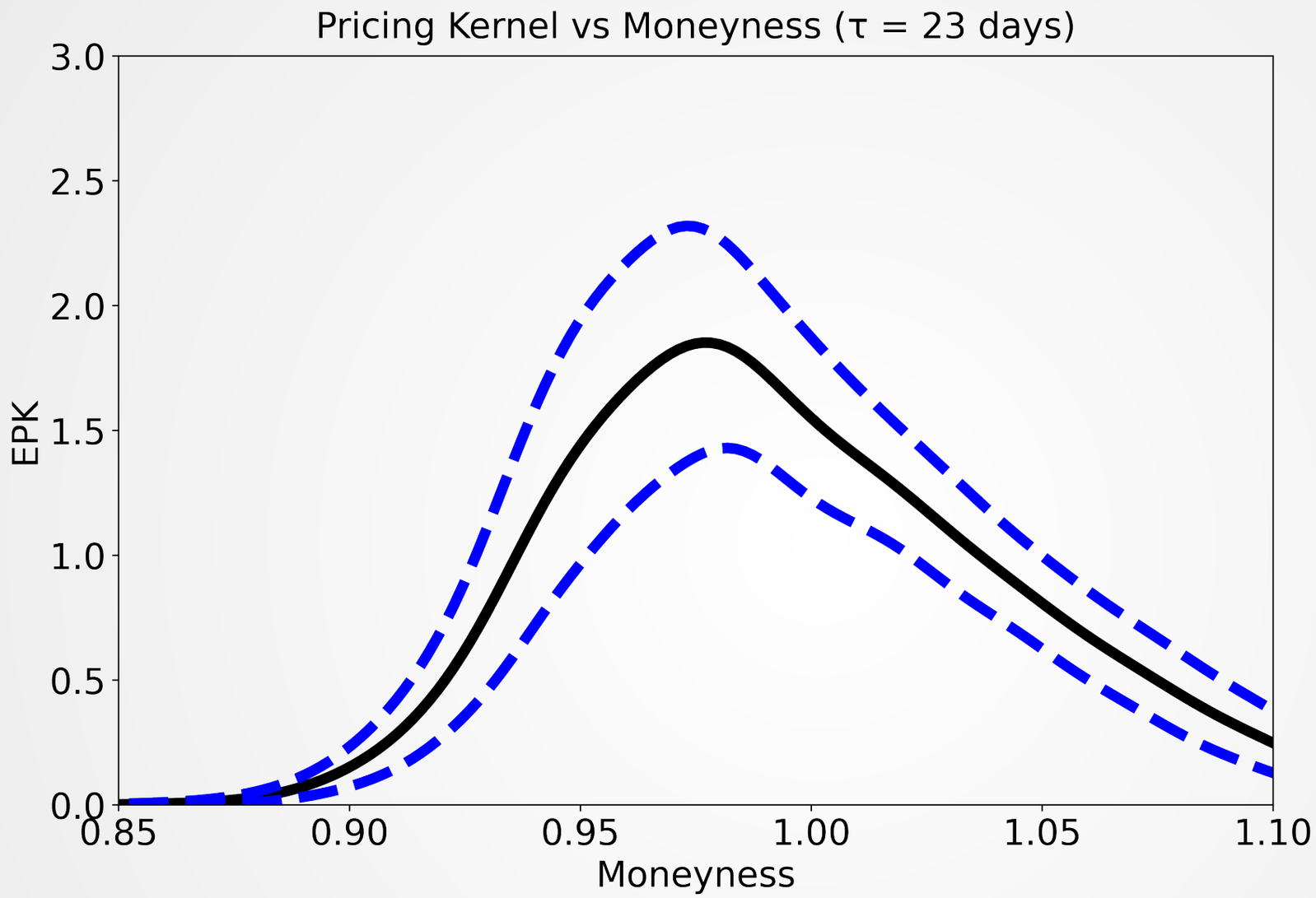
- $\sigma(t) = \begin{cases} 0.3 & \text{für } t < 2018, \\ 0.15 & \text{für } t \geq 2018. \end{cases}$

