



Article

# The Effect of Fat Tails on Rules for Optimal Pairs Trading: Performance Implications of Regime Switching with Poisson Events

Pablo García-Risueño <sup>1,\*</sup>, Eduardo Ortas <sup>2</sup> and José M. Moneva <sup>3</sup>

<sup>1</sup> General Directorate of Investments, VidaCaixa (Group CaixaBank), C/Juan Gris 2-8, 08014 Barcelona, Spain

<sup>2</sup> Accounting and Finance Department, Faculty of Business and Public Management, University of Zaragoza, Rda. Misericordia, 1, 22001 Huesca, Spain; edortas@unizar.es

<sup>3</sup> Accounting and Finance Department, Faculty of Economics and Business, University of Zaragoza, Gran Vía de Don Santiago Ramón y Cajal, 2, 50005 Zaragoza, Spain; jmmoneva@unizar.es

\* Correspondence: risueno@unizar.es or garcia.risueno@gmail.com

**Abstract:** This study examines the impact that fat-tailed distributions of the spread residuals have on the optimal orders for pairs trading of stocks and cryptocurrencies. Using daily data from selected pairs, the spread dynamics has been modeled through a mean-reverting Ornstein–Uhlenbeck process and investigates how deviations from normality affect strategy design and profitability. Specifically, we compared four fat-tailed distributions—Lévy stable, generalized hyperbolic, Johnson’s  $S_U$ , and non-centered Student’s t—and showed how they modify optimal entry and exit thresholds, and performance metrics. The main findings reveal that the proposed pairs trading strategy correctly captures some key stylized facts of residual spreads such as large jumps, skewness, and excess Kurtosis. Interestingly, we considered regime-switching behaviors to account for structural changes in market dynamics, providing empirical evidence that optimal trading rules are regime-dependent and significantly influenced by the residual distribution’s tail behavior. Unlike conventional approaches, we optimized the entry signal and link heavy tails not only to volatility clustering but also to the nonlinearity in switching regimes. These findings suggest the need to account for distributional properties and dynamic regimes when designing robust pairs trading strategies, providing a more realistic and effective framework of these strategies in highly volatile and non-normal markets.



Academic Editor: Zied Fiti

Received: 21 April 2025

Revised: 26 May 2025

Accepted: 28 May 2025

Published: 1 June 2025

**Citation:** García-Risueño, P.; Ortas, E.; & Moneva, J. M. (2025). The Effect of Fat Tails on Rules for Optimal Pairs Trading: Performance Implications of Regime Switching with Poisson Events. *International Journal of Financial Studies*, 13(2), 96. <https://doi.org/10.3390/ijfs13020096>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** pairs trading; fat-tailed distributions; Monte Carlo simulation; regime switching

## 1. Introduction

Pairs trading is a widely used trading strategy (Bağcı & Soylu, 2024, 2025; Bergmann & de Oliveira, 2025; Liou et al., 2024; Tadi & Witzany, 2025; H. Yang & Malik, 2024; Wilkens, 2025), which relies on the assumption that a known function (i.e., the spread) of the portfolio price is a stationary, mean-reverting process (Krauss, 2016; Endres, 2020). The normality assumption when fitting probability distribution to spreads’ residuals has been neglected by previous research. In this context, fat-tailed distributions may overcome the restrictions of normal spreads in the context of pairs trading performance (Göncü & Akyıldırım, 2016; Guang, 2021). Though there is an abundance in the literature on pairs trading modeled by the Ornstein–Uhlenbeck equation, there is very little research that evaluates optimal trading rules using non-normal probability distributions (Endres & Stübinger, 2019a).

In this way, a couple of studies (Stübinger & Endres, 2018; Endres & Stübinger, 2019b), which focus on S&P500 oil stocks, have modelled the central-involved random process as

a Wiener process plus a Poisson process with variable size (amplitude) each time. They concluded that the proposed pairs trading strategy achieves a higher performance than traditional approaches (i.e., based on traditional distance and time-series models). However, they focused on high-frequency data, which is often, unfortunately, only accessible to a minority of investors due to the onerous costs of the required infrastructure. [Larsson et al. \(2013\)](#) used a similar approach—but the sizes of the changes in the case of a Poisson event had distributions given by truncated Gaussians—and concluded that there is an optimal stopping boundary that may help investors to obtain superior performance in pairs trading strategies in the presence of jumps. [Wu et al. \(2020\)](#) analyzed a pair of Chinese stocks, calculating optimal trading rules for pairs trading; they modeled returns using Poisson processes with constant jump intensity. They concluded that excess jumps in the spread have a significant impact on strategy profitability and the optimal timing for opening and closing positions. [Guang \(2021\)](#) analyzed optimal trading rules using the t-student distribution for modeling the noise of the spread. It showed the results of a few pairs of stocks (e.g., Pepsi/Coca-cola, American banks, etc.) and concluded that in almost all cases, the aforementioned approach significantly improved investment performance. [Vergara and Kristjanpoller \(2024\)](#) found optimal thresholds for pairs trading of cryptocurrencies, modeling returns (rather than residuals of a spread that follows an Ornstein–Uhlenbeck equation) with a generalized hyperbolic distribution. [Rad et al. \(2016\)](#) pointed out the fact that the spreads have fat tails. They applied the copula method using t-student copulas but did not focus on optimal trading rules systematically.

Interestingly, [Göncü and Akyildirim \(2016\)](#), which is the work more closely related to this study, focused on optimizing average profits in pairs related to commodity futures. They considered dollar-neutral pairs (instead of beta-neutral pairs) and modeled spreads using generalized hyperbolic probability density functions and subclasses of them (NIG and VG density functions). Unfortunately, they neither optimized the enter value and stop-loss nor compared the effects of normal and non-normal distributions on trading rules, and they considered no regime switching (this is, they considered no change in the parameters which characterize the random variables throughout the simulation). All these features separate their work from ours. Based on the previous research findings and to the best of our knowledge, the present article is the first study of optimal trading rules for Ornstein–Uhlenbeck-based pairs trading of stocks using fat-tailed distributions whose conclusions can be applied by most investors. In this way, this paper contributes to the existing literature in several ways. Firstly, we considered four different fat-tailed distributions (i.e., Lévy stable, generalized hyperbolic, Johnson's  $S_U$ , and non-centered t-student) as well as normal distribution, and we analyzed how well they suited the empirical spreads (note that nearly all research analyzing fat tails considers only one fat-tailed distribution). We have chosen the referred fat-tailed distributions because of their popularity; they have successfully been applied to financial problems ([Carneiro et al., 2025; Göncü & Akyildirim, 2016; Simonato, 2012; Plerou et al., 2001](#)) and because of the fact that they are easily usable within Python. The comparison between normal and fat-tailed distributions provides insights into the (low) quality of the nearly ubiquitous normality assumption. Secondly, we thoroughly analyzed optimal trading rules for all thresholds (enter value, profit-taking, and stop-loss), which is seldom found in the literature. Our analysis of regime-switching is central to these optimal rules (this analysis is also rare in the literature on optimal trading rules for pairs trading). Thirdly, we performed calculations for cryptocurrencies, which have seldom been analyzed in this context due to their relative novelty. Fourthly, we provided means (open-source code) to replicate our results and to extend our analysis to arbitrary stocks and cryptocurrencies. In summary, the work that we present is expected to overcome the limitations of previous studies, whose scope was narrower and less easily applicable.

Regime-switching is often ignored in the literature. However, recent research on pairs trading considers regime-switching models and reveals their superior performance in different markets. [Chodchuangnirun et al. \(2018\)](#) focuses on the US market and show that a Markov-switching model with generalized conditional heteroskedasticity effects (MS-GARCH) performs better than other conventional methods, such as kink and threshold models. Similarly, [Namwong et al. \(2019\)](#) focus on the stocks of the SET100 index to construct the pairs and apply a MS-GARCH model, and they conclude that this approach generates positive returns, which are higher than the return from trading in individual stocks. Focusing on the SP500 shares, [Endres and Stüberger \(2019b\)](#) implement a Lévy-driven Ornstein–Uhlenbeck MS approach and find significant results with Sharpe ratios of 3.92 after transaction costs. [Chen et al. \(2014\)](#) focus on the Dow Jones Industrial Average Index stock and propose a three-regime threshold autoregressive model with GARCH effects (TAR-GARCH) and conclude that this strategy generates positive excess returns.

This paper is structured as follows. In the next section, we present the time series used in the empirical calculations. Section 3 briefly presents our calculation methods, including stochastic modeling and trading strategies, measures of performance and risk of the analyzed strategies, and assumptions underlying this work. In Section 4, we present the main results. In the interest of thoroughness, we also analyze the out-of-sample profits of the trading strategies and the effects of Poisson events. Finally, Section 5 outlines our main conclusions. Additional information on the methods and results can be found in the Supplementary Material, which also presents an analysis of single-product trading as well as theoretical explanations to provide deeper insights.

## 2. Data Description

The details of the time series used in this study are shown in Table 1. We focused on 23 stocks that correspond to well-known large capitalization enterprises traded in efficient markets—North America and Europe—during the 2017/2025 period (20 February 2017 to 20 February 2025). For representativity purposes, stocks for all 11 GICS sectors were analyzed. We also considered four among the most traded cryptocurrencies for robustness purposes, because their properties strongly differ from conventional stocks. Data for the four analyzed cryptocurrencies were collected from 2021 to 2025 (20 February 2021 to 20 February 2025). The risk-free rates considered correspond to yields of either 1-year T-bills (US) or German 1-year govies (Ger). With the aim of making the calculation replicable, we focus on the daily time series (weekdays).

Note that in our approach the trader can only buy or sell immediately before the closing time of each day; that is, in our framework, time is a discrete variable. It might be argued that by proceeding in this way we are making an error in the Ornstein–Uhlenbeck equation due to the non-infinitesimal size of the time step, which for us is one day (and, e.g., that a correction term proportional to  $\exp\left[-\frac{\theta\Delta t}{4}\right]$  must be added to correctly reproduce the variance of the process). However, we deemed that such an interpretation would be mistaken. For stocks, most of the trading activity takes place near the opening and closing times, hence properties like liquidity do strongly depend on the time of the day. If we were using time steps lower than one day in our modeling of the time series, we should give an account of such time of the day-dependent properties, which would make our model much more complex. That would be analogous to considering seasonality when modeling time series, which can be avoided if the time separation between measurements is exactly one year.

**Table 1.** Analyzed stocks and cryptocurrencies.

Sector	Company	Ticker	Exchange
Energy	TotalEnergies	TTE	NYSE
Energy	Phillips66 Company	PSX	NYSE
Energy	BP plc	BP	NYSE
Utilities	Duke Energy Corp.	DUK	NYSE
Utilities	Sempra Energy	SRE	NYSE
Consumer staples	Mondelez International	MDLZ	NasdaqGS
Consumer staples	Monster Beverage	MNST	NasdaqGS
Consumer discretionary	Booking Holdings Inc.	BKNG	NasdaqGS
Consumer discretionary	Marriott International Inc.	MAR	NasdaqGS
Materials	Brenntag SE	BNR.DE	XETRA
Materials	UPM-Kymmene Oyj	UPM.HE	Helsinki
Industrials	Caterpillar	CAT	NYSE
Industrials	Relx plc	RELX	NYSE
Information tech.	Alphabet Inc (Class C)	GOOG	NasdaqGS
Information tech.	Intuit	INTU	NasdaqGS
Communication	The Walt Disney Company	DIS	NYSE
Communication	Verizon Communications	VZ	NYSE
Healthcare	Amgen	AMGN	NasdaqGS
Healthcare	Stryker Corporation	SYK	NYSE
Financials	Bank of America Corp.	BAC	NYSE
Financials	PNC Financial Services	PNC	NYSE
Real estate	Essex Property Trust Inc.	ESS	NYSE
Real estate	Equity residential	EQR	NYSE
Cryptocurrency	Bitcoin	BTC-USD	CCC
Cryptocurrency	Tron	TRX-USD	CCC
Cryptocurrency	Dogecoin	DOGE-USD	CCC
Cryptocurrency	Ripple	XRP-USD	CCC

This table presents concise data to identify the financial products (stocks and cryptocurrencies) whose prices are used in this research. It includes the name of the financial product, its sector, the ticker to download the time series from Yahoo Finance, and the name of the market where it is traded.

There exist two main methods for choosing candidates to form a pair (i.e., a traded long–short portfolio). The first one is to take a large set of ( $N$ ) products, calculate the spreads of all  $(N \cdot (N - 1)/2)$  pairs within this set, and choose the pairs to trade depending on their properties, like stationarity, variance of the residuals, etc. The second is based on fundamental information, i.e., on selecting candidates which have common features, like sector, country, currency, market capitalization, etc. In this paper, we followed the second method, because we found it more robust. If two stocks have similar features, it is more justified that the spread of the pair reverts to a mean. We deemed that following the first method would be prone to selecting pairs which behaved well in the past by mere coincidence (data snooping), which will less likely behave so in the future. Therefore, we constructed pairs as follows. For every GICS sector, we selected those with largest market cap which are traded either in the US or in Europe. We took between 10 and 20 stocks for each sector, as well as the 11 cryptocurrencies with highest market value. We then randomly chose the two constituents of the pair among those whose spreads presented the best stationarity properties (subject to the condition that at least one of the two stocks—or cryptocurrencies—is widely known).

