Discussion Papers in Economics

No. 10/2017

Pairs trading with a mean-reverting jump-diffusion model on high-frequency data

Johannes Stübinger
University of Erlangen-Nürnberg

Sylvia Endres
University of Erlangen-Nürnberg

ISSN 1867-6707

Friedrich-Alexander-Universität Erlangen-Nürnberg
Institute for Economics
https://www.iwf.rw.fau.de/research/iwf-discussion-paper-series/

Pairs trading with a mean-reverting jump-diffusion model on high-frequency data

Johannes Stübinger[a,1], Sylvia Endres[a,1]

[a] *University of Erlangen-Nürnberg, Department of Statistics and Econometrics, Lange Gasse 20, 90403 Nürnberg, Germany*

Wednesday 3rd May, 2017

Abstract

This paper develops a pairs trading framework based on a mean-reverting jump-diffusion model and applies it to minute-by-minute data of the S&P 500 oil companies from 1998 to 2015. The established statistical arbitrage strategy enables us to perform intraday and overnight trading. Essentially, we conduct a 3-step calibration procedure to the spreads of all pair combinations in a formation period. Top pairs are selected based on their spreads' mean-reversion speed and jump behavior. Afterwards, we trade the top pairs in an out-of-sample trading period with individualized entry and exit thresholds. In the back-testing study, the strategy produces statistically and economically significant returns of 60.61 percent p.a. and an annualized Sharpe ratio of 5.30, after transaction costs. We benchmark our pairs trading strategy against variants based on traditional distance and time-series approaches and find its performance to be superior relating to risk-return characteristics. The mean-reversion speed is a main driver of successful and fast termination of the pairs trading strategy.
Keywords: Finance, statistical arbitrage, pairs trading, high-frequency data, jump-diffusion model, mean-reversion.

Email addresses: `johannes.stuebinger@fau.de` (Johannes Stübinger), `sylvia.endres@fau.de` (Sylvia Endres)

[1]The authors have benefited from many helpful discussions with Ingo Klein, Christopher Krauß, and Jonas Rende.

1. Introduction

Pairs trading is a relative-value arbitrage strategy which has been emerged by a quantitative group at Morgan Stanley in the 1980s (Vidyamurthy 2004). The strategy identifies pairs of stocks whose prices move together historically. Upon divergence, go long in the undervalued stock and go short in the overvalued stock. If history repeats itself, prices converge to their historical equilibrium and a profit can be collected. The seminal paper of Gatev et al. (2006) reports average annualized excess returns of 11 percent for U.S. CRSP securities from 1962 until 2002. Ever since this publication, academical interest in statistical arbitrage pairs trading has surged. Key contributions are provided by Vidyamurthy (2004), Elliott et al. (2005), Do and Faff (2010), Avellaneda and Lee (2010), Rad et al. (2016), and Liu et al. (2017).

Krauss (2017) identifies five streams of pairs trading research – among them is the time-series approach which focuses on mean-reverting spreads. Meaningful representatives are Elliott et al. (2005), Bertram (2009, 2010), Avellaneda and Lee (2010), Ekström et al. (2011), Cummins and Bucca (2012), Bogomolov (2013), Zeng and Lee (2014), Göncü and Akyıldırım (2016a), and Liu et al. (2017). These studies use an Ornstein-Uhlenbeck (OU) process for modeling the price spread between two stocks. Research studies on the time-series approach either focus on discussing theoretical frameworks or center the development of a trading algorithm.

Elliott et al. (2005) provide an analytic framework by describing the spread with a mean-reverting Gaussian Markov chain model, observed in Gaussian noise. Essential parts of this method are the state equation, in which the state variable follows a mean-reverting process, and the observation equation, the sum of some Gaussian noise and the state variable. Bertram (2009, 2010) and Zeng and Lee (2014) identify the optimal trading levels by maximizing the expected rate of profit. The optimal stopping problem for pairs trading is formulated and explicitly solved by Ekström et al. (2011). In Göncü and Akyıldırım (2016a), the optimal entry and exit signals are derived by maximizing the probability of successful termination of the pairs trading strategy.

Avellaneda and Lee (2010) describe relative-value models based on the OU process and conduct a back-testing framework on U.S. equities from 1997 to 2007. In Cummins and

Bucca (2012), the model of Bertram (2010) is applied to oil stocks of NYMEX and ICE from 2003 to 2010. Bogomolov (2013) extends the method of renko and kagi constructions to pairs trading spread processes, shows their theoretical profitability for the OU process, and examines the strategy performance on the American and Australian stock exchanges from 1996 until 2011. Recently, Liu et al. (2017) introduce a doubly mean-reverting process based on conditional modeling to describe spreads. For empirical study, the authors opt for oil stocks of NYSE and NASDAQ from June 2013 to April 2015 and in 2008.

Given the available literature, financial data are exposed to more than only one source of uncertainty – an obvious deficit of the OU process, where low-probability large-amplitude variations are attributed to a Gaussian framework (Barlow 2002, Carr et al. 2002, Cont and Tankov 2003, Cartea and Figueroa 2005, Meyer-Brandis and Tankov 2008, Jing et al. 2012, Jondeau et al. 2015). In consequence, modeling high-frequency dynamics with an OU process leads to unreasonable parameter estimations and disregarding of stylized facts, e.g., fat tails. This drawback is eliminated by extending the OU process with a jump term, which drives uncertainty in addition to the diffusive component (Cartea et al. 2015), creating a jump-diffusion model. Merton (1976, 1992) introduces the class of jump-diffusion models to explain stock price dynamics. It is surprising that there are only two academic studies in the context of statistical arbitrage pairs trading which generalize the OU process to allow jumps. Larsson et al. (2013) conduct an initial abstraction by formulating an optimal stopping theory. Göncü and Akyıldırım (2016b) introduce a stochastic model for daily commodity pairs trading where the noise term is driven by a Lévy-process.

We enhance the existing research in several aspects. First, our manuscript contributes to the literature by introducing a pairs selection and trading strategy based on a jump-diffusion model (JDM) in the context of high-frequency data. The existence of jumps is confirmed by a preliminary analysis on the oil sector of the S&P 500 constituents from 1998 to 2015. We construct a statistical arbitrage framework which is able to capture jumps, mean-reversion, volatility cluster, and drifts. Specifically, the spread dynamics are handled using a 3-step calibration procedure. For the reason of considering the effects of jumps during the night, our strategy is able to perform both intraday and overnight trading. Second, we benchmark our strategy based on a JDM to well-established quantitative trading strategies. Among pairs trading with the distance approach, this paper checks the performance of a strategy which models the spread with a simple OU process. Thus, we are in position to evaluate the additional benefit of regarding jumps in the context of pairs trading. The comparison of several pairs trading strategies represents a novelty in academic research with high-frequency data. Third, we conduct a large-scale empirical study on the oil companies of the S&P 500 based on minute-by-minute data from January 1998 until December 2015. The vast majority of studies about pairs trading with stochastic differential equations use daily data – a clear drawback within the framework of stochastic processes in continuous time (Bertram 2009, 2010, Avellaneda and Lee 2010, Cummins and Bucca 2012, Bogomolov 2013, Zeng and Lee 2014, Göncü and Akyıldırım 2016a). Sole exception is provided by Liu et al. (2017) who apply frequencies of 5 minutes. We present the first academic study on pairs trading using a minute-by-minute frequency over a sample period of 18 years. We find out that our strategy based on a JDM achieves statistically and economically significant returns of 60.61 percent p.a., after transaction costs. The results are far superior compared to the benchmark strategies ranging from 1.76 percent p.a. for a naive buy-and-hold strategy of the S&P 500 index to 50.45 percent p.a. for a strategy based on the OU process. In contrast to traditional pairs trading, our strategy is not adversely affected by consistently negative performance in the recent years of our sample. Fourth, we analyze the effects of strong exposure to the mean-reverting process. Spreads exhibiting a high mean-reversion speed are found to generate the best performance. Thus, we confirm the assumption that the mean-reversion speed is a main driver of the achieved returns.

The remainder of this paper is organized as follows. Section 2 briefly depicts data and software used in this study. Section 3 describes the methodology and section 4 provides the study design. In section 5, we present our results and discuss key findings in light of the existing literature. Finally, section 6 concludes and summarizes directions for further research.

2. Data and Software

For our empirical application, we opt for minute-by-minute data of the S&P 500 from January 1998 to December 2015. This highly liquid subset consists of the leading 500 companies in the U.S. stock market, covering approximately 80 percent of available market capitalization (S&P 500 Dow Jones Indices 2015). The data set serves as a crucial test for any potential capital market anomaly, given intense analyst reporting and high investor investigation. Following Krauss and Stübinger (2017), we conduct a 2-stage process with the objective of eliminating survivor bias from our data base. First, a daily constituent list for the S&P 500 from January 1998 to December 2015 is transformed into a binary matrix, indicating whether the stock is a constituent of the index in the present day or not. Second, for all stocks having ever been a constituent of the index, we download minute-by-minute data from QuantQuote (QuantQuote 2016). The data is adjusted for dividends, stock splits, and further corporate actions. By applying these two steps, we get the constituency for the S&P 500 and the respective prices over time.

The entire methodology and all relevant analyses are implemented in the programming language `R` (R Core Team 2017). Table 1 lists the additional packages for dependence modeling, data handling, and financial modeling.