Case 2: RND and HD



- RND and HD at 2020-01-02 with
  - $r = 0.00$
  - $\tau = 29$

Case 2: Classical PK

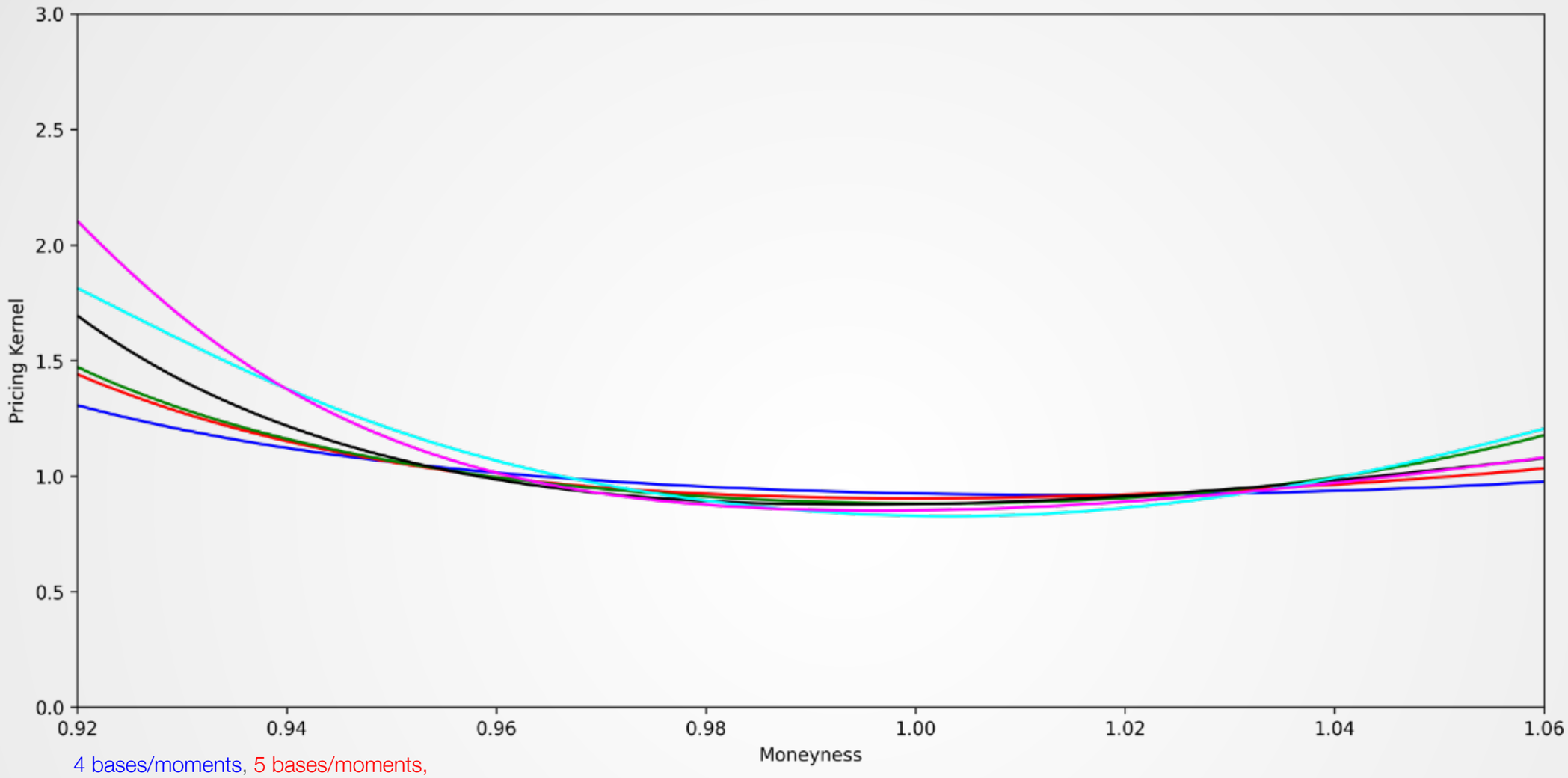


EPK,

Uniform CI Bands

- The historical density estimation is based on two distinct volatility regimes. This mixed-regime historical density significantly influences the kernel shape.

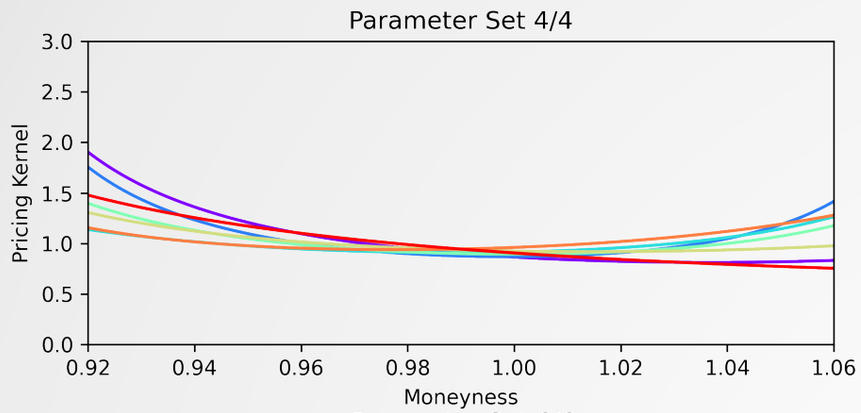
Case 2: CDI Pricing Kernel

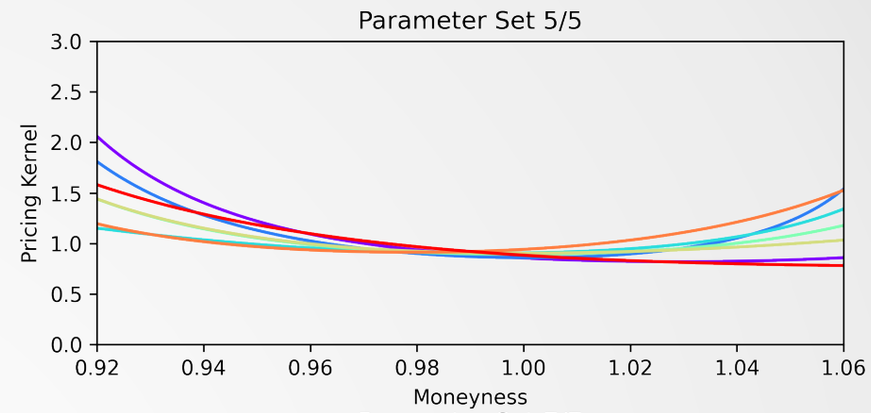


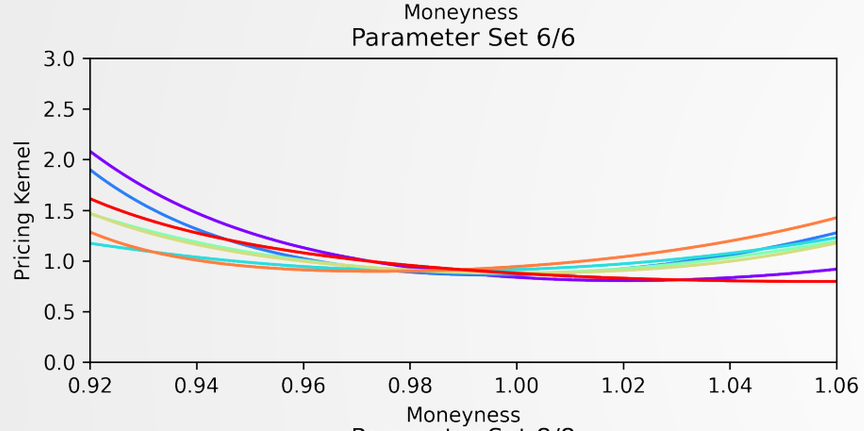
4 bases/moments, 5 bases/moments,
6 bases/moments, 7 bases/moments,
8 bases/moments, 9 bases/moments
Steps: 0.001

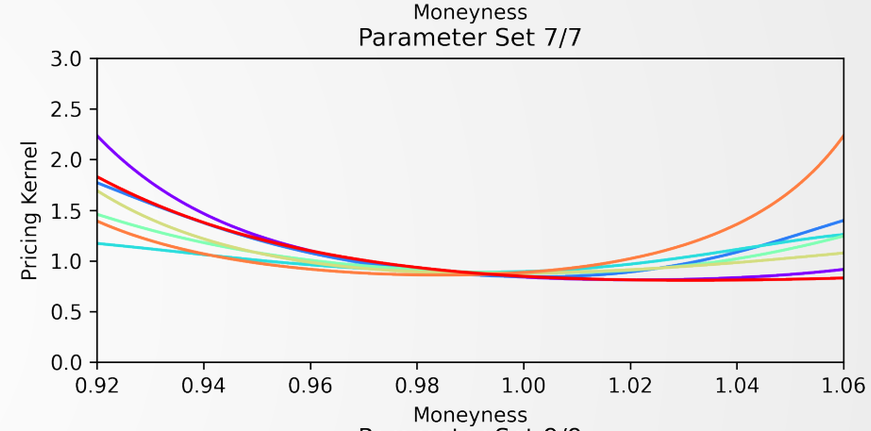
- The CDI estimates remain robust and show similarities to the baseline case.
- By design, it uses only post-2020 data, which is not interrupted by a structural break in volatility.