Table 2 shows the main descriptive statistics, normality tests, and stationarity tests for the residuals of the pairs' spreads (the way these spreads are calculated is explained in Section 3.1.1 below). The average Kurtosis for stocks is 19.7, much higher than the value obtained by Göncü and Akyildirim (2016) for futures of commodities (11.8); the Kurtosis of cryptocurrencies is still higher. As shown in Table 2, the normality assumption is clearly rejected for all the spread residuals, which also agrees with the results of Göncü and Akyildirim (2016), and strongly supports the usage of fat-tailed distributions. Though

not explicitly presented in Table 2, three different normality tests (Anderson–Darling, Shapiro–Wilk, and Kolmogorov–Smirnov) support that the spreads do not follow a normal distribution (the  $p$ -values were below  $10^{-19}$  for all normality tests for all the analyzed pairs). Accordingly, empirical features of the spreads suggest the use of several fat-tailed distributions to better capture the empirical properties of stocks' returns and spreads. The values of the skewness and Kurtosis displayed in Table 2 also support the non-normality, making it clear that the model used in the vast majority of the literature (normality of the spread residuals) fails to appropriately suit to the observed data. Table 2 also shows the test statistics of Augmented Dickey–Fuller (ADF) tests, which support the stationarity of the spread residuals. Finally, the test statistics of the Brock, Dechert, y Scheinkman (BDS) tests (shown in the last column, in which there is a  $p$ -value of zero in all cases) confirm nonlinearity of the dynamics. Graphical representations of the pairs' spreads vs. time are included in the Supplementary Materials.

**Table 2.** Descriptive statistics of pairs' spread residuals.

Sector	Pair	Mean	St. Dev.	Skewness	Kurtosis	AD Test	ADF Test	BDS Test
Materials	BNR.DE/UPM.HE	−0.0827	0.0172	0.2027	3.6685	13.8776 ***	−3.0927 **	441
Consumer discretionary	BKNG/MAR	0.0298	0.0183	0.1164	12.5727	25.7249 ***	−3.0671 **	248
Energy	PSX/TTE	0.0227	0.0186	−1.3310	22.2154	15.4386 ***	−3.514 ***	309
Energy	BP/TTE	0.1231	0.0105	0.0321	2.7763	7.2732 ***	−2.936 **	
Finance	BAC/PNC	0.3404	0.0105	0.3419	3.9468	15.7640 ***	−3.7158 ***	315
Health care	AMGN/SYK	0.0265	0.0173	−0.1989	8.2976	28.3346 ***	−3.3702 **	420
Utilities	DUK/SRE	−0.2741	0.0110	−2.0222	28.0670	36.1474 ***	−3.0831 **	333
Information tech.	GOOG/INTU	0.0239	0.0176	0.2393	4.9907	19.8525 ***	−3.1049 **	220
Real estate	EQR/ESS	−1.4987	0.0073	0.1255	4.9983	12.5093 ***	−3.5024 ***	282
Communications	DIS/VZ	0.1415	0.0204	0.2574	6.0019	20.6775 ***	−3.1276 **	207
Consumer staples	MDLZ/MNST	−0.1857	0.0153	−0.5351	11.5898	30.6926 ***	−3.5651 ***	221
Industrials	CAT/RELX	−0.1698	0.0201	0.2509	2.9576	9.9417 ***	−3.2606 **	349
Cryptocurrencies	BTC-USD/TRX-USD	−17.5508	0.0598	1.7731	19.8861	32.0512 ***	−3.5010 ***	131
Cryptocurrencies	XRP-USD/DOGE-USD	−1.2764	0.0809	0.8088	19.9873	30.8786 ***	−3.4110 **	190

This table comprises the main descriptive statistics, normality, stationarity, and non-linearity check for the pairs' spread residuals. AD refers to the Anderson and Darling normality test statistic. ADF refers to the Augmented Dickey–Fuller stationarity test. BDS refers to the Brock, Dechert and Schreinkman non-linearity test. \*\*\* significant at 1%; \*\* significant at 5%.

### 3. Methodology

#### 3.1. Stochastic Models

In this section, we briefly present the mathematical methods behind the calculations of this paper. Detailed explanations of our methods for pairs trading can be found in the Supplementary Materials.

##### 3.1.1. Pairs Trading Modeling

Pairs trading is a trading method aimed at reducing market risk where the investor builds a portfolio consisting of just two financial products, one being of long position and the other being of short position. The pair, however, must be appropriately selected, so that a known function of the portfolio price is a stationary, mean-reverting process. One of the most popular models to account for mean-reverting processes is the Ornstein–Uhlenbeck equation (Shreve, 2004), whose continuous time form is given by the following:

$$dX(t) = (E_0(1 - \varphi) - (1 - \varphi)X(t))dt + \sigma dW(t) \quad (1)$$

where  $X(t)$  is a mean-reverting stochastic process (for us, it is the spread of a long–short portfolio, i.e.,  $X(t) := s(t)$ ) and  $E_0$ ,  $\varphi$ , and  $\sigma$  are real numbers. For simplicity reasons, we assumed that  $E_0$  and  $\varphi$  were constant and that the probability distribution of  $\sigma dW(t)$

remained unchanged in certain periods.  $E_0$  represents the mean towards  $X(t)$  tends, the mean-reversion parameter  $(1 - \varphi)$ —which is positive—represents the speed of such trend, and  $\sigma$  is the multiplicative factor of the random variable  $dW(t)$ , sometimes called the volatility of the process (Göncü & Akyildirim, 2016). Such as was used in previous research, this paper considers the stochastic process representing fat-tailed random variables (Göncü & Akyildirim, 2016; Yu et al., 2017; Endres, 2020). In this way,  $X$  is called a generalized Ornstein–Uhlenbeck process or a Lévy-driven Ornstein–Uhlenbeck process.

Our pairs trading strategy defines the *spread* of a long–short portfolio formed by two financial products which presents the mean-reverting property. Such as was used in existing research (Vidyamurthy, 2004; Zeng & Lee, 2014; Göncü & Akyildirim, 2016), we defined the *spread* as the following:

$$s(t) := \log[p^A(t)] - \gamma \log[p^B(t)] \quad (2)$$

where  $p$  indicates the price of a financial product and  $A$  and  $B$  are indices of the two stocks. The weights of both stocks forming a pair are  $\pm 1$  and  $\mp \gamma$ , thus assuming that they are infinitely divisible. Dividends are assumed to be immediately reinvested. Although some authors (Huck & Afawubo, 2014; Elliott et al., 2005) prefer defining the spread by avoiding logarithms ( $p^A(t) - \gamma p^B(t)$ ), or using standardized prices (Carrasco-Blázquez et al., 2018), this paper defines the spread such as shown in Equation (2). This definition implies that the absolute return of the spread is approximately equal to the absolute return of our portfolio (see proof in Section S.2 of the Supplementary Materials). The  $\gamma$  coefficient has a strong impact on the mean-reversion property of the spread (an inappropriate choice may make the spread non-stationary). Notwithstanding, there is no unified criterion in the literature to assign a value to  $\gamma$ . Among the possible methods to do so, we shall mention three widely used approaches: (i) dollar-neutrality; (ii) Vidyamurthy's cointegration approach (which is often beta-neutral), and (iii) the Vector Error Correction Model (VECM). The most straightforward calculation is to define *dollar-neutral* pairs, where the initial values of the long and short positions fully offset each other ( $\gamma = p^A(0)/p^B(0)$ ). If the prices ( $p^A(t), p^B(t)$ ) are divided by their initial values, then this is equivalent to setting  $\gamma = 1$ . Using dollar-neutral pairs is a popular approach (Elliott et al., 2005; Galenko et al., 2012; Göncü & Akyildirim, 2016) because it requires little funding. However, the investor is exposed to noise because the prices at just one time ( $t = 0$ ) are used to calculate  $\gamma$ . This can be avoided through variations of the dollar-neutral method which uses information on prices at several different times. Between them, we cite the definition of  $\gamma$  as the slope of the ordinary least squares linear regression of  $p^A(t)$  vs.  $p^B(t)$  (Huck & Afawubo, 2014). A third approach to calculating cointegration coefficients is the VECM (Engle & Granger, 1987). Under this approach,  $\gamma$  is computed as a two-stage least squares procedure, regressing expressions that involve both prices and returns of A and B (Carrasco-Blázquez et al., 2018; Nair, 2021).

A further way to calculate  $\gamma$  is *Vidyamurthy's cointegration approach* (Vidyamurthy, 2004). This method provides high, stable, and robust returns (alpha) (Huck & Afawubo, 2014), also in pairs trading of cryptocurrencies. It has also been stated that returns from cointegration tend to present less extreme values (i.e., a lower Kurtosis coefficient) than other methods, like distance or copula methods (Rad et al., 2016). Given the aforementioned advantages, this paper focuses on this approach to calculate the value of  $\gamma$ .

Pairs trading relies on the hope that the *spread* reverts to a *mean*. To avoid a drift that shatters mean-reversion, the spread is defined so that it is stationary. This value of  $\gamma$  in Equation (2) is chosen so that, even if the time series  $\log[p^A(t)], \log[p^B(t)]$  are not stationary, their linear combination  $s(t)$  is stationary. This is the definition of *cointegration*: two non-stationary time series are said to be cointegrated if there is a linear combination of

them that is stationary (Vidyamurthy, 2004). Stocks prices  $p^A(t)$ ,  $p^B(t)$  are frequently non-stationary because stock markets tend to be bullish in the long term; therefore, the choice of  $\gamma$  is central to the profitability of pairs trading. A common way to find an appropriate value for  $\gamma$  is from the regression of *common trends*. This is based on the assumption that each of the log-prices of Equation (2) can be deconstructed into two terms: (i) a *common trend*  $n^{A,B}(t)$ , which is non-stationary and is most often related to the evolution of markets and (ii) a *random walk* term  $\varepsilon^{A,B}(t)$ , which is stationary and idiosyncratic (that is, it is unrelated to the general behavior of markets):

$$\log[p^A(t)] = n^A(t) + \varepsilon^A(t) \quad (3a)$$

$$\log[p^B(t)] = n^B(t) + \varepsilon^B(t) \quad (3b)$$

To guarantee cointegration, we needed the common trends to be proportional ( $n^B(t) \propto n^A(t)$ ). Indeed, their proportionality constant is  $\gamma$  (this is  $n^A(t) = \gamma n^B(t)$ ), which is called the *cointegration coefficient*. In the simplest case, the common trends are proportional to a single observable quantity that accounts for the market behavior (e.g., an index like the S&P 500 or the price of a share of an ETF which tracks it). Let us call  $R^I$  (for  $I = A, B$ ) to the log-returns of the stock prices ( $R^I(t) := \log[p^I(t)] - \log[p^I(t-1)]$ ),  $R_{mkt}$  to the market returns (e.g.,  $R_{mkt}(t) := \log[p^{ETF}(t)] - \log[p^{ETF}(t-1)]$ , where  $p^{ETF}$  is the price of the aforementioned share of an ETF that tracks the S&P 500 index), and  $R_f$  to a risk-free rate (e.g., the yield of the T-bills with maturity in one year).  $R_{mkt}$  corresponds to the returns of the indices displayed in the fourth column of the table presented in Section S.2 of the Supplementary Materials (for the cases with one single risk factor). We can then express the returns of stocks  $A, B$  as follows:

$$R^A(t) - R_f(t) = \beta^A(R_{mkt}(t) - R_f(t)) + r_{spec}^A(t) \quad (4a)$$

$$R^B(t) - R_f(t) = \beta^B(R_{mkt}(t) - R_f(t)) + r_{spec}^B(t) \quad (4b)$$

where  $\beta^I$  are the slopes of the linear regression of  $(R^I(t) - R_f(t))$  vs.  $(R_{mkt}(t) - R_f(t))$  (as given by the well-known Capital Asset Pricing Model, CAPM) and  $r_{spec}^I$  indicate returns which are specific (idiosyncratic) of the corresponding time series. Here we have considered a risk-free rate following the procedure indicated by Do et al. (2006), yet many authors simply assume that  $R_f = 0$ . Imposing  $\gamma = \beta^A / \beta^B$  in Equation (2) implies that the absolute returns of the spread ( $s(t) - s(t-1)$ ) will be the following:

$$s(t) - s(t-1) = (1 - \gamma)R_f(t) + r_{spec}^A(t) + r_{spec}^B(t) \quad (5)$$

where the first term of the right-hand side is expected to have a low size because frequently  $\gamma \simeq 1$ ; this is the absolute returns of the spread lack market (non-stationary) components and will hence depend on idiosyncratic components, which are assumed to be stationary. Note that this does not mean that our portfolio is perfectly beta-neutral: When we enter our position, we are long (short) one dollar in stock  $A$  and short (long)  $\gamma$  dollars (or other currency) in stock  $B$ . Since  $\gamma$  is the quotient of betas, at inception our portfolio is beta-neutral; however, after the first day the values of our two positions have varied, and their quotient will most likely not remain equal to  $\gamma$ . Hence, our portfolio will have a non-zero exposure to the market, even if the spread is perfectly beta-neutral (beta-neutral portfolios would require continuous rebalancing).