Application	**`R` package**	**Authors of the `R` package**
Dependence modeling	`fBasics`	Rmetrics Core Team et al. (2014)
	`lmtest`	Zeileis and Hothorn (2002)
	`MASS`	Venables and Ripley (2002)
	`rootSolve`	Soetaert (2009)
Data handling	`dplyr`	Wickham and Francois (2016)
	`readr`	Wickham et al. (2016)
	`readxl`	Wickham (2016)
	`texreg`	Leifeld (2013)
	`xlsx`	Dragulescu (2014)
	`xts`	Ryan and Ulrich (2014)
	`zoo`	Zeileis and Grothendieck (2005)
Financial modeling	`PerformanceAnalytics`	Peterson and Carl (2014)
	`QRM`	Pfaff and McNeil (2016)
	`quantmod`	Ryan (2016)
	`sandwich`	Zeileis (2006)
	`timeSeries`	Rmetrics Core Team et al. (2015)
	`tseries`	Trapletti and Hornik (2017)
	`TTR`	Ulrich (2016)

Table 1: `R` packages used in this paper for dependence modeling, data handling, and financial modeling.

3. Methodology

3.1. Jump-diffusion model

Pairs trading strategies aim at identifying pairs of stocks that follow an equilibrium relationship which can be achieved by focusing on mean-reverting spreads. Mathematically, the spread at time t is defined by

$$X_t = \ln\left(\frac{S_A(t)}{S_A(0)}\right) - \ln\left(\frac{S_B(t)}{S_B(0)}\right), \quad t \geq 0, \tag{1}$$

where $S_A(t)$ and $S_B(t)$ denote the prices of stocks A and B at time t. The OU process is one of the most well-known processes capturing the effect of mean-reversion, modeling the spread $\{X_t\}_{t\geq 0}$ by the following stochastic differential equation:

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t, \quad X_0 = x, \tag{2}$$

where $\theta \in \mathbb{R}^+$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, and the standard Brownian motion $\{W_t\}_{t\geq 0}$. The mean-reversion speed θ measures the degree of reversion to the equilibrium level μ, i.e., the higher the value θ is, the faster the process X_t tends back to its mean level. The OU process of equation (2) can be explicitly solved resulting in

$$X_t = xe^{-\theta t} + \mu(1 - e^{-\theta t}) + \sigma \int_0^t e^{-\theta(t-s)} dW_s.$$

Incorporating a non constant mean-reversion level $\mu(t)$ induces a time-dependent OU process. The spread X_t randomly fluctuates around the deterministic drift function $\mu(t)$:

$$dX_t = \theta(\mu(t) - X_t)dt + \sigma dW_t, \quad X_0 = x. \tag{3}$$

The solution to the stochastic differential equation (3) is given by

$$X_t = xe^{-\theta t} + \theta \int_0^t \mu(u) e^{-\theta(t-u)} du + \sigma \int_0^t e^{-\theta(t-s)} dW_s.$$

The above described OU models with their continuous paths are unlikely to produce large movements of the underlying process over a short time-period. To explain discontinuous spread variations, the OU model is extended to account for jumps in addition to the simple Gaussian shocks. Integrating a jump term into equation (3) leads to a mean-reverting JDM:

$$dX_t = \theta(\mu(t) - X_t)dt + \sigma dW_t + \ln J dN_t, \quad X_0 = x, \tag{4}$$

where $\{N_t\}_{t\geq 0}$ is a Poisson process, creating jumps at frequency $\lambda(t)$. Discontinuous path changes with randomly arriving jumps at random jump size are captured. The frequency is chosen time-dependent to account for variations in the jump occurrence. Villaplana (2003), Seifert and Uhrig-Homburg (2007), and Escribano et al. (2011) present academic studies allowing for non-constant jump intensities. In our model, the probability that a jump happens intraday is assumed to be zero. Overnight, jumps occur randomly at probability λdt. Therefore, the last component in equation (4) affects spreads only overnight. We define the varying intensity $\lambda(t)$ such that

$$\lambda(t) = \begin{cases} 0 & \text{if the observation is intraday} \\ \lambda & \text{otherwise (overnight, weekend).} \end{cases}$$

For the overnight variations it is

$$dN_t = \begin{cases} 1 & \text{with probability } \lambda dt \\ 0 & \text{with probability } 1 - \lambda dt \end{cases}$$

and intraday $dN_t = 0$ with probability 1. J is a random variable modeling the magnitudes of the jumps. Supposing $\ln J \sim \mathcal{N}(\mu_J, \sigma_J^2)$, we apply a typical assumption on the jump size distribution (Cartea and Figueroa 2005, Benth et al. 2012). Despite the jump component, the mean-reverting nature of the model is still present: After a jump, the spread does not stay in the new level, but reverts back to the equilibrium level with a speed determined by the parameter θ.

An adjusted form of equation (4) is

$$\begin{aligned} X_t &= g(t) + Y_t \\ dY_t &= -\theta Y_t dt + \sigma dW_t + \ln J dN_t, \end{aligned} \tag{5}$$

where the spread X_t is represented by a deterministic drift function $g(t)$, modeling mean variations of the spread evolution, and a stochastic process Y_t, reverting around zero. The time dependent mean-reversion level $\mu(t)$, introduced in equation (4), depends on the drift function $g(t)$ in the following way (Lucia and Schwartz 2002, Cartea and Figueroa 2005):

$$\mu(t) = \frac{1}{\theta}\frac{dg(t)}{dt} + g(t). \tag{6}$$

The solution of equation (5) is given by

$$X_t = g(t) + (x - g(0))e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)} dW_s + \int_0^t e^{-\theta(t-s)} \ln J dN_s.$$

In our empirical application, we use equation (5) for modeling the dependence structure of pairs of stocks, opting for high-frequency data with an interval length of one minute. Per pair, we observe 391 spread values each day during trading hours from 9:30 am to 4:00 pm. Following Liu et al. (2017), we denote the discretized observations of the spread process X_t on day i as

$$X_{391(i-1)+1}, X_{391(i-1)+2}, ..., X_{391i}, \quad i = 1, 2, \ldots, I$$

where I is the number of considered days. The overnight variation from the market's close of day i until the market's open of the next trading day $i+1$ is described by

$$X_{391i+1} - X_{391i}, \quad i = 1, 2, \ldots, I-1.$$

3.2. Preliminary analysis

The majority of academic research in the pairs trading context aims at capturing spread processes with strong mean-reversion by neglecting any jumps, e.g., Avellaneda and Lee (2010) and Liu et al. (2017). In the following preliminary analysis, we examine the existence of jumps in our data set – the results justify the selection of the JDM clearly. Therefore, our data set from 1998 to 2015 is portioned into disjoint sub-periods with a length of 40 days. For each sub-period, we regard the companies of the oil sector and determine the spreads of all possible pair combinations. The absolute first differences of each spread are splitted into the subsets overnight variations and intraday variations. For the overnight and intraday variations, we consider the highest $1-q$ variations for $q \in \{0.90, 0.95, 0.97, 0.99, 0.999\}$. The choice of q is inspired by Meyer-Brandis and Tankov (2008) who propose a method for spike detection where they remove a percentage of 5 percent of the highest absolute returns.

Table 2 depicts characteristics of the conditional distributions for varying q. The mean of the overnight variations ranges from 0.0251 to 0.1351 – high values compared to an interval from 0.0033 to 0.0169 for the intraday variations. This picture barely changes considering the probability mass of extreme values. In contrary, the maximum intraday variation (0.4708) exceeds the greatest overnight variation (0.3942) caused by the fact that we have a 390 times greater data base in the intraday context.

Now, the jump threshold c_q ($c_q \in \mathbb{R}^+$) is calculated based on the q-quantile of the whole data base of overnight and intraday variations together. Following Meyer-Brandis and

Tankov (2008), we estimate the jump intensity λ by the following:

$$\lambda_{\text{overnight (intraday)}} = \frac{\text{number of overnight (intraday) variations greater than } c_q}{\text{total number of overnight (intraday) variations}}.$$

Across all jump thresholds, the jump intensity overnight is clearly higher than intraday, such as, regarding the highest 0.1 percent values leads to a jump intensity of 11.83 percent for the overnight variations and 0.07 percent for the intraday variations. Our preliminary results are well in line with the literature. Jondeau et al. (2015) estimate a model containing jumps and apply it to data at tick frequency. On average, their model explains 47.7 percent of the total variation of stock returns, split into continuous innovations, intraday jumps and overnight returns. About 7 percent of those variations are represented by overnight returns, which is a substantial part in the context of tick-by-tick data.

In summary, we confirm the statements of the literature that both including a jump component is necessary and disregarding intraday jumps takes place without any significant interference.

	Overnight variations					Intraday variations				
Quantile	90%	95%	97%	99%	99.9%	90%	95%	97%	99%	99.9%
Mean	0.0251	0.0338	0.0415	0.0627	0.1351	0.0033	0.0044	0.0053	0.0078	0.0169
Minimum	0.0139	0.0199	0.0252	0.0401	0.0941	0.0019	0.0027	0.0033	0.0052	0.0111
Quartile 1	0.0162	0.0227	0.0287	0.0449	0.1065	0.0022	0.0030	0.0038	0.0057	0.0123
Median	0.0199	0.0273	0.0340	0.0525	0.1233	0.0027	0.0036	0.0044	0.0066	0.0141
Quartile 3	0.0273	0.0366	0.0449	0.0682	0.1486	0.0036	0.0047	0.0057	0.0082	0.0179
95% Quantile	0.0525	0.0682	0.0820	0.1233	0.2168	0.0066	0.0082	0.0097	0.0141	0.0311
99% Quantile	0.0941	0.1233	0.1435	0.1770	0.3462	0.0111	0.0141	0.0168	0.0247	0.0527
Maximum	0.3942	0.3942	0.3942	0.3942	0.3942	0.4708	0.4708	0.4708	0.4708	0.4708
Standard deviation	0.0169	0.0205	0.0234	0.0306	0.0450	0.0022	0.0028	0.0033	0.0047	0.0106
Jump intensity λ	0.6433	0.5655	0.5028	0.3610	0.1183	0.0986	0.0487	0.0288	0.0091	0.0007

Table 2: Preliminary analysis of the relevant S&P 500 data base from 1998 to 2015 regarding spread variations overnight and intraday.