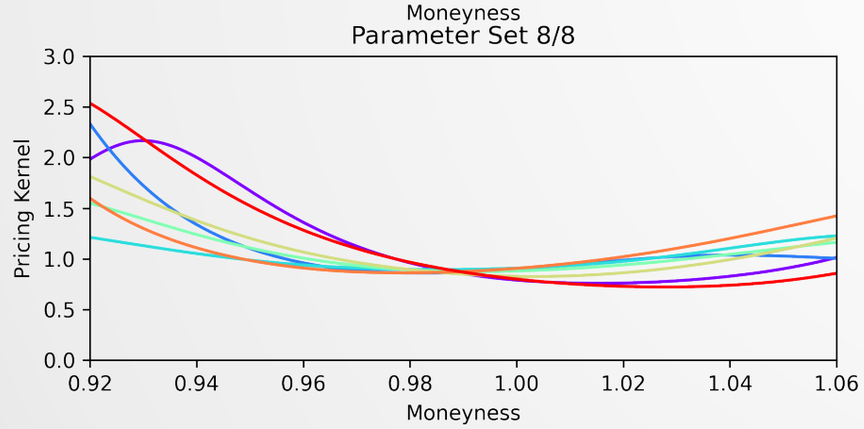
Case 2: CDI kernel variations by parameter set