Since there is no guarantee that  $r_{spec}^A(t)$ ,  $r_{spec}^B(t)$  are indeed stationary, we performed stationarity tests of the spreads (see Table 2), and those pairs whose spreads fail to pass the test are discarded.

If we considered several risk factors instead of only one, that is, if we rely on the Arbitrage Pricing Theory (APT instead of CAPM), then Equation (4) becomes the following:

$$R^A(t) - R_f(t) = \sum_{i=1}^{N_{RF}} \beta_i^A (R_i(t) - R_f(t)) + r_{spec}^A(t) \quad (6a)$$

$$R^B(t) - R_f(t) = \sum_{i=1}^{N_{RF}} \beta_i^B (R_i(t) - R_f(t)) + r_{spec}^B(t) \quad (6b)$$

where  $N_{RF}$  is the number of considered risk factors (e.g., stocks, cryptocurrencies, or commodities indexes) and  $R_i(t)$  are their respective returns (see the returns of the indices displayed in the fourth column of the table presented in Section S.2 of the Supplementary Materials). In this case,  $\gamma$  would be the slope of the linear regression ( $y$ -vs.- $x$ ) of the cloud of points  $(x, y)$  where  $x = \sum_{i=1}^{N_{RF}} \beta_i^B (R_i(t) - R_f(t))$  and  $y = \sum_{i=1}^{N_{RF}} \beta_i^A (R_i(t) - R_f(t))$ .

Note that for simplicity's sake we consider a constant value for the betas in Equations (4) and (6); an interesting discussion about the stationarity of the betas can be found in [Agrawal and Clark \(2007\)](#).

In order to obtain the parameters of the Ornstein–Uhlenbeck equation ( $E_0$ ,  $\varphi$ , and  $\sigma$  from Equation (1)), we performed a discretization such as indicated by [Göncü and Akyıldırım \(2016\)](#). We made  $dt \rightarrow \Delta t := 1$  and  $dX(t) \rightarrow X(t) - X(t - 1) := s(t) - s(t - 1)$ , and we renamed  $dW(t) \rightarrow \Delta W(t)$  as  $\varepsilon(t)$ . Equation (1) hence becomes the following:

$$s(t) = (1 - \varphi)E_0 + \varphi s(t - 1) + \sigma \varepsilon(t - 1) \quad (7)$$

where  $\sigma \varepsilon(t - 1)$  are the *residuals*, being  $\sigma$  their scaling parameter. We will later fit these residuals to different fat-tailed distributions. The ordinary least squares linear regression of  $s(t)$  (i.e.,  $y$ ) vs.  $s(t - 1)$  (i.e.,  $x$ ) provides the  $E_0$  and  $\varphi$  parameters. The residuals of the regression ( $\sigma \varepsilon$ ) can be fit to any probability distribution which is deemed appropriate. If the analyzed data present the appropriate properties (i.e., stationarity of the spread and low autocorrelation of residuals), then the selected pair can be traded with profit expectations. To evaluate the optimal rules for such trading, we set a maximum horizon, which is defined as the maximum number of days that a portfolio is held before selling. We then performed a Monte Carlo (MC) analysis for different values of the thresholds (trading rules) that determine the strategy. These are the following: (i) the *enter* threshold, which determines the minimum absolute value of the spread to build a position (i.e., to enter the trade); (ii) the *profit-taking* threshold, which determines the value of the spread to unwind (finish) the portfolio making a profit; and (iii) the *stop-loss* threshold, which determines the value at which the portfolio is unwound having a loss.

The products which form the chosen long–short pairs have important features in common: they belong to the same GICS sector and to the same region, are traded in the same market with the same currency, as well as similar regulations, and both have large market capitalizations. Hence, they are expected to depend on the same economic forces. This supports the expectation that, when the spread ( $X$  in Equation (1)) moves far away from the mean ( $E_0$ ), such market forces will move it again towards this mean.

### 3.1.2. Probability Distributions

Each pair spread is fitted to five different distributions: (i) normal, (ii) Lévy stable, (iii) generalized hyperbolic, (iv) Johnson's  $S_U$ , and (v) non-centered t-student. All distributions have a *loc* and a *scale* parameter, which account for the location and scale (size) of the

residuals. The four fat-tailed distributions also have two parameters each to account for the skewness and fatness of the tail. The generalized hyperbolic distributions that we have considered also have a fifth parameter (*p-parameter*) which accounts for the form. We called  $\mathbf{p}_d$  the set of two, four, or five parameters of the distribution plus the parameters of the Ornstein–Uhlenbeck equation<sup>1</sup>. For the normal distributions, the loc and scale parameters are the mean and standard deviation of the spread residuals. The former is always zero (by consequence of the definition of ordinary least squares linear regression), and hence the only actual parameter of the distribution is the scale ( $\varphi$  and  $E_0$  are found by simple ordinary least squares linear regression of  $s(t)$  vs.  $s(t - 1)$ ). The parameters of the four fat-tailed distributions are calculated by maximum likelihood (ML), which is the standard method to this end (Borak et al., 2011). The calculation of the parameters through ML is made by minimizing the loss function:

$$\text{loss}(\mathbf{p}_d) := -\frac{1}{N_r} \sum_{n=1}^{N_r} \ln[\text{pdf}(\mathbf{p}_d)] \quad (8)$$

where  $N_r$  stands for the number of residuals and  $\text{pdf}$  is the probability density function. We calculated the optimal parameters of the distribution (i.e., those minimizing the loss function), by using the gradient descent method (Barzilai & Borwein, 1988). This method guarantees that a local minimum of the loss function is eventually reached. Apart from performing gradient descent iterations that involve all parameters, in order to minimize the loss function we performed a series of iterations that involved only the skewness parameter or the tail parameter. This was because the gradient with respect to the scale parameter is often much higher than the rest, which can make the skewness and tail parameters move very little away from their initial values unless such individual iterations are performed. In addition, to find appropriate solutions we included some randomness in the evolution, and we tried different random and deterministic starting values for the gradient descent method, except with the Lévy stable distribution. This was because calculations of the pdf of the stable distribution are much slower than for non-centered t-student and generalized hyperbolic, which precluded promptly, performing many calculations using the stable distribution. The slowness of maximum likelihood fittings to Lévy stable distributions has already been pointed out by other authors (Kawai, 2012).

### 3.2. Trading Strategies

We searched for optimal trading strategies using synthetic data, due to their advantages (López de Prado, 2018). That is, for a given value of the set of trading rules, we generated numerous (200,000) time series. Each of these synthetic time series is one path of our Monte Carlo simulations and corresponds to the evolution of  $s(t)$  of Equation (7). The time series of  $s(t)$  are generated with up to five different probability distributions for  $\sigma\varepsilon$ : normal, Lévy stable, generalized hyperbolic, Johnson's  $S_U$ , and non-centered t-student. The trading rules consist of four numbers, which determine the thresholds for entering and exiting (unwinding) our positions. The *enter value* is the value of the spread  $s(t)$  at which we built our long–short portfolio (i.e., we bought product A and sold product B or vice versa). The limit orders are the *profit-taking* and *stop-loss* thresholds. If our position is *long in spread*, i.e., if the enter value is negative (we define it by subtracting  $E_0$ ) because we bet that the spread will increase, then the profit-taking is a value of the spread which is above the enter value, and the stop-loss is a value of the spread which is below the enter value. The converse happens if our position is *short in spread* (i.e., if the enter value is positive because we bet that the spread will decrease); in this case the profit-taking is below the enter value, and the stop-loss is above the enter value. As soon as either the profit-taking or the stop-loss threshold is exceeded, we determined that the portfolio is unwound, obtaining a profit (if

the profit-taking was surpassed) or a loss (if the stop-loss was exceeded). The last trading rule is the *maximum horizon*, which determines the uppermost time for holding a pair. If after such a period holding our pair we have reached neither the profit-taking nor the stop-loss thresholds, then the position is unwound anyway (at a profit or a loss). We did not make calculations to optimize the maximum horizon because we assumed that the trader may have had exogenous limitations for it. Therefore, each of the (e.g., 200,000) synthetic paths provides a profit (or loss), which is defined as the difference between the spread when unwinding the position minus the spread when the position was built (times  $-1$  if we were short in spread). We considered transaction costs equal to 0.5% per year for shorted stocks and 1.2% for shorted cryptocurrencies. This is a proxy for fees and commissions and slippage and market impact (see the Supplementary Materials for a discussion about this choice).

When we tested each set of parameters corresponding to a given trading strategy (horizon and enter, profit-taking, and stop-loss thresholds), we used the same set of generated random variables which account for the evolution of prices. This is aimed at avoiding statistical error (Bouchaud & Potters, 2003; Göncü & Akyildirim, 2016). Otherwise, the differences between the synthetic price time series may lead to distorted results. For simplicity's sake, we assumed that, at the initial time ( $t = 0$ ) of each MC path, the spread lies exactly at  $E_0$ . The generation of synthetic time series allows mitigating the *backtest overfitting*, which is a major common drawback of quantitative investing modeling (López de Prado, 2023). Making an analysis based solely on the observed price time series would give excessive weight to the realized values of the random variable, altogether neglecting other values, that were equally likely, but not realized.

### 3.3. Performance Measures

For each set of trading rules (enter value, profit-taking, stop-loss, and maximum horizon), our software calculates eight measures of performance and risk: the Sharpe ratio of profits ( $SR$ ), the rescaled Sharpe ratio calculated from semi-deviation ( $SR'$ ), standard deviation, semi-deviation, profit average, probability of loss, Value-at-Risk (VaR), and Expected Shortfall (ES). In this paper, we focused on the first two to account for the risk-adjusted returns, but choosing other criteria is a matter of taste. For example, the *hierarchical risk-parity method* (López de Prado, 2016) focuses on standard deviations of returns, and other authors prefer to simply use the expected profit to measure performance (Göncü & Akyildirim, 2016).

We define the profit average, the profit standard deviation, and the Sharpe ratio as the following:

$$\mu_r := \frac{1}{N_p} \sum_{i=1}^{N_p} r_i \quad (9)$$

$$\sigma_r := \sqrt{\frac{1}{N_{iter}} \sum_{i=1}^{N_p} (r_i - \mu_r)^2} \quad (10)$$

$$SR_r := \frac{\mu_r}{\sigma_r} \quad (11)$$

where  $r_i$  stands for the annualized cumulative return of the  $i$ -th Monte Carlo path (note that a path can contain several enter and exit points; the cumulative return is the addition of all of them) and  $N_p$  is the number of paths. For a given Monte Carlo path, we annualize its total return ( $r_i$ ) by multiplying it by the number of trading days of a year and dividing it by the time span of the path. In the analysis of pairs trading, we define each individual return as  $r := s(t_{out}) - s(t_{in})$ , where  $t_{out}$  is the time of closing (unwinding) our position and  $t_{in}$  is the time of building it. In the analysis of trading a single product (presented in

the Supplementary Materials), we define the return as  $r := p(t_{out}) - 1$  for calculations of a single product ( $p(t_{in}) := 1$ ).

We define the semi-deviation and its corresponding Sharpe ratio as follows:

$$\sigma'_r := \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\max[(\mu_r - r_i), 0])^2} \quad (12)$$

$$SR'_r := \frac{\mu_r}{\sqrt{2}\sigma'_r} \quad (13)$$

where only the returns that are below average are counted measuring the variability. The rescaling factor ( $\sqrt{2}$ ) is included to provide a more understandable comparison between  $SR_r$  and  $SR'_r$ ; this term would make both Sharpe ratios equal if the returns were fully symmetric. We deem the *rescaled Sharpe ratio from semi-deviation* ( $SR'_r$ ) a more reliable measure of performance than the *bare* Sharpe ratio ( $SR_r$ ) because it penalizes the variability of below-average profits, which better suits the loss aversion of investors (Thaler, 2015). Therefore, we present  $SR'$  in most of the plots (heatmaps) of this article. Note that  $SR'$  is conceptually similar to the Sortino ratio because its denominator is also a semi-deviation.