4. Study design

For our back-testing application, we follow Liu et al. (2017) and decide on the oil sector of the S&P 500 constituents from January 1998 to December 2015 (section 2). Following Jegadeesh and Titman (1993) and Gatev et al. (1999, 2006), we divide the data set into

4484 overlapping study periods (figure 1). Each study period is shifted by one day and consists of a 40-day formation period (subsection 4.1) and a 5-day trading period (subsection 4.2). The length of the formation period is consistent with Liu et al. (2017) – the first 10 days of the formation period are used for determining the jump threshold. For our intraday and overnight trading strategy, the length of the trading period follows Bowen et al. (2010). Typically, there are 32 oil companies member of the S&P 500 and each stock contains 391 minute-by-minute data points per day. Consequently, approximately $4484 \cdot 45 \cdot 32 \cdot 391 = 2,524,671,360$ stock prices are handled during one simulation run from January 1998 to December 2015.

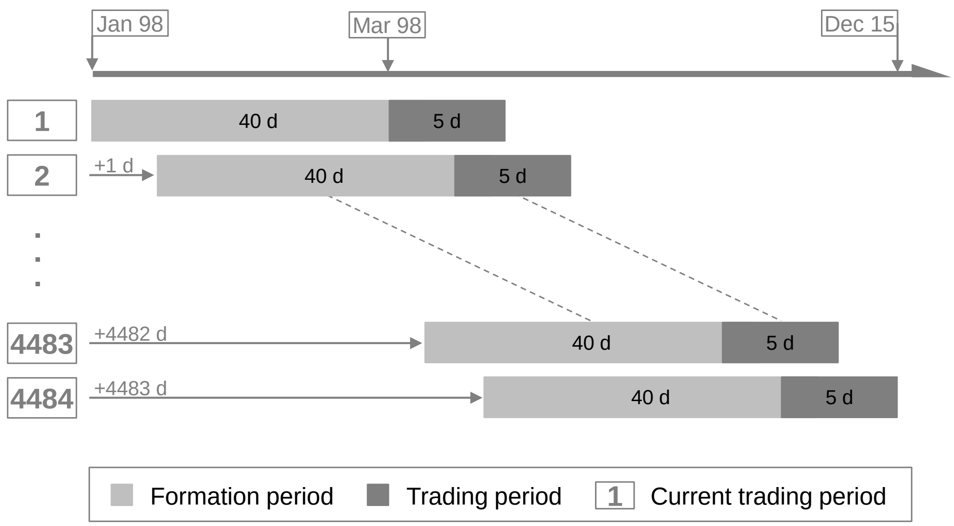


Figure 1: The back-testing application deals with 4484 overlapping study periods from January 1998 to December 2015. Each study period consists of a 40-day formation and a 5-day out-of-sample trading period.

4.1. Formation period

In the 40-day formation period T_{for}, we fit models from the type of equation (5) to all possible combinations of pairs. Therefore, we follow a 3-step calibration procedure by (i) extracting jumps, (ii) adjusting drifts, and (iii) estimating parameters. This subsection describes the 3-step logic outlined above in detail.

In the first step, we apply a threshold method for detecting and filtering the jumps. Overnight spread variations above a fixed threshold are considered to be caused by jumps. According to Meyer-Brandis and Tankov (2008), this is the most common way to separate the continuous part of a jump-diffusion process from discontinuous variations. However, the procedure is not sensitive to outliers. In the spirit of Cartea and Figueroa (2005) and Meyer-Brandis and Tankov (2008), we extract some percentage of returns with highest absolute value. The standard deviation of the remaining returns corresponds to the noise level of common fluctuations which is smaller than the standard deviation of the original series. Specifically, the formation period is divided into a 10-day initialization period and a 30-day out-of-sample training period. We calculate mean μ and standard deviation σ of all overnight variations based on the initialization period. In the remaining training period, absolute returns greater than $\mu + k\sigma$ are identified as jumps ($k \in \mathbb{R}^+$). We receive the jump adjusted time series X_t by filtering out the identified jumps from the original series.

In the second step, we align the adapted spread series by a time-varying drift adjustment. In the spirit of Kim (2003) and Ng et al. (2003), the spread at time t ($t \in T_{for}$) is subtracted by the running mean of the past 1955 minutes (5 days) resulting in the drift adjusted time series X_t.

In the third step, we estimate the parameters of the remaining process using maximum likelihood estimation. The discretization of the OU process, now reverting around zero, at time t is given by

$$X_{t+1} = X_t e^{-\theta\delta} + \sigma\sqrt{\frac{1 - e^{-2\theta\delta}}{2\theta}} Z_t, \quad t = 1, ..., N$$

with time step δ, $Z_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$, and $N = 391 \cdot 30 - 1$. In the sense of Liu et al. (2017), we approximate the interval length from one observation to the next by effective time instead of real time. Specifically, it is assumed that the overnight periods are as long as the intraday periods. This seems reasonable as we already extract the jumps over night. Therefore, each period has the length $\delta = \frac{\delta_1}{391}$ with $\delta_1 = \frac{1}{250}$ representing one day. The conditional density of X_t satisfies

$$f(X_{t+1}|X_t; \theta, \sigma) = \frac{1}{\tilde{\sigma}\sqrt{2\pi}} \exp\left(-\frac{(X_{t+1} - X_t e^{-\theta\delta})^2}{2\tilde{\sigma}^2}\right)$$

with

$$\tilde{\sigma} = \sigma\sqrt{\frac{1 - e^{-2\theta\delta}}{2\theta}}.$$

The corresponding log-likelihood function is given by

$$\mathcal{L}(\theta, \sigma; X_1, \dots, X_{N+1}) = \sum_{t=1}^{N} \ln f(X_{t+1}|X_t; \theta, \sigma)$$
$$= -\frac{N}{2}\ln(2\pi) - N\ln(\tilde{\sigma}) - \frac{1}{2\tilde{\sigma}^2}\sum_{t=1}^{N}\left(X_{t+1} - X_t e^{-\theta\delta}\right)^2.$$

The parameters θ and σ are estimated using the limited memory algorithm for bound constrained optimization by Byrd et al. (1995). We transfer the top p pairs ($p \in \mathbb{N}$) exhibiting both a high mean-reversion speed and a high number of overnight jumps to the trading period. This procedure relies on the assumption that overnight jumps create trading opportunities in addition to the Gaussian fluctuations, while high speed of mean-reversion pull the process back to its equilibrium level, where pairs trading profits are taken.

4.2. Trading period

The top pairs are transferred to the 5-day trading period T_{tra} and every newly incoming price on time t ($t \in T_{tra}$) is used to calculate the spread X_t outlined in equation (1). We suppose that spreads with a minimum of overnight variation and strong mean-reversion lead to desired performance results. If our assumption holds and the 3-step calibration procedure is feasible over time, we are in position to take permanent advantage of temporary mispricings. We strive to capture these characteristics with a trading strategy on the basis of Bollinger bands (Bollinger 1992, 2001).

For obtaining the Bollinger bands, we calculate moving mean and standard deviation of the spread series of the past 1955 minutes (see subsection 4.1). Specifically, we determine the moving mean $\mu(t)$ using equation equation (6), where $g(t)$ is estimated by the simple moving average. The moving standard deviation $\sigma(t)$ is calculated by the simple running standard deviation[2]. By adding (subtracting) k-times the running standard deviation $\sigma(t)$ to (from) the mean level $\mu(t)$, we construct the upper and lower band $\mu(t) \pm k\sigma(t)$ ($k \in \mathbb{R}^+$).

We define the following trading entry signals:

[2]Calculating the variance of the JDM analytically would require estimating the parameters of the jump size distribution. Dealing with a data base of only 30 days does not allow reasonable estimates.

- $X_t < \mu(t) - k\sigma(t)$, i.e., stock 1 is undervalued and stock 2 overvalued. Consequently, we simultaneously go long in stock 1 and go short in stock 2.

- $X_t > \mu(t) + k\sigma(t)$, i.e., stock 1 is overvalued and stock 2 undervalued. Consequently, we simultaneously go short in stock 1 and go long in stock 2.

- $\mu(t) - k\sigma(t) \leq X_t \leq \mu(t) + k\sigma(t)$, i.e., the spread does not exhibit any mispricings. Consequently, we do not execute any trades.

Upon every entry signal we buy 1 dollar worth of the undervalued stock and short 1 dollar worth of the overvalued stock. Further entry signals are neglected until the position is closed, so that only one active position per pair is allowed. Trades are held until the spread reverts back to equilibrium, i.e., crosses the time-varying mean level. We also exit the trade when the trading period ends or if one of the stocks of the respective pair is delisted.