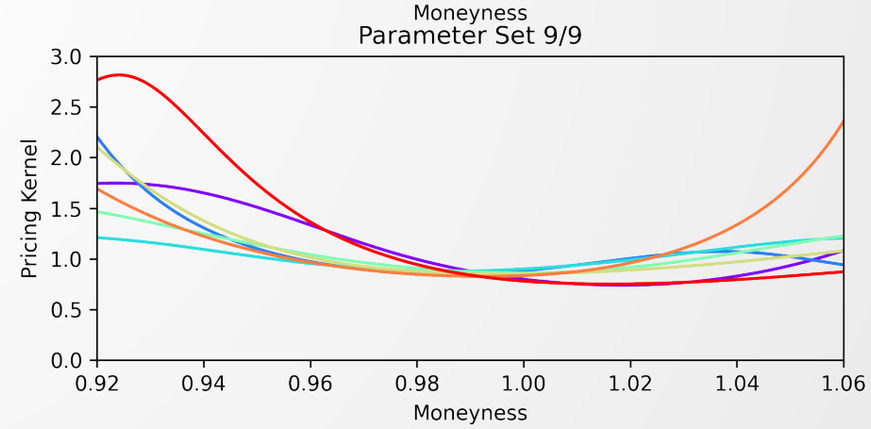






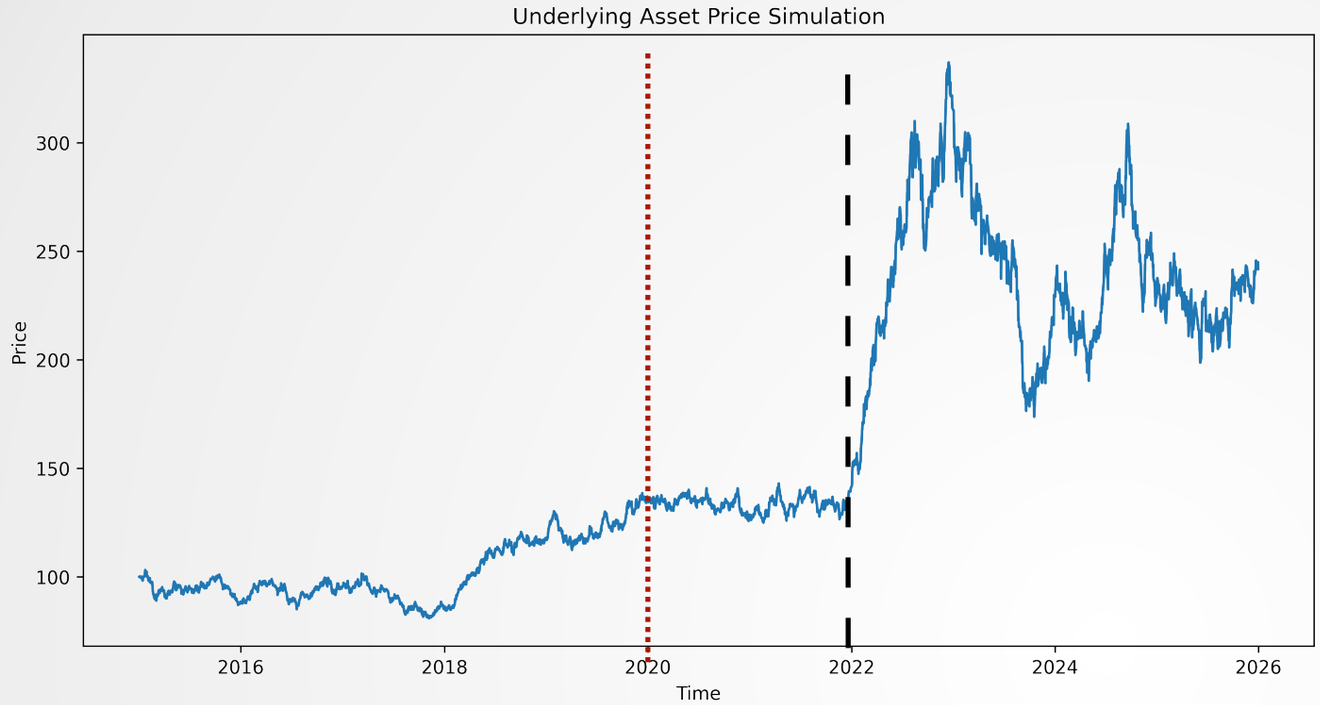




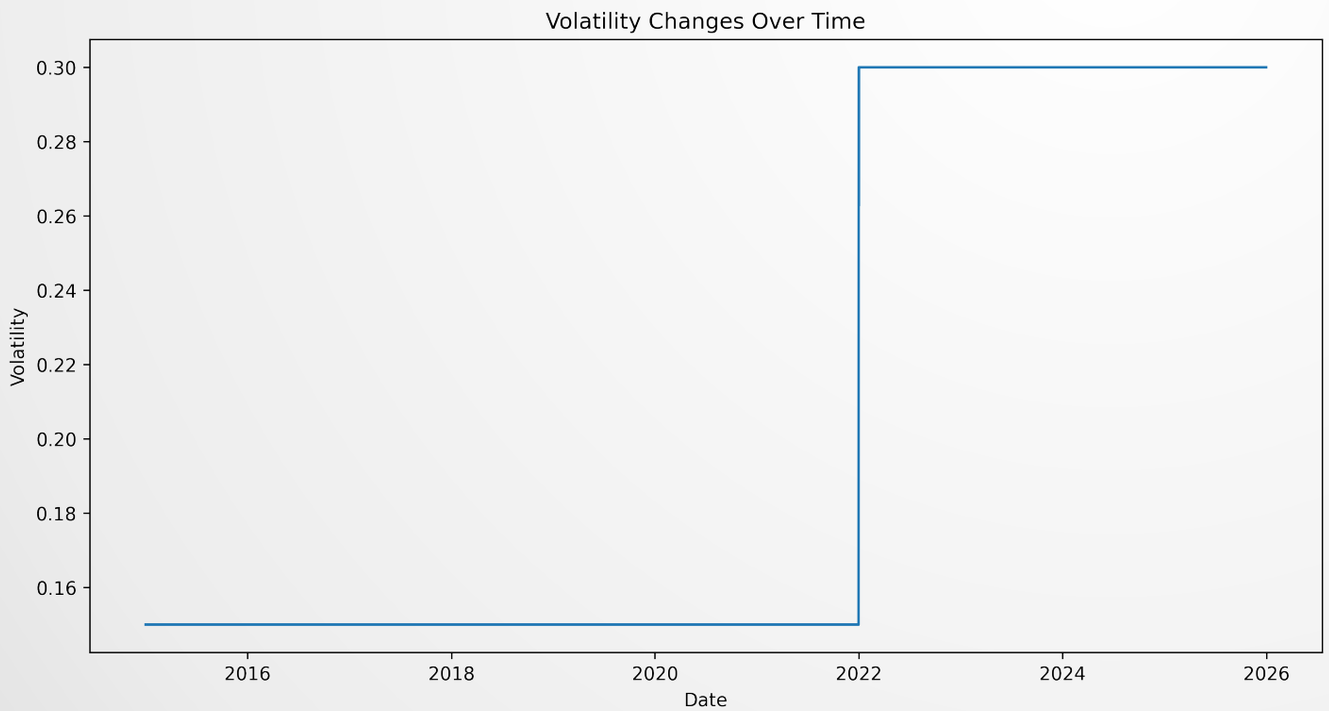


- The estimates strongly depend on the moment and basis selection, as evidenced by their behavior at the moneyness boundaries.

Case 3: Volatility jump in the "future"