We define the ES, such as indicated by the European Banking Authority, though keeping a negative sign for losses; that is the following:

$$ES = \frac{1}{\alpha N_p} \left[ \left( \sum_{i=1}^{\lfloor \alpha N_p \rfloor} r_i \right) + (\alpha N_p - \lfloor \alpha N_p \rfloor) r_{\lfloor \alpha N_p \rfloor + 1} \right] \quad (14)$$

where the  $\lfloor x \rfloor$  signs indicate the integer part (floor) of  $x$  and the indices ( $i$ ) of  $r_i$  are ordered so that  $r_i$  are monotonically increasing. For both VaR and ES, we set  $\alpha = 0.01$  (1%). Note that Equation (14) is based on Historical Simulation, which makes that formula potentially inaccurate (García-Risueño, 2025); however, we used it because its usage is widely extended due to regulatory requirements. An interesting discussion about the properties of the *ES* can be found in Gribkova et al. (2025).

In the Supplementary Materials, we also provide examples of the VaR and the probability of loss. We define the VaR using the historical method (Choudhry, 2013), that is as the  $(\alpha N_p)$ -th worst return. We define the probability of loss  $P_{loss}$  as the number of Monte Carlo paths whose return is negatively divided by the total number of paths; this is  $P_{loss} = \left( \sum_{i=1}^{N_p} 1_{r_i < 0} \right) / N_p$ .

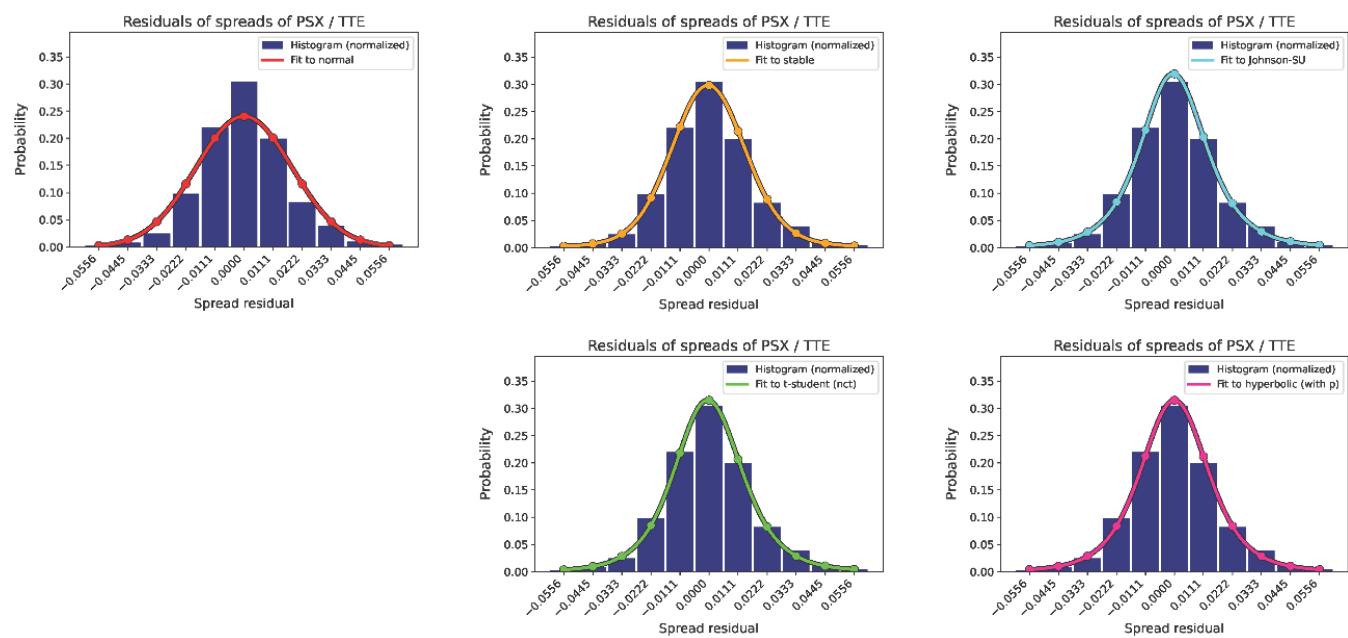
## 4. Results and Discussion

In this section, we present the calculated features of pairs trading. Complementary remarks on single-product trading can be viewed in the Supplementary Materials.

### 4.1. Fitting of Probability Distributions

We calculated spreads using the method presented in Section 3.1.1 (see Equation (2)), using pairs of the products listed in Table 1; the list of pairs is displayed in Table S1 in the Supplementary Materials. The spreads of the 13 analyzed pairs pass the stationarity test (see Table 2). The normality of the spread residuals was checked through the three normality tests, including Anderson–Darling’s (see Table 2). The results reveal that all the analyzed spreads are non-normally distributed. Then, the analyzed spreads were fitted to the normal and to four fat-tailed distributions. As an example, Figure 1 presents the histograms of the spreads (blue bars) for an example spread (pair of energy stocks, PSX/TTE). In each subplot, the curve represents the best fit found for a given distribution. As expected, the normal distribution fits much worse to the histogram than the fat-tailed distributions. This property holds for all the analyzed pairs (see the rest of the figures in the Supplementary Material).

The subplots of our Figure 1 that correspond to normal and to generalized hyperbolic distributions present the same behaviors reported by Figures 2 and 4–6 of Göncü and Akyıldırım (2016).



**Figure 1.** Fitted spreads to the probability distributions (TTE/PSX example). The blue bars present the histogram of the spread residuals. The curves in each subplot correspond to the best-found fitting to it for each of the five analyzed probability distributions. The normal probability density function (red curve) fits to the histogram much worse than the fat-tailed density functions (yellow, blue, green, and pink curves).

Note that, for the analyzed periods, all the spreads listed in Table 2 (except BP/TTE) qualitatively display oscillation around one single value of the mean (see Figures S1–S3 in the Supplemental Materials), while the BP/TTE spread displays several different values around which the spread oscillates (see Figure 4 below in this document). Therefore, we fitted the spread residuals of every pair (except BP/TTE) to a unimodal distribution. Concerning BP/TTE, we fitted the spreads of different time windows to different unimodal distributions.

In Tables 3–7, we showed the fitting parameters of the spread residuals. These values clearly reveal that the tails of the residuals' distribution are notably fat; for example, the average value of the number of degrees of freedom ( $N_{df}$ ) of t-student distributions is 3.84 for stocks and 2.06 for cryptocurrencies (note that the heavy-tailed Cauchy distribution has  $N_{df} = 1$ , and the normal distribution has  $N_{df} = \infty$ ), the average  $a$ -param of the generalized hyperbolic distribution is 0.141 for stocks, and the average value of the  $\alpha$  parameter of the stable distribution is 1.70 for stocks (note that the Cauchy distribution has  $\alpha = 1$ , and the normal distribution has  $\alpha = 2$ ). Tables 3–7 also show the loss function, which takes similar values for the four fat-tailed distributions, though among them the stable is the worst in most cases. Specifically, in six cases the generalized hyperbolic distribution provides the best fitting (i.e., lowest loss function) among all five analyzed distributions; in five cases, the non-centered, t-student function presents the best fitting; in two cases the Johnson  $S_U$  fits best, in all cases considering the minimum loss function as the decision for determining the quality of the distribution. In Table 8, we displayed a comparison using the Akaike information criterion (AIC), which includes a penalty depending on the number of parameters of the distribution (which are two for the normal, four for

non-centered t-Student, Johnson's  $S_U$ , and Lévy stable, and are five for the generalized hyperbolic distribution). The number displayed in Table 8 (the  $AIC$ ) is the following:

$$AIC = 2N_{params} + 2n \text{ loss}(p_d)$$

where  $N_{params}$  is the number of parameters of the distribution,  $n$  is the number of points of the time series (spread), and  $\text{loss}(p_d)$  is the loss function as defined in Equation (8). Interestingly, the Akaike criterion relegates the generalized hyperbolic distribution for all stocks: for them, the best choice according to  $AIC$  would be the non-centered t-student distribution in all cases except one, where Johnson's  $S_U$  should be preferred.

**Table 3.** Spreads fitting to normal distribution.

Spread Name	$\varphi$	$E_0$	$R^2$	Loc	Scale	Loss
TTE/PSX	0.985	-0.440	0.970	$-4 \cdot 10^{-17}$	$1.84 \cdot 10^{-2}$	-2.5761
BAC/PNC	0.979	-1.622	0.959	$3 \cdot 10^{-17}$	$1.05 \cdot 10^{-2}$	-3.1409
BKNZ/MAR	0.988	2.714	0.977	$3 \cdot 10^{-16}$	$1.83 \cdot 10^{-2}$	-2.5800
UPM.HE/BNR.DE	0.991	-0.511	0.984	$-8 \cdot 10^{-17}$	$1.72 \cdot 10^{-2}$	-2.6441
AMGN/SYK	0.990	0.623	0.979	$3 \cdot 10^{-18}$	$1.73 \cdot 10^{-2}$	-2.6410
DUK/SRE	0.986	0.597	0.972	$-3 \cdot 10^{-17}$	$1.10 \cdot 10^{-2}$	-3.0897
GOOG/INTU	0.991	-1.093	0.983	$-8 \cdot 10^{-17}$	$1.68 \cdot 10^{-2}$	-2.6669
ESS/EQR	0.986	1.309	0.972	$-1 \cdot 10^{-16}$	$7.31 \cdot 10^{-3}$	-3.4989
DIS/VZ	0.990	-0.169	0.980	$1 \cdot 10^{-17}$	$2.04 \cdot 10^{-2}$	-2.4756
MNST/MDLZ	0.988	-0.720	0.978	$-2 \cdot 10^{-17}$	$1.53 \cdot 10^{-2}$	-2.7616
CAT/RELX	0.991	1.403	0.982	$-2 \cdot 10^{-16}$	$2.01 \cdot 10^{-2}$	-2.4889
TRX-USD/BTC-USD	0.963	-11.906	0.942	$-2 \cdot 10^{-15}$	$5.98 \cdot 10^{-2}$	-1.3982
DOGE-USD/XRP-USD	0.976	1.552	0.949	$-2 \cdot 10^{-16}$	$7.60 \cdot 10^{-2}$	-1.1575

This table shows the fitting parameters of the spread residuals to the normal distribution. The loss function values are shown in the last column.

**Table 4.** Spreads fitting to non-centered t-student distribution.

Spread Name	Loc	Scale	sk.param.	df	Loss
TTE-vs.-PSX	$-1.69 \cdot 10^{-3}$	$1.33 \cdot 10^{-2}$	$1.03 \cdot 10^{-1}$	4.49	-2.6668
BAC-vs.-PNC	$-1.80 \cdot 10^{-3}$	$7.48 \cdot 10^{-3}$	$1.92 \cdot 10^{-1}$	3.97	-3.2098
BKNG-vs.-MAR	$-4.833 \cdot 10^{-4}$	$1.20 \cdot 10^{-2}$	$3.37 \cdot 10^{-2}$	3.54	-2.7029
UPM.HE-vs.-BNR.DE	$1.523 \cdot 10^{-3}$	$1.25 \cdot 10^{-2}$	$9.60 \cdot 10^{-2}$	4.10	-2.7055
AMGN-vs.-SYK	$-4.86 \cdot 10^{-4}$	$1.10 \cdot 10^{-2}$	$3.80 \cdot 10^{-2}$	3.24	-2.7640
DUK-vs.-SRE	$1.01 \cdot 10^{-5}$	$6.78 \cdot 10^{-3}$	$1.13 \cdot 10^{-2}$	3.35	-3.2582
GOOG-vs.-INTU	$1.11 \cdot 10^{-3}$	$1.15 \cdot 10^{-2}$	$-7.40 \cdot 10^{-2}$	3.64	-2.7530
ESS-vs.-EQR	$4.71 \cdot 10^{-4}$	$5.42 \cdot 10^{-3}$	$-7.30 \cdot 10^{-2}$	4.40	-3.5599
DIS-vs.-VZ	$-2.44 \cdot 10^{-3}$	$1.39 \cdot 10^{-2}$	$1.37 \cdot 10^{-1}$	3.63	-2.5672
MNST-vs.-MDLZ	$3.33 \cdot 10^{-5}$	$9.57 \cdot 10^{-3}$	$3.06 \cdot 10^{-3}$	3.14	-2.8922
CAT-vs.-RELX	$-6.20 \cdot 10^{-4}$	$1.54 \cdot 10^{-2}$	$2.97 \cdot 10^{-2}$	4.73	-2.53437
TRX-USD-vs.-BTC-USD	$1.71 \cdot 10^{-3}$	$2.70 \cdot 10^{-2}$	$-7.62 \cdot 10^{-2}$	1.86	-1.6071
DOGE-USD-vs.-XRP-USD	$4.87 \cdot 10^{-3}$	$3.73 \cdot 10^{-2}$	$8.28 \cdot 10^{-2}$	2.26	-1.3909

This table shows the fitting parameters of the spread residuals to the non-centered t-student distribution. The loss function values are shown in the last column.