According to Miao (2014) and Krauss et al. (2017), we focus on a portfolio consisting of the top $p = 10$ pairs. For constructing the Bollinger bands, we follow the vast majority of literature and set $k = 2$, a value suggested by Bollinger (1992, 2001) and applied by Avellaneda and Lee (2010), Clegg and Krauss (2016), and Stübinger et al. (2016).

Return computation follows Gatev et al. (2006). Specifically, we relate the sum of daily payoffs across all pairs to the sum of invested capital at the end of the previous day. We show both the return on committed capital (invest one dollar for each pair) and the return on employed capital (invest one dollar for each active pair). Following Avellaneda and Lee (2010), Stübinger et al. (2016), and Liu et al. (2017), we depict transaction costs of 5 bps per share per half-turn. This procedure is feasible in light of our high-frequency data in a highly liquid investment universe.

To assess the additional benefit of our JDM-based strategy, we benchmark it with pairs trading variants based on the (i) classic distance model (CDM), (ii) Bollinger bands model (BBM), (iii) Ornstein-Uhlenbeck model (OUM), and a (iv) S&P 500 buy-and-hold strategy (MKT) – all well-established quantitative strategies in the literature. Data and general framework are set identical to the JDM. In the following, we depict the key facts of the four benchmarks.

Classic distance model (CDM). With the seminal paper of Gatev et al. (2006), interest for pairs trading has surged in the academic community. We follow their approach and implement the strategy on our circumstances. Pairs are determined possessing the smallest sum of squared deviations between normalized prices during the formation period. In the subsequent trading period, pairs are opened if the spread diverges more than two standard deviations in absolute value. The trade is closed at the next crossing of prices. For further details about this approach, see Gatev et al. (2006) and Do and Faff (2010, 2012).

Bollinger bands model (BBM). For the second benchmark, we enlarge the CDM using time-varying trading thresholds in the spirit of Bollinger (1992, 2001). Again, we select the pairs with the minimal sum of squared deviations. The fixed trading thresholds of Gatev et al. (2006) are replaced by time-varying entry and exit signals. Specifically, the upper (lower) Bollinger band are determined by adding (subtracting) 2-times the running standard deviation to (from) the running mean. We calculate the running ratios of the past 1955 minutes to be in accord with subsection 4.2. By using Bollinger bands, we aim to capture drifts and volatility clusters – typical characteristics of financial time series (Ou and Penman 1989, Lux and Marchesi 2000, Cont 2007).

Ornstein-Uhlenbeck model (OUM). In spirit of Elliott et al. (2005), Avellaneda and Lee (2010), and Göncü and Akyıldırım (2016a), the dynamics of the spread are described by a mean-reverting OU process. Similar to the JDM, we select the pairs based on the highest mean-reversion speed and the highest variance. The second selection criterion is motivated by Liu et al. (2017), who aim at capturing volatile intraday movements and thus many trading opportunities by a high short-term variance. Trading thresholds are identical to the JDM. Summarizing, the OUM is a reduced version of the JDM with the deficit of being not able to capture overnight price changes (Kappou et al. 2010).

S&P 500 buy-and-hold strategy (MKT). Last but not least, we compare our JDM to a naive S&P 500 buy-and-hold strategy. We buy the S&P 500 index in March 1998 and hold it during the complete trading period. This passive investment strategy runs regardless of any market conditions.

5. Results

We follow Krauss and Stübinger (2017) and conduct a fully-fledged performance evaluation on the JDM from March 1998 to December 2015 – compared to the CDM, BBM, OUM, and MKT. The key results for the top $p = 10$ pairs are depicted in two panels – before and after transaction costs. First, we analyze the performance of all strategies (subsection 5.1) and execute a sub-period analysis (subsection 5.2). The majority of the used performance metrics is regarded by Bacon (2008). Second, we investigate the exposure to common systematic risk factors (subsection 5.3) and check the robustness of the JDM (subsection 5.4). Finally, we examine the influence of the mean-reversion speed on the performance of the JDM (subsection 5.5).

5.1. Strategy performance

Table 3 reports daily return characteristics and corresponding risk metrics for the top 10 pairs per strategy from March 1998 until December 2015. We observe statistically significant returns for the CDM, BBM, OUM, and JDM, with Newey-West (NW) t-statistics above 11.76 before transaction costs and above 7.00 after transaction costs. This picture barely changes considering the economic perspective – the mean of daily returns varies from 0.04 percent for the CDM to 0.19 percent for the JDM after transaction costs. The return distribution of the JDM achieves right skewness – following Cont (2001) a desired characteristic for any investor. In line with Mina and Xiao (2001), we report historical Value at Risk (VaR) measures. Tail risk after transaction costs is greatly reduced for the JDM in contrast to the general market, e.g., the historical VaR (1%) is -1.49 percent for the JDM versus -3.50 percent for the buy-and-hold strategy. The maximum drawdown confirms this statement – the decline from a historical peak is at a very low level for the JDM (15.08%), compared to the CDM (20.29%), BBM (30.95%), OUM (65.47%), and MKT (64.33%). Also, the hit rate of the JDM outperforms clearly with approximately 64 percent after transaction costs. Summarizing, the JDM achieves convincing return characteristics and risk metrics – this statement remains true even after transaction costs. We have to survey the robustness of this strategy to systematic sources of risk.

	Before transaction costs				After transaction costs				
	CDM	BBM	OUM	JDM	CDM	BBM	OUM	JDM	MKT
Mean return	0.0007	0.0015	0.0025	0.0027	0.0004	0.0009	0.0017	0.0019	0.0001
Standard error (NW)	0.0001	0.0001	0.0002	0.0001	0.0001	0.0001	0.0002	0.0001	0.0002
t-Statistic (NW)	11.7621	16.5427	12.6959	20.1463	7.0059	10.1834	8.8740	14.6688	0.8889
Minimum	-0.0203	-0.1706	-0.1560	-0.0442	-0.0209	-0.1719	-0.1563	-0.0447	-0.0947
Quartile 1	-0.0005	-0.0003	-0.0018	-0.0008	-0.0007	-0.0008	-0.0025	-0.0015	-0.0056
Median	0.0004	0.0011	0.0024	0.0024	0.0002	0.0006	0.0017	0.0016	0.0005
Quartile 3	0.0015	0.0029	0.0071	0.0057	0.0012	0.0022	0.0063	0.0048	0.0061
Maximum	0.0408	0.0471	0.1034	0.0983	0.0388	0.0462	0.1012	0.0942	0.1096
Standard deviation	0.0024	0.0043	0.0112	0.0070	0.0024	0.0042	0.0110	0.0068	0.0126
Skewness	2.6553	-13.0216	-1.3027	1.4535	2.4223	-14.3790	-1.4063	1.3595	-0.1983
Kurtosis	30.5440	571.4109	22.6129	15.1016	28.8833	641.5760	23.4490	14.9369	7.5250
Historical VaR 1%	-0.0048	-0.0055	-0.0289	-0.0142	-0.0052	-0.0060	-0.0298	-0.0149	-0.0350
Historical CVaR 1%	-0.0066	-0.0121	-0.0503	-0.0192	-0.0069	-0.0126	-0.0510	-0.0198	-0.0506
Historical VaR 5%	-0.0024	-0.0027	-0.0125	-0.0070	-0.0027	-0.0032	-0.0132	-0.0077	-0.0197
Historical CVaR 5%	-0.0039	-0.0055	-0.0245	-0.0116	-0.0042	-0.0059	-0.0252	-0.0123	-0.0302
Maximum drawdown	0.0471	0.1706	0.2703	0.1083	0.2029	0.3095	0.6547	0.1508	0.6433
Share with return > 0	0.6258	0.6895	0.6530	0.6992	0.5612	0.6095	0.6175	0.6412	0.5306

Table 3: Daily return characteristics and risk metrics for the top 10 pairs of the CDM, BBM, OUM, and JDM, compared to a S&P 500 long-only benchmark (MKT) from March 1998 until December 2015. NW denotes Newey-West standard errors with five-lag correction and CVaR the Conditional Value at Risk.

Table 4 depicts summary statistics on trading frequency. The number of actually traded pairs is vastly different for the CDM (6.13), compared to the BBM, OUM, JDM (above 9.47) – this dissimilarity is originated by the two different trading strategies. Higher number of tradings generate increasing transaction costs, such as, the difference of daily mean return before and after transaction costs amounts to 0.03 percentage points for the CDM versus 0.08 percentage points for the JDM (table 3). The trade duration of approximately 2 days across all systems illustrates the importance of considering overnight effects in financial data (Kappou et al. 2010). The half of the pairs have to be closed at the end of the trading period. Specifically, for the CDM, 5.14 pairs are closed of necessity – a value similar to the finding of Clegg and Krauss (2016) (10.86 out of 20 pairs).

	CDM	BBM	OUM	JDM
Average number of pairs traded per 5-day period	6.1321	9.4795	9.8173	9.8450
Average number of round-trip trades per pair	1.1266	1.7162	1.9533	2.0956
Standard deviation of number of round-trip trades per pair	0.4453	1.1236	1.1162	1.4975
Average time pairs are open in days	2.3403	1.7755	1.9216	1.8471
Standard deviation of time open, per pair, in days	1.4277	1.5153	1.3389	1.3146
Average number of pairs where closing is forced	5.1381	5.3913	6.0874	5.9871

Table 4: Trading statistics for the top 10 pairs of the CDM, BBM, OUM, and JDM per 5-day trading period.