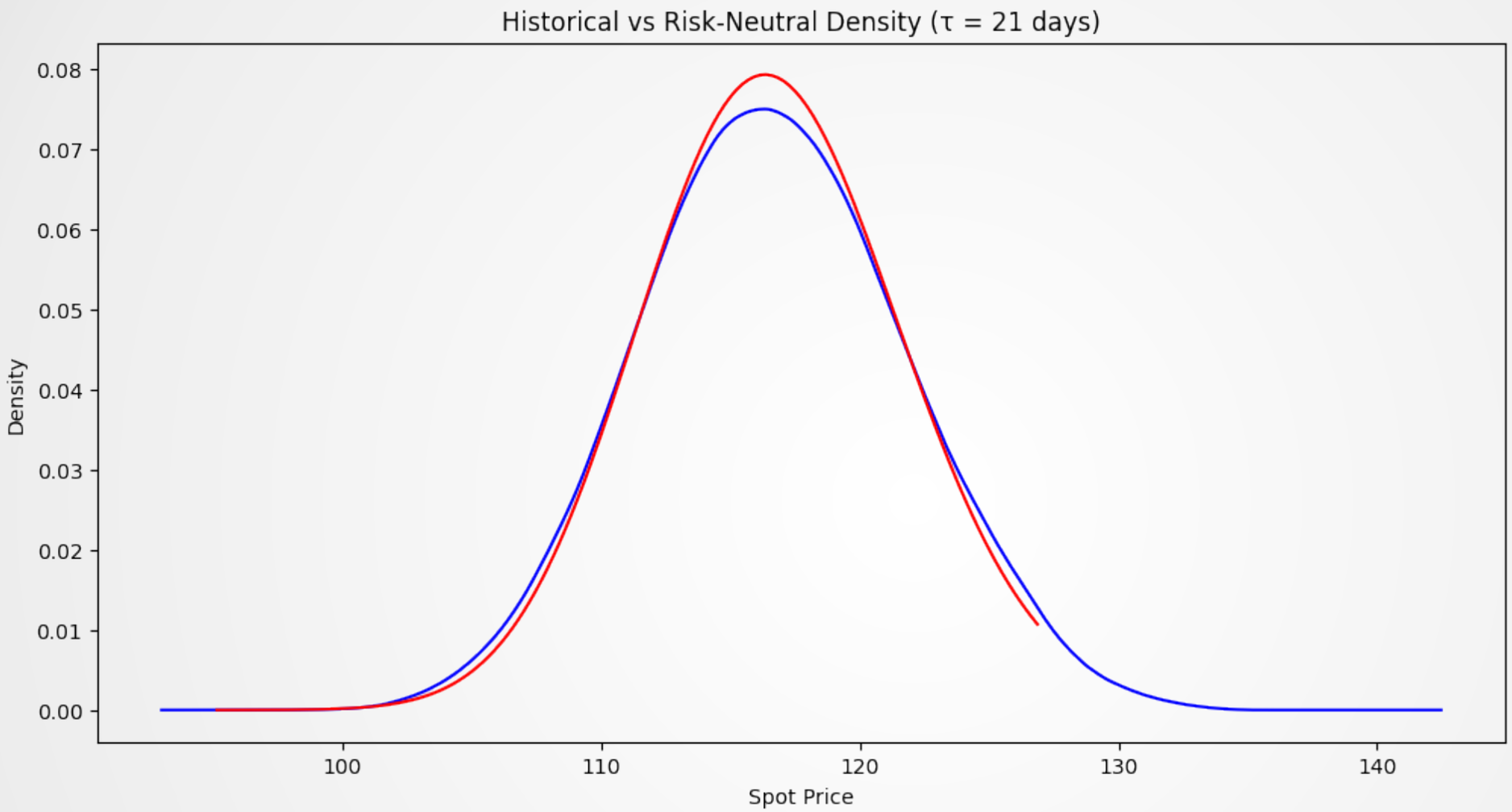


- Estimation of the PK
- Volatility jump



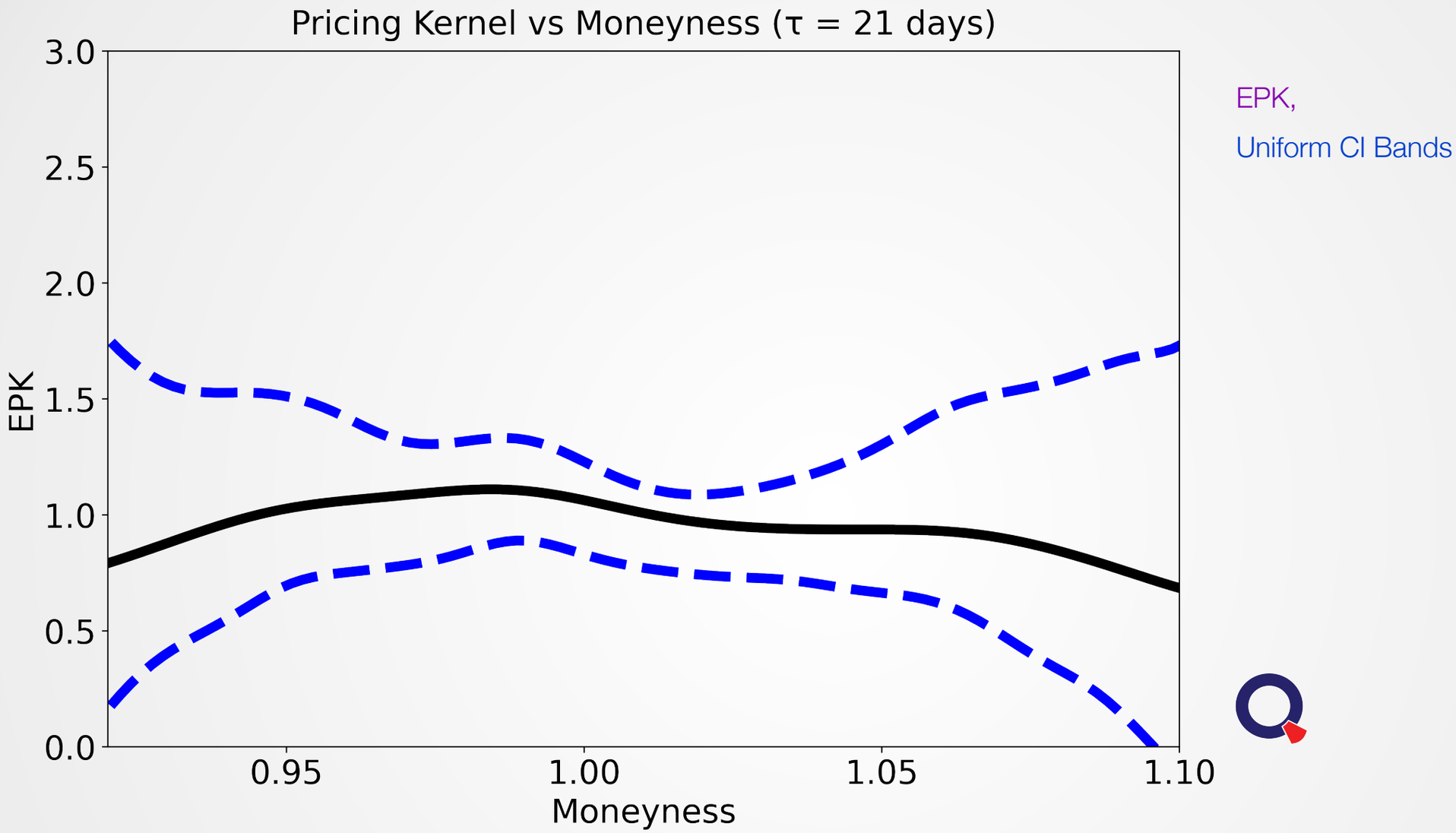
- We simulate a volatility regime shift at 2022-01-01
- $\sigma(t) = \begin{cases} 0.15 & \text{für } t < 2022, \\ 0.3 & \text{für } t \geq 2022. \end{cases}$