**Table 5.** Spreads fitting to Johnson- $S_U$  distribution.

Spread Name	Loc	Scale	a Param.	b Param.	Loss
TTE-vs.-PSX	$-1.175 \cdot 10^{-3}$	$2.048 \cdot 10^{-2}$	$-6.594 \cdot 10^{-2}$	1.477	-2.6659
BAC-vs.-PNC	$-1.261 \cdot 10^{-3}$	$1.085 \cdot 10^{-2}$	$-1.242 \cdot 10^{-1}$	1.384	-3.2097
BKNG-vs.-MAR	$-2.802 \cdot 10^{-4}$	$1.616 \cdot 10^{-2}$	$-1.804 \cdot 10^{-2}$	1.277	-2.7016
UPM.HE-vs.-BNR.DE	$-1.051 \cdot 10^{-3}$	$1.857 \cdot 10^{-2}$	$-6.136 \cdot 10^{-2}$	1.416	-2.7055
AMGN-vs.-SYK	$-2.521 \cdot 10^{-4}$	$1.403 \cdot 10^{-2}$	$-1.866 \cdot 10^{-2}$	1.208	-2.7633
DUK-vs.-SRE	$1.459 \cdot 10^{-4}$	$8.843 \cdot 10^{-3}$	$6.114 \cdot 10^{-3}$	1.234	-3.2562
GOOG-vs.-INTU	$8.333 \cdot 10^{-4}$	$1.583 \cdot 10^{-2}$	$5.055 \cdot 10^{-2}$	1.305	-2.7527
ESS-vs.-EQR	$3.032 \cdot 10^{-4}$	$8.352 \cdot 10^{-3}$	$4.389 \cdot 10^{-2}$	1.474	-3.5596
DIS-vs.-VZ	$-1.670 \cdot 10^{-3}$	$1.904 \cdot 10^{-2}$	$-8.551 \cdot 10^{-2}$	1.305	-2.5668
MNST-vs.-MDLZ	$8.512 \cdot 10^{-5}$	$1.204 \cdot 10^{-2}$	$2.073 \cdot 10^{-3}$	1.188	-2.8919
CAT-vs.-RELX	$-4.178 \cdot 10^{-4}$	$2.470 \cdot 10^{-2}$	$-1.845 \cdot 10^{-2}$	1.541	-2.53439
TRX-USD-vs.-BTC-USD	$1.291 \cdot 10^{-3}$	$2.422 \cdot 10^{-2}$	$5.110 \cdot 10^{-2}$	0.834	-1.6151
DOGE-USD-vs.-XRP-USD	$3.380 \cdot 10^{-3}$	$3.712 \cdot 10^{-2}$	$4.822 \cdot 10^{-2}$	0.938	-1.3937

This table shows the fitting parameters of the spread residuals to the Johnson- $S_U$  distribution. The loss function values are shown in the last column.

**Table 6.** Spreads fitting to generalized hyperbolic distribution.

Spread Name	Loc	Scale	b Param.	a Param	p Param	Loss
TTE-vs.-PSX	$-7.546 \cdot 10^{-5}$	$2.825 \cdot 10^{-2}$	$6.617 \cdot 10^{-3}$	$7.487 \cdot 10^{-3}$	-2.248	-2.6666
BAC-vs.-PNC	$-6.262 \cdot 10^{-4}$	$1.416 \cdot 10^{-2}$	$8.099 \cdot 10^{-2}$	$4.401 \cdot 10^{-1}$	-1.762	-3.2099
BKNG-vs.-MAR	$-2.742 \cdot 10^{-4}$	$2.263 \cdot 10^{-2}$	$6.031 \cdot 10^{-3}$	$7.532 \cdot 10^{-3}$	-1.770	-2.7029
UPM.HE-vs.-BNR.DE	$9.823 \cdot 10^{-5}$	$2.529 \cdot 10^{-2}$	$5.633 \cdot 10^{-3}$	$2.000 \cdot 10^{-2}$	-2.043	-2.7053
AMGN-vs.-SYK	$-3.966 \cdot 10^{-5}$	$1.978 \cdot 10^{-2}$	$2.498 \cdot 10^{-3}$	$2.002 \cdot 10^{-2}$	-1.618	-2.7639
DUK-vs.-SRE	$1.648 \cdot 10^{-4}$	$1.242 \cdot 10^{-2}$	$-1.782 \cdot 10^{-2}$	$1.803 \cdot 10^{-2}$	-1.676	-3.2582
GOOG-vs.-INTU	$3.324 \cdot 10^{-4}$	$2.080 \cdot 10^{-2}$	$-2.430 \cdot 10^{-2}$	$3.315 \cdot 10^{-1}$	-1.624	-2.7531
ESS-vs.-EQR	$1.499 \cdot 10^{-4}$	$1.128 \cdot 10^{-2}$	$-3.160 \cdot 10^{-2}$	$2.122 \cdot 10^{-1}$	-2.159	-3.5599
DIS-vs.-VZ	$-6.468 \cdot 10^{-4}$	$2.588 \cdot 10^{-2}$	$4.018 \cdot 10^{-2}$	$2.112 \cdot 10^{-1}$	-1.738	-2.5671
MNST-vs.-MDLZ	$1.268 \cdot 10^{-4}$	$1.588 \cdot 10^{-2}$	$-8.602 \cdot 10^{-2}$	$2.241 \cdot 10^{-1}$	-1.401	-2.8923
CAT-vs.-RELX	$-3.954 \cdot 10^{-4}$	$3.342 \cdot 10^{-2}$	$3.207 \cdot 10^{-2}$	$5.813 \cdot 10^{-2}$	-2.360	-2.53443
TRX-USD-vs.-BTC-USD	$1.214 \cdot 10^{-3}$	$1.419 \cdot 10^{-2}$	$-5.196 \cdot 10^{-3}$	$1.873 \cdot 10^{-1}$	0.066	-1.6208
DOGE-USD-vs.-XRP-USD	$2.150 \cdot 10^{-4}$	$4.020 \cdot 10^{-2}$	$2.130 \cdot 10^{-3}$	$2.490 \cdot 10^{-1}$	-0.592	-1.3921

This table shows the fitting parameters of the spread residuals to the generalized hyperbolic distribution. The loss function values are shown in the last column.

**Table 7.** Spreads fitting to Lévy stable distribution.

Spread Name	Loc	Scale	$\beta$	$\alpha$	Loss
TTE-vs.-PSX	$1.076 \cdot 10^{-4}$	$1.055 \cdot 10^{-2}$	$1.057 \cdot 10^{-1}$	1.771	-2.6640
BAC-vs.-PNC	$4.212 \cdot 10^{-5}$	$5.948 \cdot 10^{-3}$	$1.327 \cdot 10^{-1}$	1.715	-3.2061
BKNG-vs.-MAR	$9.657 \cdot 10^{-5}$	$9.714 \cdot 10^{-3}$	$4.331 \cdot 10^{-2}$	1.697	-2.7014
UPM.HE-vs.-BNR.DE	$-8.969 \cdot 10^{-6}$	$9.861 \cdot 10^{-3}$	$6.902 \cdot 10^{-2}$	1.707	-2.7002
AMGN-vs.-SYK	$1.190 \cdot 10^{-4}$	$8.941 \cdot 10^{-3}$	$4.463 \cdot 10^{-2}$	1.653	-2.7609
DUK-vs.-SRE	$1.720 \cdot 10^{-4}$	$5.466 \cdot 10^{-3}$	$4.284 \cdot 10^{-2}$	1.662	-3.2568
GOOG-vs.-INTU	$5.256 \cdot 10^{-5}$	$9.231 \cdot 10^{-3}$	$-3.348 \cdot 10^{-2}$	1.687	-2.7497
ESS-vs.-EQR	$-2.419 \cdot 10^{-5}$	$4.277 \cdot 10^{-3}$	$-5.984 \cdot 10^{-2}$	1.747	-3.5558
DIS-vs.-VZ	$7.915 \cdot 10^{-5}$	$1.110 \cdot 10^{-2}$	$1.107 \cdot 10^{-1}$	1.682	-2.5636
MNST-vs.-MDLZ	$7.616 \cdot 10^{-5}$	$7.723 \cdot 10^{-3}$	$1.385 \cdot 10^{-3}$	1.606	-2.8875
CAT-vs.-RELX	$-9.711 \cdot 10^{-5}$	$1.207 \cdot 10^{-2}$	$3.309 \cdot 10^{-2}$	1.768	-2.5299
TRX-USD-vs.-BTC-USD	$-3.380 \cdot 10^{-3}$	$2.265 \cdot 10^{-2}$	$-0.637 \cdot 10^{-2}$	1.284	-1.597
DOGE-USD-vs.-XRP-USD	$-5.851 \cdot 10^{-4}$	$3.160 \cdot 10^{-2}$	$-6.865 \cdot 10^{-2}$	1.447	-1.381

This table shows the fitting parameters of the spread residuals to the Lévy stable distribution. The loss function values are shown in the last column.

**Table 8.** Akaike information criterion for the analyzed distributions.

Spread Name	Normal	t-Student	Johnson's $S_U$	Gen. hyperb.	Lévy Stable
TTE-vs.-PSX	−10,073.7	<b>−10,424.5</b>	−10,421.0	−10,421.7	−10,413.6
BAC-vs.-PNC	−12,622.4	<b>−12,895.4</b>	−12,895.0	−12,893.8	−12,880.5
BKNG-vs.-MAR	−10,367.6	<b>−10,857.7</b>	−10,852.4	−10,855.7	−10,851.6
UPM.HE-vs.-BNR.DE	−10,223.4	<b>−10,456.9</b>	−10,456.9	−10,454.1	−10,436.4
AMGN-vs.-SYK	−10,612.8	<b>−11,103.3</b>	−11,100.5	−11,100.9	−11,090.8
DUK-vs.-SRE	−12,416.6	<b>−13,090.0</b>	−13,081.9	−13,088	−13,084.3
GOOG-vs.-INTU	−10,716.9	<b>−11,059.1</b>	−11,057.9	−11,057.5	−11,045.8
ESS-vs.-EQR	−14,061.6	<b>−14,302.8</b>	−14,301.6	−14,300.8	−14,286.3
DIS-vs.-VZ	−9947.9	<b>−10,312.1</b>	−10,310.5	−10,309.7	−10,297.7
MNST-vs.-MDLZ	−11,097.6	<b>−11,618.6</b>	−11,617.4	−11,617.1	−11,599.8
CAT-vs.-RELX	−10,001.4	−10,180.17	<b>−10,180.24</b>	−10,178.4	−10,162.1
TRX-USD-vs.-BTC-USD	−2800.8	−3215.8	−3231.9	<b>−3241.3</b>	−3195.6
DOGE-USD-vs.-XRP-USD	−2317.9	−2782.2	<b>−2787.8</b>	−2782.6	−2762.3

This table shows the values of the Akaike information criterion for each of the analyzed time series (spreads) and each of the 5 analyzed probability distributions. Bold numbers indicate the best fitting (minimum Akaike information criterion).