Table 5 contains return characteristics and risk metrics of trades where closing is forced. As expected, the mean return of all strategies is negative before and after transaction costs – latter ranges from -0.45 percent for the CDM to -2.66 percent for the OUM. This finding is not surprising since we only close a trade during the trading period if the spread converges back to the supposed equilibrium level. We observe an asymmetry of the return distributions when closing is forced – across all strategies, the absolute value of the minimum is approximately two to four times higher than the maximum. Specifically, adding the jump component in our model seems to have a strong positive effect on the risk of loss – the minimum return of the JDM (-24.58 percent) is much greater than the minimum return of the OUM (-59.45 percent).

	Before transaction costs				After transaction costs			
	CDM	BBM	OUM	JDM	CDM	BBM	OUM	JDM
Mean return	-0.0025	-0.0070	-0.0247	-0.0177	-0.0045	-0.0090	-0.0266	-0.0197
Minimum	-0.5074	-1.0000	-0.5937	-0.2443	-0.5084	-1.0000	-0.5945	-0.2458
Median	-0.0008	-0.0024	-0.0142	-0.0105	-0.0028	-0.0044	-0.0162	-0.0125
Maximum	0.2877	0.2264	0.1853	0.1072	0.2851	0.2239	0.1829	0.1050
Standard deviation	0.0343	0.0326	0.0399	0.0277	0.0342	0.0326	0.0398	0.0277

Table 5: Return characteristics of trades where closing is forced for the top 10 pairs of the CDM, BBM, OUM, and JDM, from March 1998 until December 2015.

Table 6 summarizes annualized risk-return measures for all four strategies. The JDM achieves annualized returns of 98.55 percent before and 60.61 percent after transaction costs – classic pairs trading strategies and a naive buy-and-hold strategy are clearly outperformed. Modeling overnight jumps pays off – compared to the OUM, the JDM produces higher returns at approximately half the standard deviation, resulting in Sharpe ratios after transaction costs of 5.30 for the JDM and 2.71 for the OUM. The mean returns and Sharpe ratios for the BBM, OUM, and JDM show similar results for employed and committed capital – not surprising since the top pairs open in almost all cases (table 4).

	Before transaction costs				After transaction costs				
	CDM	BBM	OUM	JDM	CDM	BBM	OUM	JDM	MKT
Mean return	0.1784	0.4375	0.8324	0.9855	0.0990	0.2382	0.5045	0.6061	0.0176
Mean excess return	0.1549	0.4089	0.7960	0.9461	0.0771	0.2136	0.4746	0.5742	-0.0027
Standard deviation	0.0386	0.0685	0.1773	0.1116	0.0376	0.0668	0.1749	0.1084	0.2005
Downside deviation	0.0179	0.0460	0.1168	0.0514	0.0198	0.0477	0.1205	0.0558	0.1441
Sharpe ratio	4.0105	5.9734	4.4907	8.4750	2.0520	3.1986	2.7131	5.2982	-0.0136
Sortino ratio	9.9497	9.5119	7.1270	19.1603	4.9942	4.9894	4.1882	10.8676	0.1218
Mean return on employed capital	0.2862	0.4535	0.8518	0.9994	0.1512	0.2429	0.5164	0.6133	0.0176
Sharpe ratio on employed capital	4.7664	6.0554	4.5592	8.5225	2.3925	3.1874	2.7563	5.3154	-0.0136

Table 6: Annualized risk-return measures for the top 10 pairs of the CDM, BBM, OUM, and JDM, compared to a S&P 500 long-only benchmark (MKT) from March 1998 until December 2015.

5.2. Sub-period analysis

Do and Faff (2010), Bowen and Hutchinson (2015), and Krauss et al. (2017) report varying performance over time of their pairs trading strategies. Table 7 analyzes the annualized risk-return measures of the four strategies during sub-periods of 3 years.

The first period ranges from 1998 to 2000 and corresponds with the growth of the dot-com bubble. We observe that all strategies achieve much better results compared to the overall period in table 6. Specifically, the JDM and the OUM with annualized returns of 247.62 percent and 261.23 percent after transaction costs outperform clearly simple pairs trading. Returns are most likely driven by bid-ask bounces in consequence of fractional pricing during this time.

The second period ranges from 2001 to 2003 and includes the dot-com crash, the September 11 attacks and the start of the Iraq war. In contrast to the general market with annualized returns of -7.81 percent, strategy returns are still far above zero – even after transaction costs. We note that the JDM is the only strategy that reduces significantly its downside deviation in comparison to the previous sub-period, resulting in a Sortino ratio of 28.77.

The third period ranges from 2004 to 2006 and describes the time of moderation. Prices are leveling out reducing the standard deviation of the general market to 10.46 percent.

Since pairs trading takes profit from temporal deviations between the stocks, we may carefully conclude that pairs trading strategies have hardly a chance to perform during this financial recovery. We observe satisfying performance results, such as, annualized returns after transaction costs range from 58.03 percent for the OUM to 6.66 percent for the CDM.

The forth period ranges from 2007 to 2009 and is in accord with the global financial crisis. Contrary to the market, all strategies generate positive returns varying from 70.14 percent of the JDM to 13.70 for the CDM. This fact is not surprising since Do and Faff (2010), Krauss et al. (2017), and Liu et al. (2017) show that pairs trading outperforms during bear markets.

The fifth period ranges from 2010 to 2012 and specifies the time of deterioration. In contrast to the market with annualized returns of 6.71 percent, negative returns range from -2.36 percent for the JDM to -15.57 percent for the OUM. The JDM still achieves positive returns before transaction costs of 16.49 percent. In conclusion, the model detects structure, but the resulting profits are not large enough considering the impact of transaction costs.

The sixth period ranges from 2013 to 2015 and characterizes the period of comebacks. Across all strategies, the mean excess returns equal the mean returns because the daily risk free rate is zero during this sub-period. The market with annualized returns of 11.82 percent outperforms all strategies varying from 5.07 percent for the JDM to -15.21 percent for the OUM. The majority of academic research shows declining returns for the recent years, e.g., Clegg and Krauss (2016) and Krauss et al. (2017). However, the JDM produces positive mean returns prior and after transaction costs.

		Before transaction costs				After transaction costs				
		CDM	BBM	OUM	JDM	CDM	BBM	OUM	JDM	MKT
1998-2000	Mean return	0.4248	1.1973	3.5260	3.4726	0.3193	0.8545	2.6123	2.4762	0.0625
	Mean excess return	0.3554	1.0905	3.3065	3.2558	0.2550	0.7643	2.4370	2.3075	0.0107
	Standard deviation	0.0438	0.0705	0.2190	0.1475	0.0427	0.0681	0.2158	0.1434	0.2056
	Downside deviation	0.0166	0.0196	0.1069	0.0566	0.0184	0.0218	0.1109	0.0605	0.1443
	Sharpe ratio	8.1124	15.4669	15.0954	22.0789	5.9727	11.2152	11.2926	16.0888	0.0519
	Sortino ratio	25.5682	61.1825	32.9859	61.3774	17.3495	39.2157	23.5488	40.9134	0.4327
2001-2003	Mean return	0.2592	0.6743	1.3355	1.7033	0.1746	0.4199	0.8600	1.0794	-0.0781
	Mean excess return	0.2324	0.6387	1.2859	1.6459	0.1496	0.3897	0.8205	1.0353	-0.0978
	Standard deviation	0.0349	0.1128	0.2067	0.1052	0.0336	0.1119	0.2032	0.1000	0.2184
	Downside deviation	0.0146	0.1002	0.1447	0.0327	0.0162	0.1014	0.1477	0.0375	0.1538
	Sharpe ratio	6.6585	5.6607	6.2223	15.6511	4.4494	3.4829	4.0387	10.3494	-0.4478
	Sortino ratio	17.7391	6.7319	9.2320	52.0796	10.7860	4.1416	5.8236	28.7702	-0.5080
2004-2006	Mean return	0.1430	0.3961	0.9527	0.7613	0.0666	0.2028	0.5803	0.4241	0.0787
	Mean excess return	0.1099	0.3558	0.8964	0.7105	0.0358	0.1681	0.5347	0.3830	0.0475
	Standard deviation	0.0286	0.0437	0.1089	0.0790	0.0279	0.0418	0.1067	0.0765	0.1046
	Downside deviation	0.0160	0.0188	0.0616	0.0367	0.0178	0.0219	0.0659	0.0412	0.0720
	Sharpe ratio	3.8430	8.1335	8.2287	8.9896	1.2801	4.0237	5.0115	5.0081	0.4542
	Sortino ratio	8.9274	21.0538	15.4739	20.7273	3.7324	9.2449	8.8101	10.2841	1.0935
2007-2009	Mean return	0.2228	0.4830	0.9379	1.0620	0.1370	0.2789	0.6069	0.7014	-0.1177
	Mean excess return	0.1977	0.4526	0.8982	1.0198	0.1136	0.2526	0.5740	0.6665	-0.1358
	Standard deviation	0.0542	0.0610	0.1478	0.1268	0.0526	0.0586	0.1459	0.1242	0.2995
	Downside deviation	0.0223	0.0225	0.0835	0.0581	0.0242	0.0257	0.0878	0.0623	0.2209
	Sharpe ratio	3.6481	7.4212	6.0780	8.0404	2.1613	4.3118	3.9328	5.3656	-0.4534
	Sortino ratio	9.9951	21.4665	11.2390	18.2726	5.6629	10.8624	6.9143	11.2638	-0.5328
2010-2012	Mean return	0.0298	0.0710	-0.0040	0.1649	-0.0340	-0.0589	-0.1557	-0.0236	0.0671
	Mean excess return	0.0290	0.0702	-0.0047	0.1640	-0.0347	-0.0596	-0.1564	-0.0244	0.0663
	Standard deviation	0.0280	0.0432	0.1595	0.0847	0.0276	0.0422	0.1588	0.0835	0.1856
	Downside deviation	0.0178	0.0281	0.1269	0.0577	0.0201	0.0313	0.1310	0.0625	0.1341
	Sharpe ratio	1.0363	1.6241	-0.0297	1.9366	-1.2562	-1.4122	-0.9847	-0.2920	0.3572
	Sortino ratio	1.6690	2.5311	-0.0313	2.8572	-1.6882	-1.8826	-1.1889	-0.3778	0.5004
2013-2015	Mean return	0.0478	0.1072	0.0023	0.2571	-0.0196	-0.0331	-0.1521	0.0507	0.1182
	Mean excess return	0.0478	0.1072	0.0023	0.2571	-0.0196	-0.0331	-0.1521	0.0507	0.1182
	Standard deviation	0.0319	0.0398	0.1832	0.0902	0.0312	0.0381	0.1824	0.0889	0.1282
	Downside deviation	0.0191	0.0238	0.1501	0.0597	0.0211	0.0272	0.1538	0.0642	0.0905
	Sharpe ratio	1.4989	2.6953	0.0124	2.8503	-0.6293	-0.8683	-0.8342	0.5700	0.9222
	Sortino ratio	2.5055	4.5038	0.0151	4.3036	-0.9298	-1.2165	-0.9895	0.7888	1.3058