Case 3: RND and HD



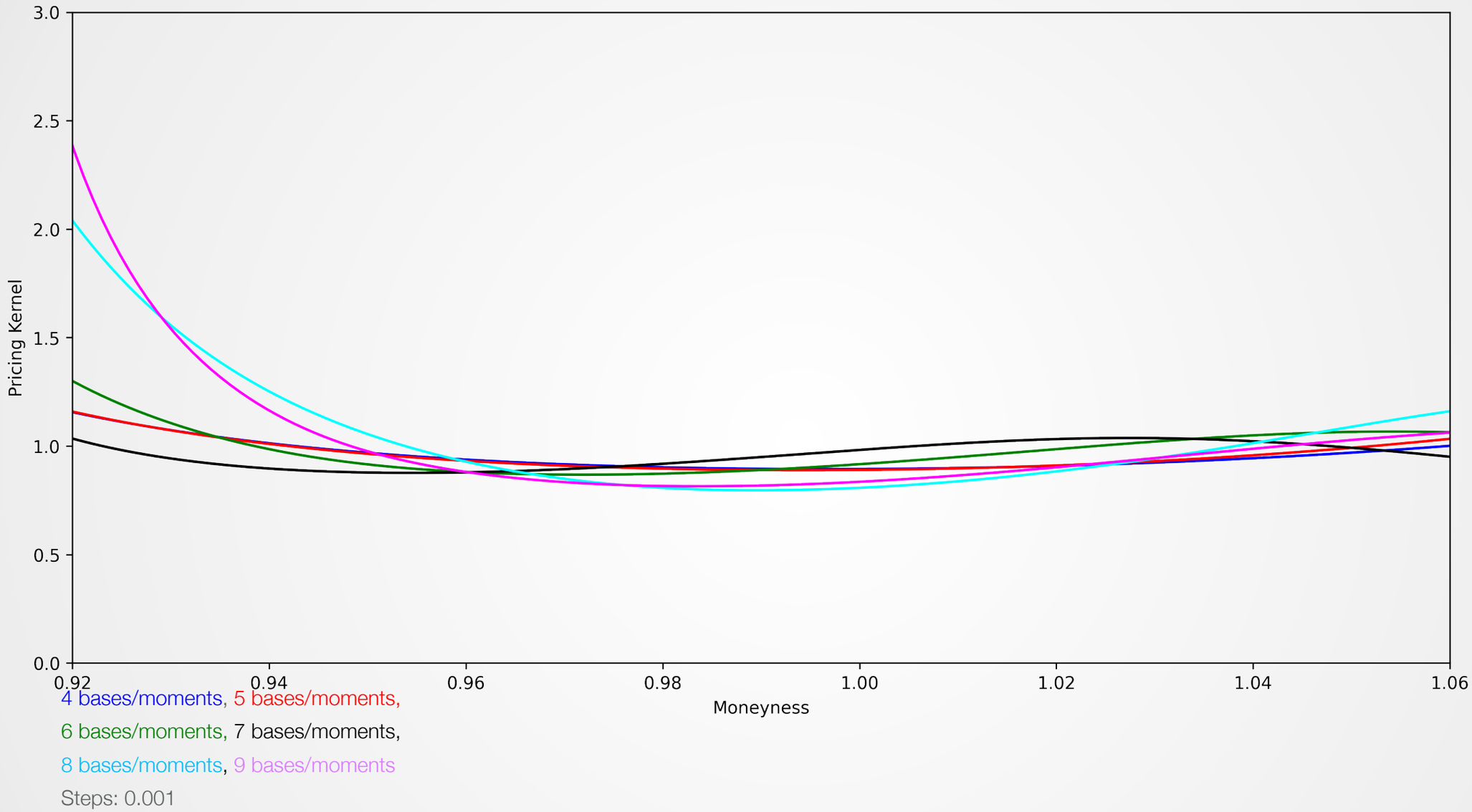
- RND and HD at 2020-01-10 with
  - $r = 0.00$
  - $\tau = 21$

Case 3: Classical Pricing Kernel



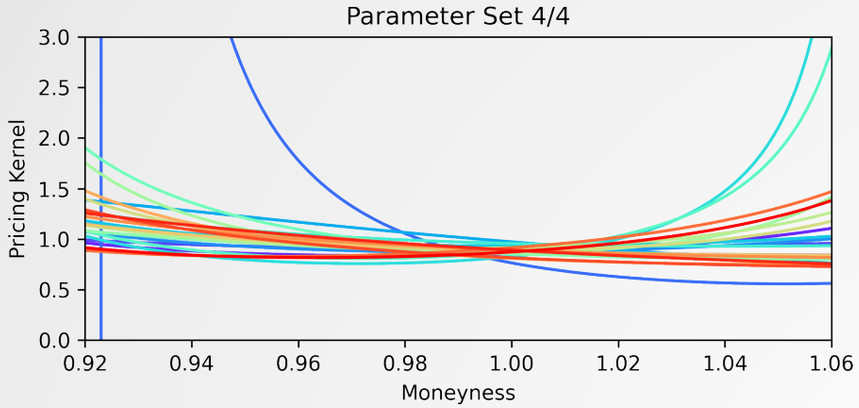
- A volatility jump in the *future* has no effect on the characteristics of the PK, as only *past* data is being used for the estimation.
- The PK is similar to the baseline kernel.

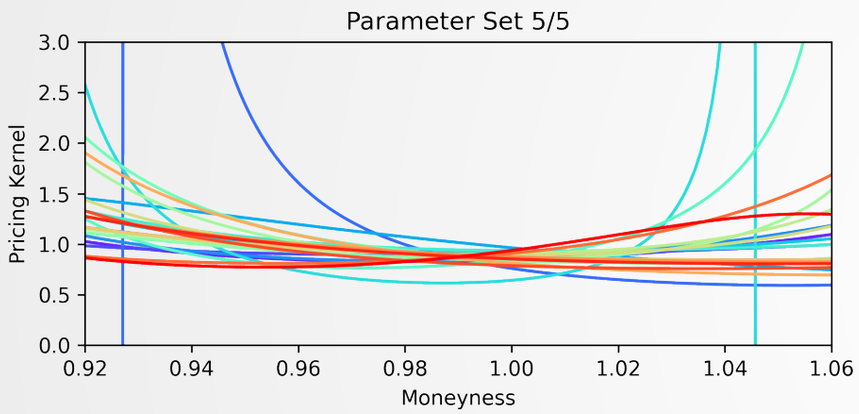
Case 3: CDI Pricing Kernel

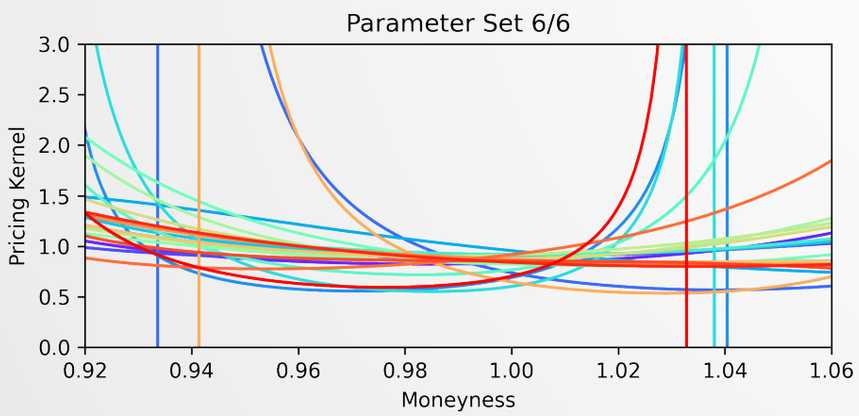


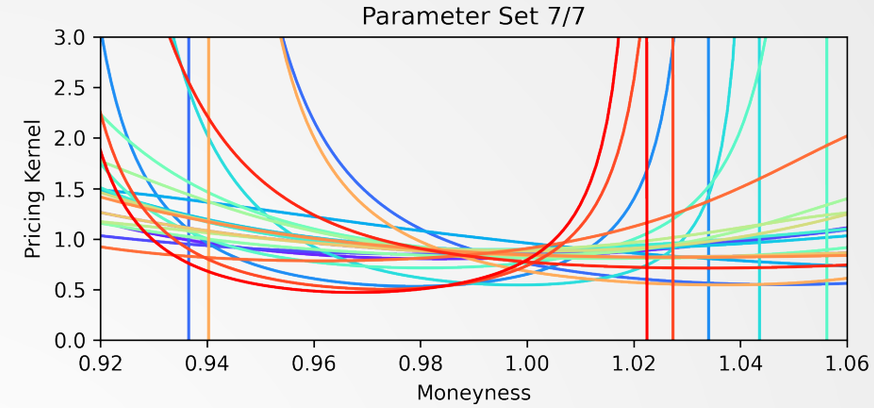
- CDI kernels have non-monotonic segments across all parameter sets