#### 4.2. Trading Rules

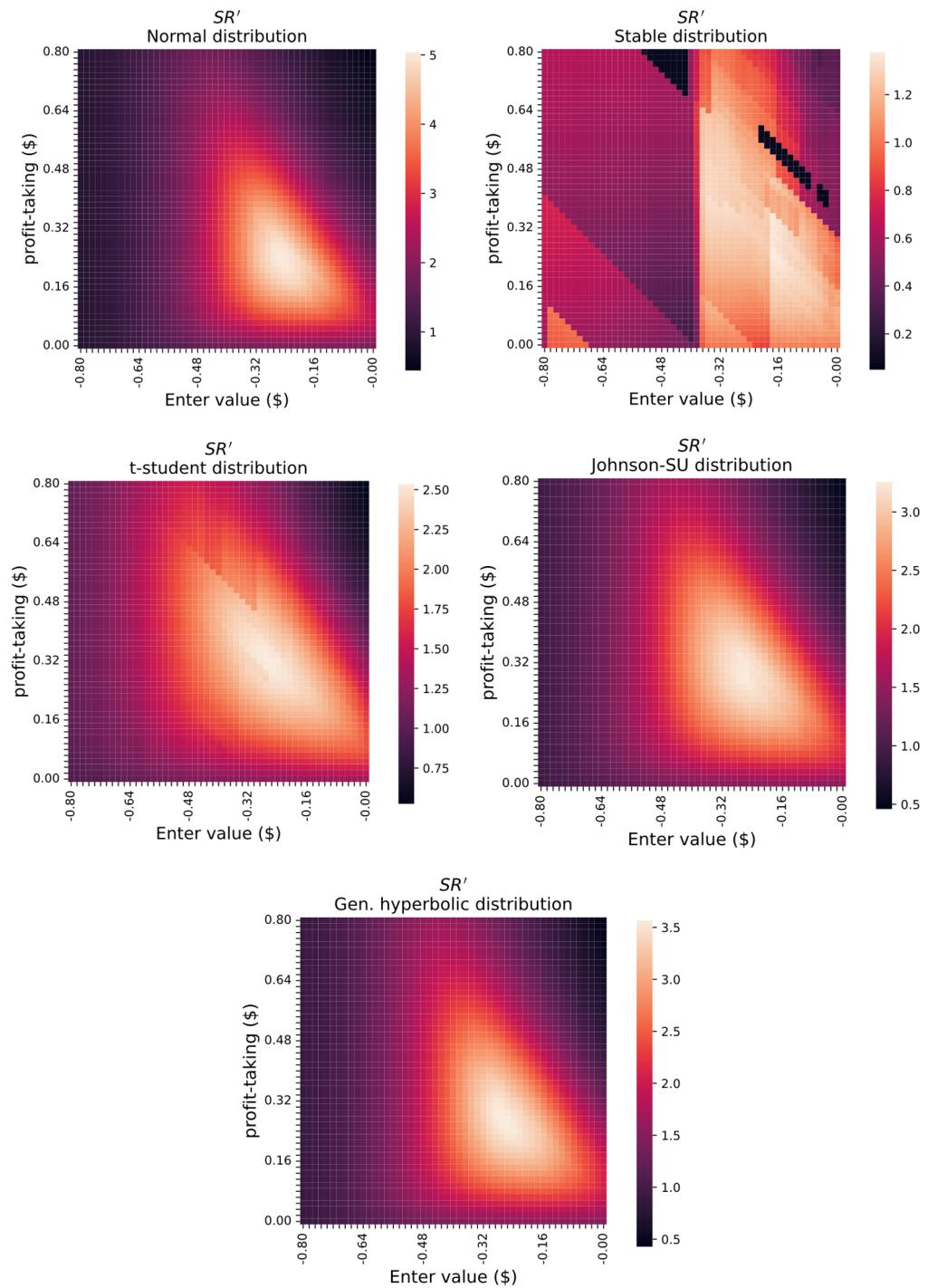
In this section, we presented the performance of different trading strategies. We first (Section 4.2.1) displayed the case that is commonly assumed in the literature, which corresponds to the Ornstein–Uhlenbeck equation with constant parameters. Section 4.2.2 presents an attempt to overcome such limitations of the model, considering discrete changes in the mean ( $E_0$ ) through Poisson events.

##### 4.2.1. Mean Reversion with Constant Parameters

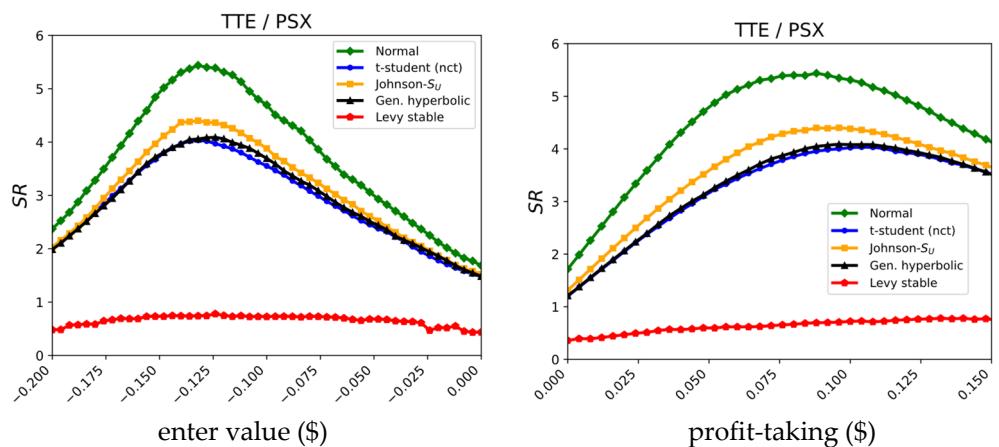
Figure 2 shows the heatmaps indicating profitability of the trading strategies. Each subplot corresponds to calculations made using synthetic data to generate time series of the spreads. A different distribution was used to generate the residuals of the spreads in each case; all were carried out using 200,000 Monte Carlo paths. They present the rescaled Sharpe ratio from semi-deviation ( $SR'$ ), which measures risk-adjusted profitability (see the Supplementary Materials for other performance metrics). The maximum horizon of all figures corresponds to one year. The  $x$  (enter value) and  $y$  (profit-taking) axes of Figure 2 represent the value of the spread when the position is built and unwound, respectively. Note that the heatmaps of this paper present values of the profit-taking and stop-loss thresholds which correspond to additive values with respect to the enter value, and the enter values correspond to additive values with respect to the corresponding  $E_0$ .

As can be noted in Figure 2, different distributions lead to differences in the maximum value of the Sharpe ratios ( $SR'$ ). Hence, the choice of distribution to fit the spread residuals has non-trivial consequences on the trading strategy. This is more clearly seen in Figure 3 (see Supplementary Materials for examples corresponding to other pairs). Each point of the curve of the left subplot corresponds to the maximum value  $SR$  sweeping all profit-taking thresholds and a given enter value. The right subplot corresponds to the maximum value  $SR$  sweeping all enter values for a given profit-taking threshold (measured with respect to the enter value). These plots inform us about several interesting findings: (i) The value of the maximum  $SR$  strongly depends on the distribution chosen to model the residuals of the spread (it is nearly six for the normal distribution, and four for the t-student and generalized hyperbolic distributions). (ii) The optimal profit-taking threshold also strongly differs for different distributions (it is 0.064 for the normal distribution and 0.048 for the t-student distribution). Accordingly, this will probably have an impact on the number of times a trade is entered. These results correspond to an analysis where no stop-loss orders

are in force (i.e., the stop-loss parameter is set to  $\pm\infty$ ). Our analyses indicate that farther (higher size) stop-loss parameters lead to higher values of the Sharpe ratios, and thus the optimal trading rules consist of setting them to  $-\infty$  (if the bet is long in spread) or to  $+\infty$  (if the bet is short in spread).



**Figure 2.** Sharpe ratios' heatmaps from fitting to each probability distribution (XRP-USD/DOGE-USD example): heatmaps which present the Sharpe ratio  $SR'$  as a function of different values of the trading rules (enter value and profit-taking thresholds). Each point of the heatmaps was calculated using synthetic data (200,000 Monte Carlo paths) for the time series of the residuals of the spread. Each subplot corresponds to the synthetic data being generated with a different probability density function (in all cases, using the parameters of the optimal fitting of the observed residuals to the corresponding distribution).



**Figure 3.** Trading rules performance (TTE/PSX example). (**Left**): Sharpe ratios vs. enter value (maximum for all profit-taking thresholds) for the five analyzed distributions; (**Right**): Sharpe ratios vs. profit-taking threshold (maximum for all enter values).

Although the analysis shown in Figures 2 and 3 is based on static parameters of the Ornstein–Uhlenbeck equation ( $E_0, \varphi$ ), in a real-life trading action, the investor would probably set a finite stop-loss threshold. In the next section, we overcome this drawback, allowing a regime switching process for the  $E_0$  parameter.

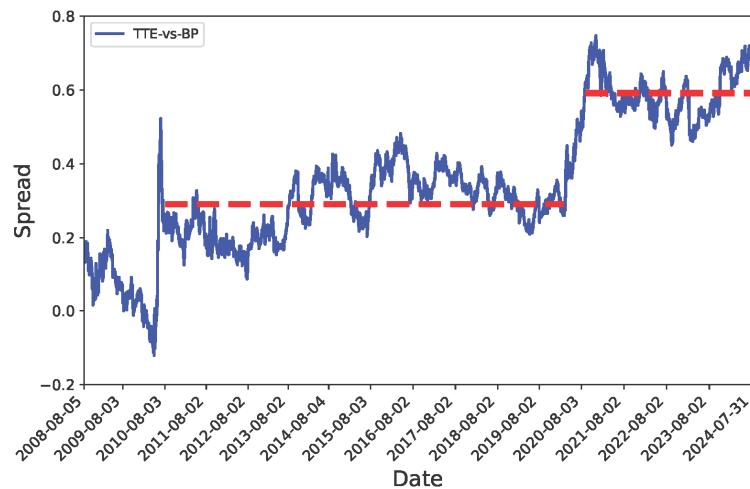
#### 4.2.2. Regime Switching and Stop-Loss Orders

This section comprises the analysis of optimal trading rules with regime switching, modeling the switch as a Poisson process and the noise using the previously considered fat-tailed distributions. This analysis complements previous research (Bai & Wu, 2017; Y. Yang et al., 2017; Altay et al., 2018), which considers optimal trading rules in the presence of regime switching, but not the effect of fat-tailed distributions. This analysis also extends the work of Endres and Stüberger (2019b), which does so through jump components (rather than fat-tailed distributions), but it does not optimize the trading rules as we do (their enter value is  $(E_0 \pm \sigma(t)/2)$ , and the profit-taking threshold is the opposite of the enter value around  $E_0$ ).

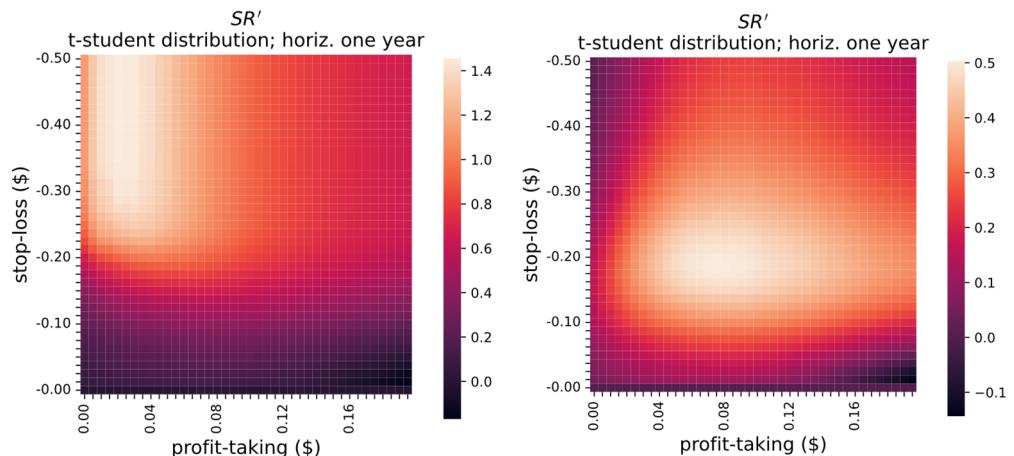
Figure 4 shows the BP/TTE spread calculated from 2008 to 2024. As can be seen, different regions where the spread oscillates around distinct values exist (approximately indicated with red horizontal lines). Recalculating the spread for the period from 2010 to 2020 indicates that it is stationary, and hence it could be considered a candidate for pairs trading before March 2020. However, the regime switching it then undergoes advises against it. The change of the long-term mean level of the spread ( $E_0$ ) has a dramatic effect on the optimal trading rules. As stated above, if one generates synthetic data naively, ignoring changes in  $E_0$ , then the most profitable strategies omit a stop-loss order. They involve holding your portfolio until the spread finally reverts to the mean, no matter how large your temporary loss is.

Nevertheless, a more realistic approach which includes regime switching for synthetic data leads to optimal trading rules with a non-infinite stop-loss threshold, as we will see below. Let us assume that the long-term mean level of the spread ( $E_0$ ) can change. We modeled the transition as a Poisson event. Results can be seen in Figure 5. The intermediate period is a time window from August 2010 to February 2020, which corresponds to 2310 trading days. Hence, we set a daily probability of the Poisson event of  $1/2310 \approx 0.0004329$ . We also assume that, if such an event takes place, there is a 50% probability of  $E_0$  increasing or dropping; we set the size of the variation of  $E_0$  equal to 0.62 because that is the difference between the values of the long-term mean level separately calculated using data from the two time stretches marked with red lines in Figure 4 (the set of residuals of each time

window is fitted to a different unimodal distribution). Accordingly, Figure 5 shows the heatmaps ignoring (left graph) and considering (right graph) Poisson events (it corresponds to the t-student distribution, see Supplementary Materials for other distributions. The optimal thresholds are those which maximize the  $SR'$ . This figure shows that for a given enter value the optimal stop-loss in the absence of changes in  $E_0$  (Poisson events) has an infinite size (stop-loss =  $\pm\infty$ ), while in the presence of changes in  $E_0$  it has a finite value. It can be stated that regime switching reduces the maximum attainable  $SR'$ , though the value with regime switching is still relatively high (about 0.5).



**Figure 4.** BPE/TTE spread in the presence of regime switching. The blue line represents the observed spread (calculated using all the data of the period between 2008 and 2024). The dashed red lines present the approximate values of  $E_0$  the spread oscillates around. The clear difference between both dashed lines indicates that a regime switching happened.