Table 7: Annualized risk-return measures for the top 10 pairs of the CDM, BBM, OUM, and JDM, compared to a S&P 500 long-only benchmark (MKT) for sub-periods of 3 years from March 1998 until December 2015.

5.3. Common risk factors

Table 8 explores the systematic risk exposure for the top 10 pairs of the JDM after transaction costs. We follow Krauss and Stübinger (2017) and conduct three types of regression. First, we capture the return anomalies by the three-factor model (FF3) of Fama and French (1996). The model measures the sensitivity to the overall market, small minus big capitalization stocks (SMB), and high minus low book-to-market stocks (HML). Second, we depict the Fama-French 3+2 factor model (FF3+2) in line with Gatev et al. (2006). It augments the baseline model by the additional factors momentum and short-term reversal. Third, we extend the first model by adding two factors, i.e., portfolios of stocks with robust minus weak profitability (RMW) and with conservative minus aggressive (CMA) investment behavior. According to Fama and French (2015), we call this model Fama-French five-factor model (FF5). We download the data related to these models from Kenneth R. French's website[3]. Irrespective of the applied Fama-French model, we find statistically and economically significant alphas of 0.18 percent per day. Since our strategy is dollar-neutral, it is not surprising that the returns show slight exposure to the general market – therefore FF3+2 and FF5 indicate no loading. Loadings on SMB, HML, SMB5, HML5, RMW5, and CMA5 are insignificant and very close to zero. We observe a statistical significant positive loading on the reversal factor – a clear evidence that our strategy buys short-term losers and shorts short-term winners. As expected, loading on the momentum factor is small and statistically insignificant. The FF3+2 model has the highest explanatory power caused by the short-term reversal factor. Summarizing, the JDM produces statistically and economically significant returns, obtains almost no loading on systematic sources of risk, and outperforms classic pairs trading approaches.

[3]We thank Kenneth R. French for providing all relevant data for these models on his website.

	FF3	FF3+2	FF5
(Intercept)	0.0018***	0.0018***	0.0018***
	(0.0001)	(0.0001)	(0.0001)
Market	0.0186*	0.0006	0.0171
	(0.0080)	(0.0089)	(0.0093)
SMB	−0.0139	−0.0137	
	(0.0162)	(0.0162)	
HML	−0.0256	−0.0106	
	(0.0152)	(0.0163)	
Momentum		0.0087	
		(0.0113)	
Reversal		0.0736***	
		(0.0114)	
SMB5			−0.0109
			(0.0175)
HML5			−0.0183
			(0.0173)
RMW5			0.0019
			(0.0226)
CMA5			−0.0152
			(0.0277)
R^2	0.0019	0.0111	0.0019
Adj. R^2	0.0012	0.0100	0.0008
Num. obs.	4484	4484	4484
RMSE	0.0068	0.0068	0.0068

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 8: Exposure to systematic sources of risk after transaction costs for the daily returns of the top 10 pairs of the JDM from March 1998 until December 2015. Standard errors are depicted in parentheses.

5.4. Robustness checks

Whenever strategies produce high returns it appears the suspicion of data snooping. Therefore, we run a series of robustness checks on our input parameters.

First of all, we contrast the performance of the JDM with the results of 200 random bootstrap tradings. In the spirit of Gatev et al. (2006), we combine each original trading signal of the JDM with two random securities of the oil sector at that time. As expected, the average daily returns of bootstrapped pairs account for -0.01 percent per day – a reasonable value and well in line with the findings of Gatev et al. (2006). The JDM produces daily returns of 0.27 percent before transaction costs which are far superior to the results of random bootstrap trading. Hence, our strategy identifies temporal variations and exploits market inefficiencies.

In subsection 4.2, the input parameters are motivated based on the literature – we set a number of 10 top pairs ($p = 10$) and a trading threshold of two standard deviations ($k = 2$). Table 9 depicts annualized mean returns and Sharpe ratios for the JDM after transaction costs varying the input parameters p and k in two directions. Furthermore, we consider varying d-day trading periods ($d \in \{1, 2, 3, 4, 5\}$). First of all, a smaller number of top pairs leads to a better performance indicating that our pairs selection algorithm introduced in section 4 is meaningful. Higher annualized returns and Sharpe ratios can generally be found at lower levels of k – higher transaction costs in consequence of increasing trading frequency are compensated by rising returns. Regarding overnight effects in context of high-frequency data pays off – we observe that a larger trading period increases the profitability of our strategy. Overall, the initial parameter setting hits not the optimum, our model identifies correct pairs, and considering overnight effects has a positive impact on the trading results.

		Return					Sharpe ratio				
	$k \setminus d$	1	2	3	4	5	1	2	3	4	5
Top 5	1	0.8537	1.0629	1.1244	1.1489	1.1568	4.6786	6.0104	6.3887	6.6330	6.7488
	2	0.6162	0.6522	0.6695	0.6731	0.6686	4.5233	4.8030	4.8267	4.8896	4.8829
	3	0.3172	0.3273	0.3343	0.3420	0.3419	3.4314	3.3932	3.3729	3.4420	3.4389
Top 10	1	0.6501	0.8843	0.9787	1.0138	1.0204	4.3839	6.0263	6.6895	6.9871	7.0872
	2	0.4993	0.5646	0.5972	0.6084	0.6061	4.5429	5.0761	5.2433	5.3209	5.2982
	3	0.2528	0.2681	0.2860	0.2953	0.2982	3.4568	3.4713	3.5547	3.6304	3.6267
Top 20	1	0.5466	0.7421	0.8112	0.845	0.8587	4.2517	5.8778	6.4085	6.6511	6.7839
	2	0.4361	0.4780	0.4976	0.5079	0.5085	4.6075	4.9659	5.0366	5.0793	5.0707
	3	0.2257	0.2303	0.2369	0.2454	0.2493	3.6147	3.5175	3.4520	3.5098	3.5207

Table 9: Annualized mean return and Sharpe ratio after transaction for a varying number of top pairs (p), the k-times of the standard deviation, and the length of the trading period in days (d) from March 1998 until December 2015.

Figure 2 depicts the annualized returns after transaction costs (left axis), the average number of pairs traded (right axis), and the average number of pairs where closing is forced (right axis) for a varying trading period measured in days for our initial parameter setting. We observe that a larger trading period increases the annualized returns ranging from 49.93

percent for a 1-day trading period to 60.61 percent for a 5-day trading period. As expected, the average number of pairs opened during the trading period shows a similar picture – the number increases from 6.18 pairs (1-day trading period) to 9.85 pairs (5-day trading period). In contrast, the average number of pairs where closing is forced decreases for a longer trading period – a favorable property for investors (subsection 5.1). Overall, we conclude that overnight trading reduces downside risks and improves the performance of our strategy.