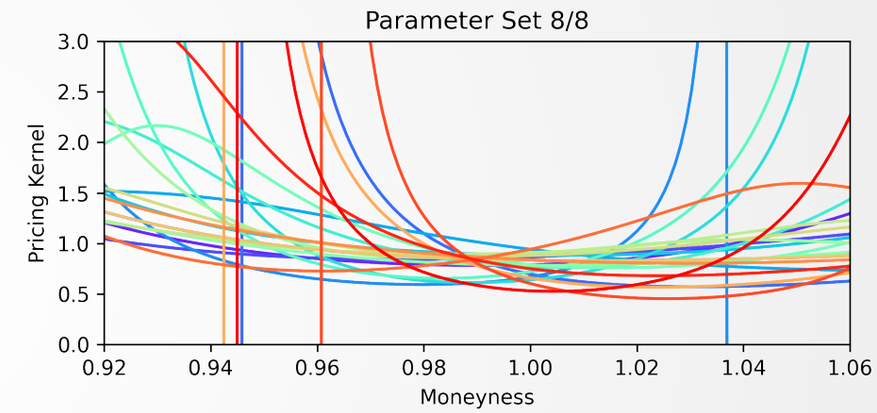
Case 3: CDI kernel variations by parameter set

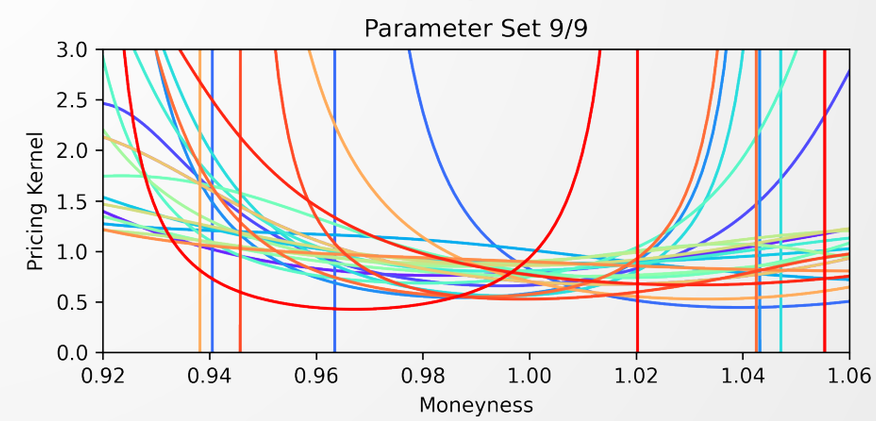






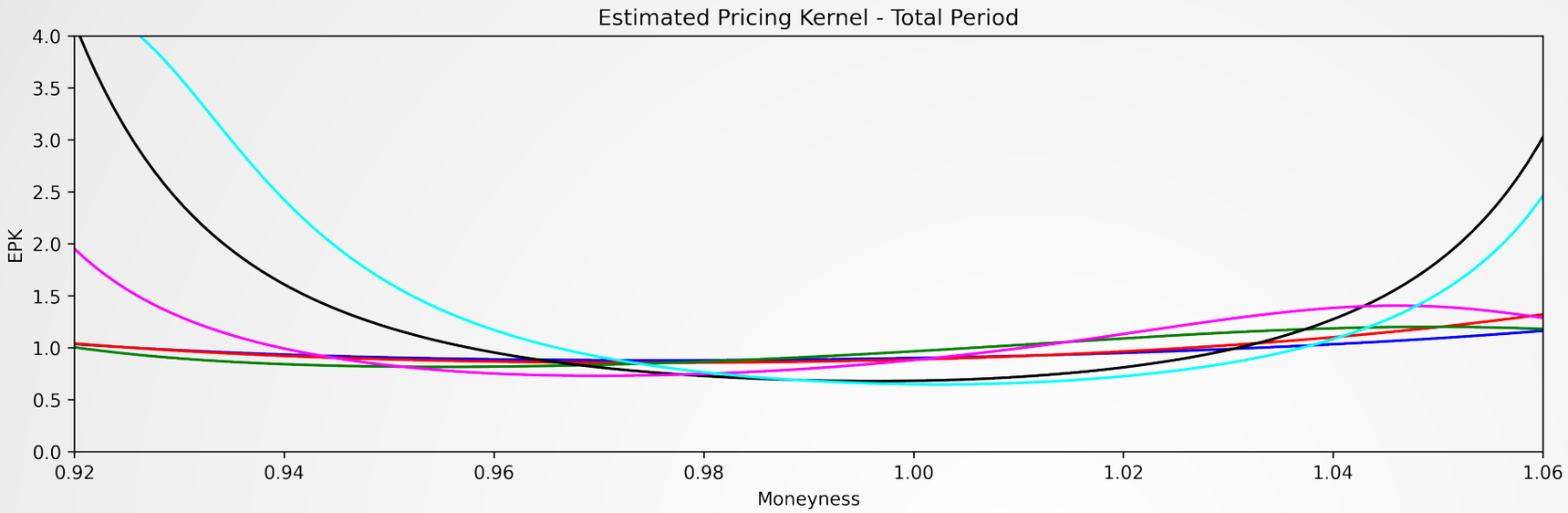




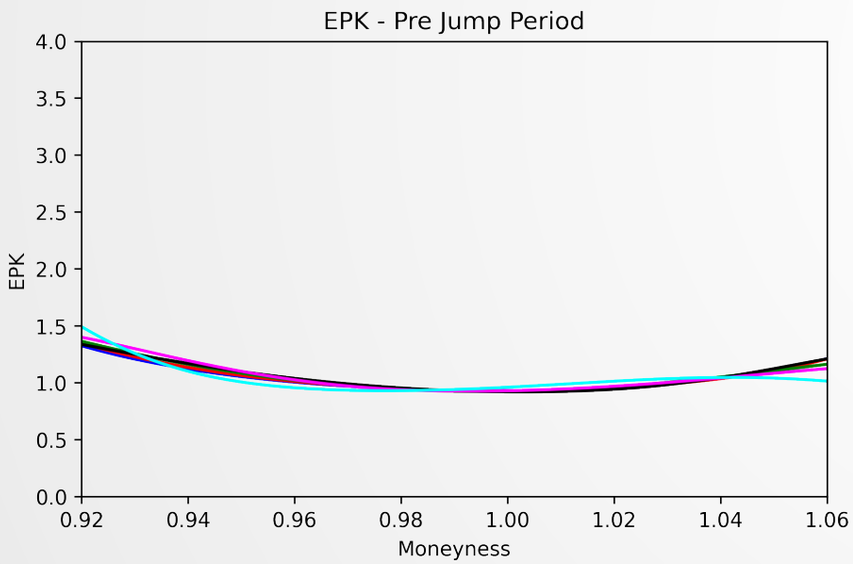


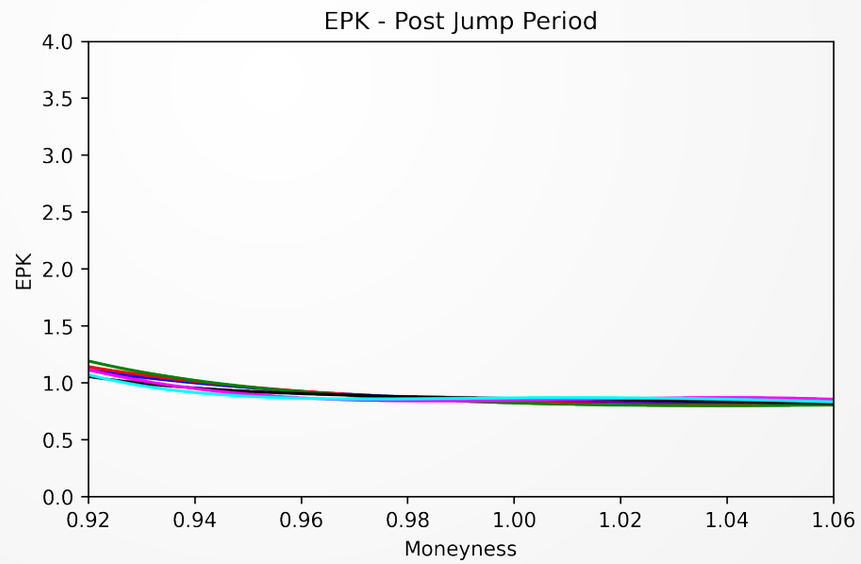
- CDI estimates show systematic failures across all parameter combinations, returing erratic, non-monotonic kernels, with increasing segments at moneyness <1.

Case 3: Large sample - Pre-/ Post-/ Total CDI kernel



4 bases/moments, 5 bases/moments,
6 bases/moments, 7 bases/moments,
8 bases/moments, 9 bases/moments
Steps: 0.001





- To investigate the sample size effect, we simulated data from 2000-2025 with a volatility jump (15% to 30%) in 2013, and estiamted the CDI kernel in 2000.
- Persistent non-monotonicity and U-shaped patterns can be observed across parameter sets.
- To isolate the impact of the volatility jump, we analyzed pre-jump (2000–2013) and post-jump (2013–2025) sub-periods. Both periods yield monotonically decreasing kernel across all parameter sets.

Result

- CDI shows significant sensitivity to future volatility regime changes.
  - Non-monotonically decreasing functions and erratic pricing kernels are observed, independent of the sample size