**Figure 5.** Effect of the regime switching in the context of Poisson events: heatmaps which present the Sharpe ratio obtained from synthetic data of the spread residuals modeled with a t-student distribution. Each point corresponds to a given pair of thresholds (profit-taking and stop-loss). The (left) graph shows the  $SR'$  without Poisson events for changes in the long-term mean level ( $E_0$ ) of the Ornstein–Uhlenbeck equation; the (right) considers changes in the long-term mean level. In the former case, the stop-loss which maximizes the Sharpe ratio is infinite; in the latter case, it is finite.

The result of Figure 5-left agrees with previous research (López de Prado, 2018), which states that the optimal rules for pairs trading consist of unwinding the position as soon as a small profit is realized and waiting (eventually for long periods) until your position is in-the-money, in case it is temporarily out-of-the-money (in practice, this is equivalent to

setting a stop-loss threshold of infinite size, i.e., never to make an stop-loss order). However, such a bogus conclusion is due to naïve modeling, i.e., to the unrealistic assumption that the  $E_0$  must forcedly remain unchanged. Note that in practice the price of a given stock can abruptly change for many reasons, like the company launching a new product, its board choosing a mistaken strategy, etc. This phenomenon can easily cause that  $E_0$  takes a new value or even make both prices no longer cointegrated. The results displayed in Figure 5-right show that when the naïve modeling is avoided, i.e., when we consider regime switching, the optimal stop-loss order is no longer an infinite size. Figure 5-right also reveals that considering regime switching severely reduces the maximum  $SR'$ . This indicates that spreads that are expected to be especially prone to changes in  $E_0$  should be discarded in actual pairs trading.

#### 4.3. Out-of-Sample Calculations

For the sake of completeness, we also present the out-of-sample profits of trading rules calculated using the methods presented in this paper. These results correspond to the actual profits that would have been obtained if an investor had traded the spreads of the analyzed pairs following optimal trading rules calculated using different probability density functions to fit the spread residuals. For every pair of stocks, we considered carrying out pairs trading between 20 February 2021 and 20 February 2025; for pairs of cryptocurrencies, we have considered carrying out pairs trading between 20 February 2023 and 20 February 2025. For every 3 months, we took the data of the spreads in the 5 years immediately before the beginning of that period. We fitted the residuals of the spreads of that 5-year period to each chosen probability density function and used the found probability density function to calculate optimal trading rules (heatmaps), calculated using 20,000 Monte Carlo paths. In each case, we set an arbitrary stop-loss parameter depending on the size of the past variations of the spread. Our trading rules cover both the positions short in spread (*positive spread*) and long in spread (*negative spread*). The trading rules specify when we enter and when we unwind the pair; the profit is calculated as the difference between the observed prices on both dates (note that, if the spread reaches a profit-taking threshold but unwinding the position results in a loss, then we do not unwind the position, because that would be contradictory to the concept of profit-taking). Whenever we enter a trade, we buy one dollar (or euro) of the long stock, and we sell short  $\gamma$  dollars (or euros) of the short stock (or the converse). Therefore, the cost of entering our position can be either positive or negative (if the size, measured in currency units, of the shorted position is higher than the size of the long position). We set a maximum horizon of 2 years; in all cases we consider transaction costs of 0.5%.

The results are shown in Table 9 (for stocks) and Table 10 (for cryptocurrencies). Columns with label “#” indicate the number of times that a trade was entered; columns with label *Cost* indicate the average cost of entering our position. The column with label *Profit* is the profit corresponding to the whole considered period measured in currency units (USD or euros). For example, if the price of a given entered long position is 1 dollar (it can be either 1 or  $\gamma$ , which is usually approximately 1) then a total profit of, for example, 0.16 implies that the profit was 16% of the cost of entering the long position.

**Table 9.** Stocks' out-of-sample pairs trading profitability.

Sector	Pair	Distribution	Negative Spread			Positive Spread			Total		
			#	Cost	Profit	#	Cost	Profit	#	Cost	Profit
Consumer discretionary	BKNG /MAR	Normal	2	−0.008	0.109	3	0.008	0.382	5	0.002	0.491
		t-student	2	−0.008	0.153	2	0.008	0.154	4	0.000	0.307
		Johnson- $S_U$	5	−0.008	0.438	2	0.008	0.249	7	−0.003	0.687
		Gen. Hyperb.	4	−0.011	0.335	0		0.000	4	−0.011	0.335
		Levy stable	1	−0.011	0.001	1	0.011	0.059	2	0.000	0.060
Materials	UPM.HE /BNR.DE	Normal	3	0.071	0.060	4	−0.067	0.271	7	−0.008	0.332
		t-student;	3	0.071	0.170	3	−0.074	0.148	6	−0.001	0.318
		Johnson- $S_U$	3	0.075	0.080	3	−0.074	0.179	6	0.001	0.259
		Gen. Hyperb.;	3	0.071	0.062	3	−0.063	0.145	6	0.004	0.207
		Levy stable	3	0.060	0.092	2	−0.061	0.155	5	0.012	0.248
Industrials	RELX /CAT	Normal	1	−0.091	0.117	2	0.127	0.011	3	0.054	0.128
		t-student	1	−0.105	0.231	3	0.130	0.021	4	0.071	0.252
		Johnson- $S_U$	0		0.000	2	0.129	0.011	2	1.129	0.011
		Gen. Hyperb.	2	−0.143	0.330	1	0.208	−0.004	3	−0.026	0.326
		Levy stable	1	−0.184	0.416	2	0.206	0.153	3	0.109	0.569
Communication services	VZ/DIS	Normal	1	−0.214	0.103	2	0.214	0.012	3	0.072	0.115
		t-student	1	−0.214	0.074	2	0.214	0.012	3	0.072	0.086
		Johnson- $S_U$	1	−0.214	0.074	2	0.214	0.012	3	0.072	0.086
		Gen. Hyperb.	1	−0.214	0.053	2	0.214	0.012	3	0.072	0.065
		Levy stable	1	−0.214	0.034	3	0.214	0.024	4	0.107	0.058
Consumer staples	MDLZ /MNST	Normal	1	−0.051	0.185	3	0.051	0.097	4	0.026	0.282
		t-student	1	−0.051	0.222	2	0.051	0.160	3	0.017	0.382
		Johnson- $S_U$	1	−0.051	0.194	2	0.051	0.141	3	0.017	0.334
		Gen. Hyperb.	1	−0.051	0.185	2	0.051	0.027	3	0.017	0.212
		Levy stable	1	−0.051	0.106	4	0.051	0.053	5	0.031	0.159
Healthcare	SYK /AMGN	Normal	7	0.170	0.006	2	−0.170	0.163	9	0.095	0.168
		t-student	3	0.170	−0.251	1	−0.170	0.005	4	0.085	−0.245
		Johnson- $S_U$	4	0.170	−0.019	1	−0.170	0.005	5	0.102	−0.014
		Gen. Hyperb.	5	0.170	0.070	1	−0.170	0.157	6	0.073	0.227
		Levy stable	2	0.170	0.064	4	−0.170	0.285	6	−0.057	0.349
Utilities	SRE /DUK	Normal	4	0.048	0.061	3	−0.048	0.149	7	0.007	0.210
		t-student	3	0.048	−0.021	2	−0.048	0.112	5	0.010	0.091
		Johnson- $S_U$	4	0.048	0.040	2	−0.048	0.099	6	0.016	0.139
		Gen. Hyperb.	4	0.048	0.040	2	−0.048	0.099	6	0.016	0.139
		Levy stable	3	0.048	−0.017	1	−0.048	0.032	4	0.024	0.015
Energy	TTE /PSX	Normal	2	−0.038	0.052	6	0.031	0.496	8	0.014	0.547
		t-student	1	−0.038	0.027	8	0.032	0.718	9	0.024	0.745
		Johnson- $S_U$	1	−0.038	0.027	5	0.034	0.480	6	0.022	0.507
		Gen. Hyperb.	1	−0.038	0.027	7	0.032	0.639	8	0.023	0.666
		Levy stable	3	−0.038	0.184	6	0.023	0.340	9	0.003	0.524
Financials	PNC /BAC	Normal	2	−0.041	0.049	4	0.041	0.214	6	0.014	0.263
		t-student	2	−0.041	0.055	3	0.041	0.136	5	0.014	0.191
		Johnson- $S_U$	2	−0.041	0.049	3	0.041	0.136	5	0.014	0.185
		Gen. Hyperb.	5	−0.052	0.185	0		0.000	5	−0.052	0.185
		Levy stable	4	−0.019	0.199	4	0.019	0.219	8	0.000	0.418
Information Technology	INTU /GOOG	Normal	4	0.058	0.121	4	−0.058	0.192	8	0.000	0.313
		t-student	4	0.058	0.100	5	−0.058	0.226	9	−0.006	0.327
		Johnson- $S_U$	4	0.058	0.101	4	−0.058	0.271	8	0.000	0.372
		Gen. Hyperb.	4	0.058	0.128	4	−0.058	0.307	8	0.000	0.434
		Levy stable	2	0.058	0.234	4	−0.058	0.328	6	−0.019	0.562
Real State	EQR/ESS	Normal	7	−0.020	0.304	4	0.020	0.144	11	−0.006	0.448
		t-student	4	−0.020	0.213	4	0.020	0.192	8	0.000	0.405
		Johnson- $S_U$	6	−0.004	0.180	2	0.004	0.106	8	−0.002	0.286
		Gen. Hyperb.	4	−0.004	0.121	2	0.004	0.106	6	−0.002	0.227
		Levy stable	1	−0.004	0.050	4	0.004	0.221	5	0.003	0.271

This table presents the out-of-sample profits of the 11 analyzed pairs of stocks (one for each GICS sector). The profits depend on the optimal trading rules, which were calculated using synthetic data with spread residuals following one of the five analyzed probability distributions; hence each block has five rows. The first vertical block indicates the sector, tickers of the two stocks of the long–short pair, and name of the distribution. The last three blocks correspond to the bets which are short in spread, long in spread, and sum of both, respectively. For each, we displayed the number of times a bet was entered, the average profit, and the cost of entering the position (which can be negative because the pair is long–short).

**Table 10.** Cryptocurrencies' out-of-sample pairs trading profitability.

Spread	Distribution	Negative Spread			Positive Spread			Total		
		#	Cost	Profit	#	Cost	Profit	#	Cost	Profit
BTC-USD /TRX-USD	Normal	0			0			0		
	t-student	3	−0.099	0.562	0			3	−0.099	0.562
	Johnson- $S_U$	0			0			0		
	Gen. Hyperb.	0			0			0		
	Levy stable	2	0.099	0.190	2	−0.099	0.738	4	0.000	0.927
XRP-USD /DOGE-USD	Normal	1	−0.054	0.117	2	0.054	1.124	3	0.018	1.242
	t-student	1	−0.054	0.183	2	0.054	1.354	3	0.018	1.537
	Johnson- $S_U$	1	−0.054	0.073	2	0.054	1.124	3	0.018	1.197
	Gen. Hyperb.	1	−0.054	0.182	2	0.054	1.124	3	0.018	1.307
	Levy stable	4	−0.054	0.469	5	0.054	2.371	9	0.018	2.841

This table presents the out-of-sample profits of the 2 analyzed pairs of cryptocurrencies. The profits depend on the optimal trading rules, which were calculated using synthetic data with spread residuals following one of the five analyzed probability distributions; hence each block has five rows. The first vertical block indicates the sector, tickers of the two cryptocurrencies of the long–short pair, and name of the distribution. The last three blocks correspond to the bets which are short in spread, long in spread, and sum of both, respectively. For each, we display the number of times a bet was entered, the average profit, and the cost of entering the position.

The results displayed in Tables 9 and 10 indicate that the cointegration method produces consistent positive returns, as indicated by previous research (see e.g., Göncü & Akyıldırım, 2016; Rad et al., 2016). Tables 9 and 10 also indicate that the out-of-sample profits obtained usually differ for trading rules derived using fittings of the spread residuals to different probability density functions. This is reasonable because different functions lead to different trading rules, which lead to results that can clearly differ for one single set of historical prices. Nevertheless, for a given pair the total profits are usually positive and are usually the same order of magnitude.

## 5. Conclusions

This study determines optimal trading rules for Ornstein–Uhlenbeck-based pairs trading strategies applied to stocks and cryptocurrencies. The spread residuals are modeled by using fat-tailed distributions (i.e., Lévy stable, generalized hyperbolic, Johnson's  $S_U$ , and non-centered t-student), covering a gap in existing research where such distributional assumptions are narrowly considered in the context of trading thresholds. The model determines efficient entry, profit-taking, and stop-loss levels, addressing more realistic statistical properties of the spread residuals than the naïve normality and non-regime switching assumptions, which are widely present in the literature.

The empirical model assesses the adequacy of the selected fat-tailed distributions to account for heavy tails and skewness. After that, we extended existing frameworks by introducing regime-switching dynamics in trade rules optimization. Specifically, we used a Poisson-based, Markov-switching process that can capture structural breaks and changing market conditions, thereby increasing the strategy's practical relevance.