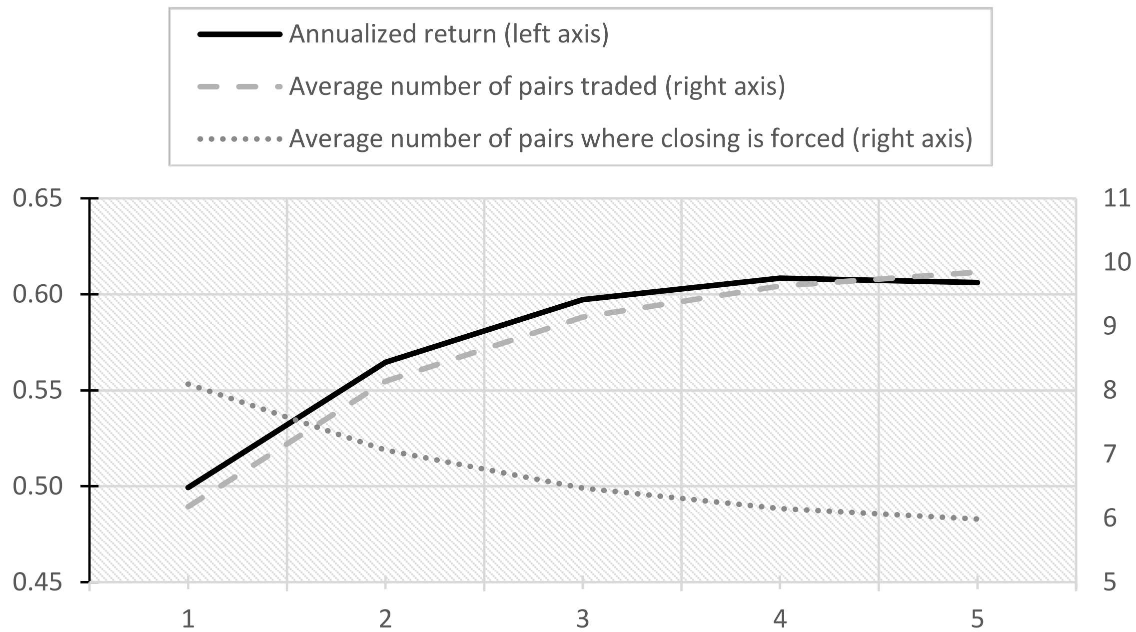


Figure 2: Annualized return after transaction costs (left axis), average number of pairs traded (right axis), and average number of pairs where closing is forced (right axis) for a varying trading period measured in days for our initial parameter setting.

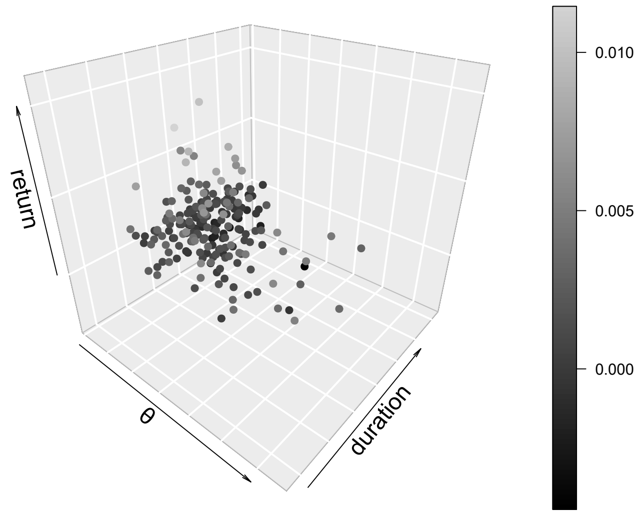
5.5. Influence of mean-reversion speed

Our trading strategy outlined in subsection 4.1 relies on the mean-reversion of the underlying spread process (Leung and Li 2015, Cartea et al. 2015). Spreads of pairs with a strong exposure to mean-reversion provide the highest process' predictability and thus, generate profits from trading. The joint information from two return series is used to create exit signals, which are mainly driven by the mean-reverting nature of their linear combination – the spread. Specifically, spreads with a high mean-reversion speed θ converge fast back to equilibrium and thus produce the best performance (Cartea et al. 2015).

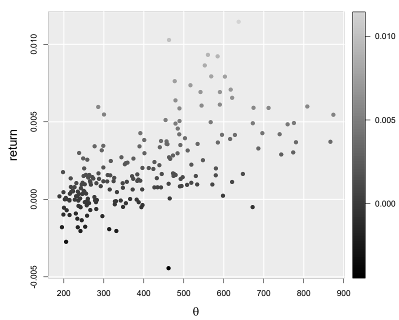
Besides a successful termination of the pairs trading strategy (Göncü and Akyıldırım 2016a), the speed of trading plays a fundamental role. Bertram (2010) emphasizes the importance of time in financial markets. The author measures the strategy's return by dividing the return per trade by the time to complete a trade cycle. Consequently, the time over which a return takes place should also be taken into consideration. This applies for our strategy – pairs that revert too slow are closed at the end of the trading period and may produce losses.

In the following, we vindicate the theoretical construct relying on mean-reversion by analyzing the relationship between achieved returns, exposure to the mean-reverting process, and the factor of time represented by the period over which a return takes place. Specifically, the attribute "return" describes the achieved daily returns after transaction costs for the top 10 pairs of the JDM, "θ" defines the average mean-reversion speed of the selected top pairs, and "duration" specifies the average holding period per round-trip trade.

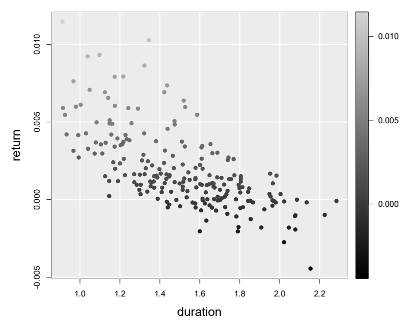
Figure 3 summarizes the attributes "return", "θ", and "duration" – the daily data base is monthly aggregated. Most interesting is the fact that the three variables depict a strong multivariate dependence, illustrated in plot (a). Furthermore, we analyze the relationship of two variables c.p., i.e., the third variable is held constant. We observe in plot (b) that "return" is powerfully linked to "θ". This fact confirms the assumption that mean-reversion is a driving component of positive returns. Thus, we may carefully infer that our pairs selection algorithm outlined in subsection 4.1 is meaningful. Considering plot (c), "return" and "duration" show a strong relationship. This is an interesting finding since the strategy's return in the sense of Bertram (2010) is optimized in two ways at the same time: For an increasing return, the time over which the return takes place becomes shorter and vice versa. Third, "duration" is strongly associated with "θ", which is illustrated in plot (d). It is not surprising that an increasing mean-reversion speed θ leads directly to a lower holding period. Concluding, the results are in line with the literature – mean-reversion is a driver for successful and fast termination of our pairs trading strategy.



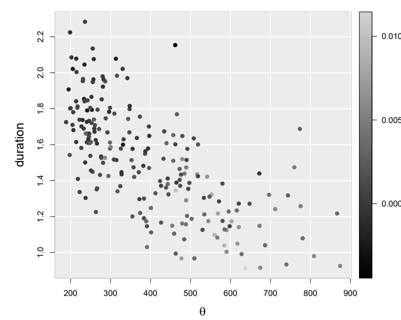
(a)



(b)



(c)



(d)

Figure 3: Monthly aggregated data of daily returns after transaction costs (return), average holding period per round-trip trade per pair (duration), and the mean-reversion rate of the top pairs (θ) in three-dimensional scatter plot (a) and two-dimensional scatter plots (b), (c), (d). The color scheme in each scatter plot represents the monthly aggregated daily returns.

6. Conclusion

In this paper, we introduce an integrated pairs trading framework based on a JDM and deploy it on minute-by-minute data of the S&P 500 oil sector from January 1998 to December 2015. In this respect, we make three contributions to the literature.

The first contribution relies on the developed pairs trading framework based on a mean-reverting JDM. To our knowledge, we are the first authors considering a jump component for pairs formation and trading in the context of high-frequency data – this fact enables us to include intraday and overnight effects. In a preliminary study, we re-confirm the necessity of considering overnight jumps. Our strategy selects pairs based on their mean-reversion speed and jump behavior, thereby receiving profitable pairs. Finally, we implement individualized trading rules based on Bollinger bands.

The second contribution focuses on the performance evaluation of our trading strategy and the implemented benchmark approaches. We find that JDM-based pairs trading outperforms traditional distance and time-series approaches. Specifically, our strategy achieves statistically and economically significant returns of 60.61 percent p.a. after transaction costs. These returns yield to an annualized Sharpe ratio of 5.30 after transaction costs – pairs trading with the distance and time-series approach ranges between 2.05 and 3.20. The returns can partially be attributed to systematic risk exposure, mostly driven by the short-term reversal factor, but daily alpha still remains at 0.18 percent after transaction costs. A series of robustness checks confirms the necessity of regarding jumps in spread modeling.

The third contribution is rooted in the influence of the mean-reversion speed on the performance of the strategy. We find that successful termination as well as fast trading speed of the pairs trading strategy are strongly influenced by the exposure to mean-reversion.

For further research, we identify three possible directions: First, the model may be extended by integrating a general Lévy-process, which is able to capture stylized facts in a more common way than the Poisson process. Second, the model performance should be evaluated by applying the pairs trading strategy on other stock universes. Third, the Bollinger bands may be refined by an exponential moving average to consider the time structure.

Bibliography

Avellaneda, M., Lee, J.-H., 2010. Statistical arbitrage in the US equities market. Quantitative Finance 10 (7), 761–782.

Bacon, C. R., 2008. Practical portfolio performance: Measurement and attribution, 2nd Edition. Wiley Finance. Wiley, Chichester, England and Hoboken, NJ.

Barlow, M. T., 2002. A diffusion model for electricity prices. Mathematical Finance 12 (4), 287–298.

Benth, F. E., Kiesel, R., Nazarova, A., 2012. A critical empirical study of three electricity spot price models. Energy Economics 34 (5), 1589–1616.

Bertram, W. K., 2009. Optimal trading strategies for Itô diffusion processes. Physica A: Statistical Mechanics and its Applications 388 (14), 2865–2873.

Bertram, W. K., 2010. Analytic solutions for optimal statistical arbitrage trading. Physica A: Statistical Mechanics and its Applications 389 (11), 2234–2243.

Bogomolov, T., 2013. Pairs trading based on statistical variability of the spread process. Quantitative Finance 13 (9), 1411–1430.