- This challenges economic theory and Linn et al.'s monotonicity proposition.
- CDI proofs to have severe parameter sensitivity.
  - Lower parameter sets (4–6) return monotonical kernels in stable markets. Higher parameterizations (7–9) introduce instability and non-monotonic patterns.

CDI may addresses the information set mismatch, but shares fundamental limitations with the classical approach when confronting structural changes in the market volatility.

Monotonicity proposition critique

- Monotonicity must hold simultaneously across the entire domain, Linn et al. (2018) rely on pointwise confidence intervals.
- Pointwise CIs assess local variability, not global structure.

$$P\left(m(x) \in [L_1(x), U_1(x)]\right) = 1 - \alpha$$

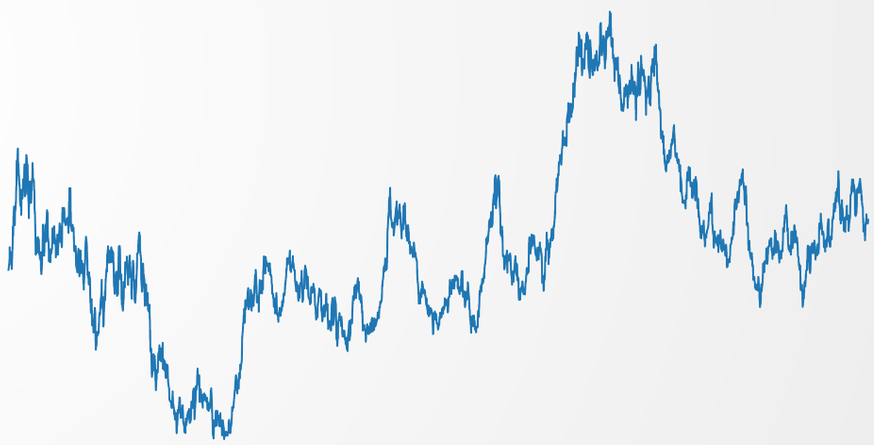
- Fails to account for joint uncertainty. Confidence bands (Härdle et al., 2004) are needed.

$$P\left(m(x) \in [L_2(x), U_2(x)] \text{ for all } x \in [a, b]\right) = 1 - \alpha$$

- Linn et al.'s claim they "cannot reject monotonicity" but lack a formal test.

Volatility Jumps and Pricing Kernel Features

Wolfgang Karl Härdle
Levin Schach



Ladislaus von Bortkiewicz Professor of Statistics
Humboldt-Universität zu Berlin
theIDA.net
ASE, UEdinburgh, NYCU 玉山學者