The main findings reveal that the choice of distribution to model the spread residuals has a significant effect on optimal trading thresholds, especially profit-taking levels. Under the mean reversion model with constant parameters, larger stop-loss thresholds improve profitability. However, the need to shorten the finite stop-loss thresholds in real trading highlights the need to add regime-switching processes to pairs trading rules optimization. When a Poisson process accounting for regime-switching is considered, the trading thresholds are more conservative but better reflect real market constraints. Although the introduction of regime shifts reduces expected profitability, it confirms that optimal pairs trading strategies require unwinding positions with reduced performance.

Future research can benefit from the results of this study by focusing on more precisely addressing the magnitude and persistence of regime shifts when modeling residual spreads with fat-tailed distributions. Therefore, it could consider time-varying parameters or higher-frequency data to better address transitions and tail behavior, as well as more sophisticated models, like GARCH (to give account of volatility clustering) or copulas, which were omitted from this paper for the sake of simplicity. We will also consider tweaking the fitting to Lévy stable distributions. Another planned research line is to apply the knowledge from fat tails presented in this paper in more sophisticated trading strategies, e.g., using further market data, like Value Line ranks, as made in Waggle et al. (2001). The present research is not free from limitations. The assumption of specific parametric fat-tailed distributions and Poisson-based regime switching processes may not capture all the features and complexities of real market dynamics. Furthermore, while transaction costs are considered in the form of explicit fees, other parameters such as slippage, bid-ask spreads, and liquidity constraints are not explicitly addressed and should be considered by further research.

**Supplementary Materials:** The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/ijfs13020096/s1>.

**Author Contributions:** Conceptualization, P.G.-R.; methodology, P.G.-R.; software, P.G.-R.; validation, P.G.-R.; formal analysis, P.G.-R. and E.O.; investigation, P.G.-R., E.O. and J.M.M.; resources, P.G.-R.; data curation, P.G.-R.; writing—original draft preparation, P.G.-R. and E.O.; writing—review and editing, P.G.-R., E.O. and J.M.M.; visualization, P.G.-R. and E.O.; supervision, E.O. and J.M.M.; project administration, E.O. and J.M.M.; funding acquisition, E.O. and J.M.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** Eduardo Ortas and José M. Moneva are grateful for the financial help from the Spanish Ministry of Science, Innovation and Universities (research project PID2023-146084OB-I00) and from the Government of Aragon, Spain (grant number S33\_20R).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The Python code used in this paper is freely available in [https://github.com/pablogr/FIN\\_optimal\\_trading\\_rules\\_fat\\_tails](https://github.com/pablogr/FIN_optimal_trading_rules_fat_tails) (accessed on 27 May 2025). This code can be used to download the raw input data from Yahoo Finance and to make every single calculation whose results are displayed in this paper. The corresponding author is at the disposal of the readers for discussion. Further data are contained within the article and Supplementary Materials.

**Acknowledgments:** We are grateful to Lorién López-Villellas for useful technical support and to Jo Robinson for proofreading.

**Conflicts of Interest:** Pablo García-Risueño discloses that all work corresponding to this article has been carried out in his non-working (free) time and using his own resources (books, internet searches, etc.), i.e., it is totally independent from his labor and paid activities in financial companies. The authors report that there are no competing interests to declare.

## Note

<sup>1</sup> For all distributions, we used the implementations present in the `scipy.stats` v. 1.10.1 Python module.

## References

Aggrawal, P., & Clark, J. M. (2007). ETF betas: A study of their estimation sensitivity to varying time intervals. *Institutional Investor New York*, 41(10), 96.

Altay, S., Colaneri, K., & Eksi, Z. (2018). Pairs trading under drift uncertainty and risk penalization. *International Journal of Theoretical and Applied Finance*, 21(7), 1850046. [\[CrossRef\]](#)

Bağcı, M., & Soylu, P. K. (2024). Classification of the optimal rebalancing frequency for pairs trading using machine learning techniques. *Borsa Istanbul Review*, 24, 83–90. [\[CrossRef\]](#)

Bağcı, M., & Soylu, P. K. (2025). The optimal threshold selection for high-frequency pairs trading via supervised machine learning algorithms. *Computational Economics*. online first. [\[CrossRef\]](#)

Bai, Y., & Wu, L. (2017). Analytic value function for optimal regime-switching pairs trading rules. *Quantitative Finance*, 18(4), 637–654. [\[CrossRef\]](#)

Barzilai, J., & Borwein, J. M. (1988). Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1), 141–148. [\[CrossRef\]](#)

Bergmann, D., & de Oliveira, M. A. (2025). Tail dependence and equilibrium reversion in brazilian pairs trading: A copula-based analysis. *Preprint*. [\[CrossRef\]](#)

Borak, S., Misiorek, A., & Weron, R. (2011). Models for heavy-tailed asset returns. *Statistical Tools for Finance and Insurance*, 2, 21–55.

Bouchaud, J. P., & Potters, M. (2003). *Theory of financial risk and derivative pricing: From statistical physics to risk management*. Cambridge University Press.

Carneiro, L., Gomes, L., Lopes, C., & Pereira, C. (2025). Spillovers between euronext stock indices: The COVID-19 effect. *International Journal of Financial Studies*, 13(2), 66. [\[CrossRef\]](#)

Carrasco-Blázquez, M., De la Orden De la Cruz, C., & Prado-Román, C. (2018). Pairs trading techniques: An empirical contrast. *European Research on Management and Business Economics*, 24(3), 160–167. [\[CrossRef\]](#)

Chen, C. W. S., Chen, M., & Chen, S. Y. (2014). Pairs trading via three-regime threshold autoregressive GARCH models. In V. N. Huynh, V. Kreinovich, & S. Sriboonchitta (Eds.), *Modeling dependence in econometrics. Advances in intelligent systems and computing* (Vol. 251). Springer. [\[CrossRef\]](#)

Chodchuangnirun, B., Zhu, K., & Yamaka, W. (2018). Pairs trading via nonlinear autoregressive GARCH models. In V. N. Huynh, M. Inuiguchi, D. Tran, & T. Denoeux (Eds.), *Integrated uncertainty in knowledge modelling and decision making. IUKM 2018* (Vol. 10758). Lecture Notes in Computer Science. Springer. [\[CrossRef\]](#)

Choudhry, M. (2013). *An introduction to value-at-risk*. John Wiley & Sons.

Do, B., Faff, R., & Hamza, K. (2006, June 28–July 1). *A new approach to modeling and estimation for pairs trading*. 2006 Financial Management Association European Conference, Madrid, Spain.

Elliott, R. J., van der Hoek, J., & Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, 5(3), 271–276. [\[CrossRef\]](#)

Endres, S. (2020). Review of stochastic differential equations in statistical arbitrage pairs trading. *Managerial Economics*, 20(2), 71–118. [\[CrossRef\]](#)

Endres, S., & Stübinger, J. (2019a). A flexible regime switching model with pairs trading application to the S&P 500 high-frequency stock returns. *Quantitative Finance*, 19(10), 1727–1740. [\[CrossRef\]](#)

Endres, S., & Stübinger, J. (2019b). Optimal trading strategies for Lévy-driven Ornstein–Uhlenbeck processes. *Applied Economics*, 51(29), 3153–3169. [\[CrossRef\]](#)

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2), 251–276. [\[CrossRef\]](#)

Galenko, A., Popova, E., & Popova, I. (2012). Trading in the presence of cointegration. *Journal of Alternative Investments*, 15(1), 85–97. [\[CrossRef\]](#)

García-Risueño, P. (2025). Historical simulation systematically underestimates the expected shortfall. *Journal of Risk and Financial Management*, 18(1), 34. [\[CrossRef\]](#)

Göncü, A., & Akyildirim, E. (2016). A stochastic model for commodity pairs trading. *Quantitative Finance*, 16(12), 1843–1857. [\[CrossRef\]](#)

Gribkova, N., Wang, M., & Zitikis, R. (2025). Fundamentals of non-parametric statistical inference for integrated quantiles. *Preprint*. [\[CrossRef\]](#)

Guang, Z. (2021). Pairs trading with general state space models. *Quantitative Finance*, 21(9), 1567–1587. [\[CrossRef\]](#)

Huck, N., & Afawubo, K. (2014). Pairs trading and selection methods: Is cointegration superior? *Applied Economics*, 47(6), 599–613. [\[CrossRef\]](#)

Kawai, R. (2012). Continuous-time modeling of random searches: Statistical properties inference. *Journal of Physics A: Mathematical and Theoretical*, 45(23), 235004. [\[CrossRef\]](#)

Krauss, C. (2016). Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, 31(2), 513–545. [\[CrossRef\]](#)

Larsson, S., Lindberg, C., & Warfheimer, M. (2013). Optimal closing of a pair trade with a model containing jumps. *Applications of Mathematics*, 58(3), 249–268. [\[CrossRef\]](#)

Liou, J. H., Liu, Y. T., & Cheng, L. C. (2024). Price spread prediction in high-frequency pairs trading using deep learning architectures. *International Review of Financial Analysis*, 96, 103793. [\[CrossRef\]](#)

López de Prado, M. (2016). Building diversified portfolios that outperform out of sample. *The Journal of Portfolio Management*, 42(4), 59–69. [\[CrossRef\]](#)

López de Prado, M. (2018). *Advances in financial machine learning*. Wiley.

López de Prado, M. (2023). *Causal factor investing*. Cambridge University Press.

Nair, S. T. G. (2021). Pairs trading in cryptocurrency market: A long-short story. *Investment Management and Financial Innovations*, 18(3), 127–141. [\[CrossRef\]](#)

Namwong, N., Yamaka, W., & Tansuchat, R. (2019). Trading signal analysis with Pairs trading strategy in the stock exchange of Thailand. In V. Kreinovich, & S. Sriboonchitta (Eds.), *Structural changes and their econometric modeling. TES 2019* (Vol. 808). Studies in Computational Intelligence. Springer. [\[CrossRef\]](#)

Plerou, V., Gopikrishnan, P., Gabaix, X., Amaral, L. A. N., & Stanley, H. E. (2001). Price fluctuations, market activity and trading volume. *Quantitative Finance*, 1(2), 262. [\[CrossRef\]](#)

Rad, H., Low, R. K. Y., & Faff, R. (2016). The profitability of pairs trading strategies: Distance, cointegration and copula methods. *Quantitative Finance*, 16(10), 1541–1558. [\[CrossRef\]](#)

Shreve, S. E. (2004). *Stochastic calculus for finance II: Continuous-time models*. Springer.

Simonato, J. G. (2012). GARCH processes with skewed and leptokurtic innovations: Revisiting the Johnson Su case. *Finance Research Letters*, 9(4), 213–219. [\[CrossRef\]](#)

Stübinger, J., & Endres, S. (2018). Pairs trading with a mean-reverting jump–diffusion model on high-frequency data. *Quantitative Finance*, 18(10), 1735–1751. [\[CrossRef\]](#)

Tadi, M., & Witzany, J. (2025). Copula-based trading of cointegrated cryptocurrency Pairs. *Financial Innovation*, 11(1), 40. [\[CrossRef\]](#)

Thaler, R. H. (2015). *Misbehaving. The making of behavioral economics*. W. W. Norton and Company.

Vergara, G., & Kristjanpoller, W. (2024). Deep reinforcement learning applied to statistical arbitrage investment strategy on cryptomarket. *Applied Soft Computing*, 153, 111255. [\[CrossRef\]](#)

Vidyamurthy, G. (2004). *Pairs trading: Quantitative methods and analysis* (Vol. 217). Wiley.

Waggle, D., Agrawal, P., & Johnson, D. (2001). Interaction between value line's timeliness and safety ranks. *Journal of Investing*, 10(1), 53–62. [\[CrossRef\]](#)

Wilkens, S. (2025). Pairs trading in the German stock market: Is there still life in the old dog? *Financial Markets and Portfolio Management. online first*. [\[CrossRef\]](#)

Wu, L., Zang, X., & Zhao, H. (2020). Analytic value function for a pairs trading strategy with a Lévy-driven Ornstein–Uhlenbeck process. *Quantitative Finance*, 20(8), 1285–1306. [\[CrossRef\]](#)

Yang, H., & Malik, A. (2024). Optimal market-neutral multivariate pair trading on the cryptocurrency platform. *International Journal of Financial Studies*, 12(3), 77. [\[CrossRef\]](#)

Yang, Y., Göncü, A., & Pantelous, A. (2017). Pairs trading with commodity futures: Evidence from the Chinese market. *China Finance Review International*, 7(3), 274–294. [\[CrossRef\]](#)

Yu, Q., Shen, G., & Cao, M. (2017). Parameter estimation for Ornstein–Uhlenbeck processes of the second kind driven by  $\alpha$ -stable Lévy motions. *Communications in Statistics-Theory and Methods*, 46(21), 10864–10878. [\[CrossRef\]](#)

Zeng, Z., & Lee, C. G. (2014). Pairs trading: Optimal thresholds and profitability. *Quantitative Finance*, 14(11), 1881–1893. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.