Bollinger, J., 1992. Using bollinger bands. Stocks & Commodities 10 (2), 47–51.

Bollinger, J., 2001. Bollinger on Bollinger bands. McGraw Hill Professional.

Bowen, D., Hutchinson, M. C., O'Sullivan, N., 2010. High frequency equity pairs trading: Transaction costs, speed of execution and patterns in returns. The Journal of Trading 5 (3), 31–38.

Bowen, D. A., Hutchinson, M. C., 2015. Pairs trading in the UK equity market: Risk and return. The European Journal of Finance 22 (14), 1363–1387.

Byrd, R. H., Lu, P., Nocedal, J., Zhu, C., 1995. A limited memory algorithm for bound constrained optimization. SIAM Journal on Scientific Computing 16 (5), 1190–1208.

Carr, P., Geman, H., Madan, D. B., Yor, M., 2002. The fine structure of asset returns: An empirical investigation. The Journal of Business 75 (2), 305–332.

Cartea, Á., Figueroa, M. G., 2005. Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. Applied Mathematical Finance 12 (4), 313–335.

Cartea, Á., Jaimungal, S., Penalva, J., 2015. Algorithmic and high-frequency trading. Cambridge University Press.

Clegg, M., Krauss, C., 2016. Pairs trading with partial cointegration. FAU Discussion Papers in Economics, University of Erlangen-Nürnberg.

Cont, R., 2001. Empirical properties of asset returns: Stylized facts and statistical issues. Quantitative Finance 1 (2), 223–236.

Cont, R., 2007. Volatility clustering in financial markets: Empirical facts and agent-based models. In: Teyssière, G., Kirman, A. P. (Eds.), Long Memory in Economics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 289–309.

Cont, R., Tankov, P., 2003. Financial modelling with jump processes. Vol. 2. CRC Press.

Cummins, M., Bucca, A., 2012. Quantitative spread trading on crude oil and refined products markets. Quantitative Finance 12 (12), 1857–1875.

Do, B., Faff, R., 2010. Does simple pairs trading still work? Financial Analysts Journal 66 (4), 83–95.

Do, B., Faff, R., 2012. Are pairs trading profits robust to trading costs? Journal of Financial Research 35 (2), 261–287.

Dragulescu, A. A., 2014. xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files.

Ekström, E., Lindberg, C., Tysk, J., 2011. Optimal liquidation of a pairs trade. In: Di Nunno, G., Øksendal, B. (Eds.), Advanced Mathematical Methods for Finance. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 247–255.

Elliott, R. J., van der Hoek, J., Malcolm, W. P., 2005. Pairs trading. Quantitative Finance 5 (3), 271–276.

Escribano, A., Ignacio Peña, J., Villaplana, P., 2011. Modelling electricity prices: International evidence. Oxford Bulletin of Economics and Statistics 73 (5), 622–650.

Fama, E. F., French, K. R., 1996. Multifactor explanations of asset pricing anomalies. The Journal of Finance 51 (1), 55–84.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116 (1), 1–22.

Gatev, E., Goetzmann, W. N., Rouwenhorst, K. G., 1999. Pairs trading: Performance of a relative value arbitrage rule. Working paper, Yale School of Management's International Center for Finance.

Gatev, E., Goetzmann, W. N., Rouwenhorst, K. G., 2006. Pairs trading: Performance of a relative-value arbitrage rule. Review of Financial Studies 19 (3), 797–827.

Göncü, A., Akyıldırım, E., 2016a. Statistical arbitrage with pairs trading. International Review of Finance 16 (2), 307–319.

Göncü, A., Akyıldırım, E., 2016b. A stochastic model for commodity pairs trading. Quantitative Finance 16 (12), 1843–1857.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance 48 (1), 65–91.

Jing, B.-Y., Kong, X.-B., Liu, Z., 2012. Modeling high-frequency financial data by pure jump processes. Annals of Statistics, 759–784.

Jondeau, E., Lahaye, J., Rockinger, M., 2015. Estimating the price impact of trades in a high-frequency microstructure model with jumps. Journal of Banking & Finance 61 (2), 205–224.

Kappou, K., Brooks, C., Ward, C., 2010. The S&P500 index effect reconsidered: Evidence from overnight and intraday stock price performance and volume. Journal of Banking & Finance 34 (1), 116–126.

Kim, K.-J., 2003. Financial time series forecasting using support vector machines. Neurocomputing 55 (1), 307–319.

Krauss, C., 2017. Statistical arbitrage pairs trading strategies: Review and outlook. Journal of Economic Surveys 31 (2), 513–545.

Krauss, C., Do, X. A., Huck, N., 2017. Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. European Journal of Operational Research 259 (2), 689–702.

Krauss, C., Stübinger, J., 2017. Non-linear dependence modelling with bivariate copulas: Statistical arbitrage pairs trading on the S&P 100. Applied Economics 23 (1), 1–18.

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.

Leifeld, P., 2013. texreg: Conversion of statistical model output in R to HTML tables. Journal of Statistical Software 55 (8), 1–24.

Leung, T., Li, X., 2015. Optimal mean reversion trading: Mathematical analysis and practical applications. World Scientific.

Liu, B., Chang, L.-B., Geman, H., 2017. Intraday pairs trading strategies on high frequency data: The case of oil companies. Quantitative Finance 17 (1), 87–100.

Lucia, J. J., Schwartz, E. S., 2002. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. Review of Derivatives Research 5 (1), 5–50.

Lux, T., Marchesi, M., 2000. Volatility clustering in financial markets: A microsimulation of interacting agents. International Journal of Theoretical and Applied Finance 3 (4), 675–702.

Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

Merton, R. C., 1992. Continuous-time finance, revised edition. Blackwell Publishers.

Meyer-Brandis, T., Tankov, P., 2008. Multi-factor jump-diffusion models of electricity prices. International Journal of Theoretical and Applied Finance 11 (5), 503–528.

Miao, G. J., 2014. High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. International Journal of Economics and Finance 6 (3), 96–110.

Mina, J., Xiao, J. Y., 2001. Return to RiskMetrics: The evolution of a standard. RiskMetrics Group.

Ng, H. S., Lam, K. P., Lam, S. S. (Eds.), 2003. Incremental genetic fuzzy expert trading system for derivatives market timing: 2003 IEEE International Conference on Computational Intelligence for Financial Engineering, 2003. Proceedings.

Ou, J. A., Penman, S. H., 1989. Financial statement analysis and the prediction of stock returns. Journal of Accounting and Economics 11 (4), 295–329.

Peterson, B. G., Carl, P., 2014. PerformanceAnalytics: Econometric tools for performance and risk analysis.

Pfaff, B., McNeil, A., 2016. QRM: Provides R-Language code to examine quantitative risk management concepts.

QuantQuote, 2016. QuantQuote market data and software. https://www.quantquote.com/.

R Core Team, 2017. R package stats: A language and environment for statistical computing.

Rad, H., Low, R. K. Y., Faff, R. W., 2016. The profitability of pairs trading strategies: Distance, cointegration and copula methods. Quantitative Finance, 1–18.

Rmetrics Core Team, Wuertz, D., Setz, T., Chalabi, Y., 2014. fBasics: Rmetrics - Markets and basic statistics.

Rmetrics Core Team, Wuertz, D., Setz, T., Chalabi, Y., 2015. timeSeries: Rmetrics - Financial time series objects.

Ryan, J. A., 2016. quantmod: Quantitative financial modelling framework.

Ryan, J. A., Ulrich, J. M., 2014. xts: eXtensible time series.

Seifert, J., Uhrig-Homburg, M., 2007. Modelling jumps in electricity prices: Theory and empirical evidence. Review of Derivatives Research 10 (1), 59–85.

Soetaert, K., 2009. rootSolve: Nonlinear root finding, equilibrium and steady-state analysis of ordinary differential equations.

S&P 500 Dow Jones Indices, 2015. Equity S&P 500. http://us.spindices.com/indices/equity/sp-500.

Stübinger, J., Mangold, B., Krauß, C., 2016. Statistical arbitrage with vine copulas. FAU Discussion Papers in Economics, University of Erlangen-Nürnberg.

Trapletti, A., Hornik, K., 2017. tseries: Time series analysis and computational finance.

Ulrich, J., 2016. TTR: Technical trading rules.

Venables, W. N., Ripley, B. D., 2002. Modern applied statistics with S, 4th Edition. Springer, New York.

Vidyamurthy, G., 2004. Pairs trading: Quantitative methods and analysis. John Wiley & Sons, Hoboken, N.J.

Villaplana, P., 2003. Pricing power derivatives: A two-factor jump-diffusion approach. SSRN Electronic Journal.

Wickham, H., 2016. readxl: Read Excel files.

Wickham, H., Francois, R., 2016. dplyr: A grammar of data manipulation.

Wickham, H., Hester, J., Francois, R., 2016. readr: Read tabular data.

Zeileis, A., 2006. Object-oriented computation of Sandwich estimators. Journal of Statistical Software 16 (9), 1–16.

Zeileis, A., Grothendieck, G., 2005. zoo: S3 infrastructure for regular and irregular time series. Journal of Statistical Software 14 (6), 1–27.

Zeileis, A., Hothorn, T., 2002. Diagnostic checking in regression relationships. R News 2 (3), 7–10.

Zeng, Z., Lee, C.-G., 2014. Pairs trading: Optimal thresholds and profitability. Quantitative Finance 14 (11), 1881–1893.