Graphical Abstract

Mean reversion trading on the naphtha crack

Briac Turquet, Pierre Bajgrowicz, Olivier Scaillet

Highlights

Mean reversion trading on the naphtha crack

Briac Turquet, Pierre Bajgrowicz, Olivier Scaillet

- We provide insights into OTC trading of petroleum derivatives and identify causes for short term market inefficiencies of the naphtha crack.
- We characterise these inefficiencies using non-parametric estimation of the drift and diffusion coefficients of the arithmetic Brownian motion.
- We provide a framework for filtering the grid search results in calibration and identify successful mean reversion trading strategies.

Mean reversion trading on the naphtha crack

Briac Turquet[a], Pierre Bajgrowicz[b], Olivier Scaillet[c]

[a]*ETH Zürich, D-MTEC, Weinbergstrasse 56/58, Zürich, 8006, Zürich, Switzerland*
[b]*Axpo Solutions AG, Parkstrasse 23, Baden, 5401, Baden, Switzerland*
[c]*University of Geneva, Geneva Finance Research Institute, Swiss Finance Institute, Bd du Pont d'Arve 40, Geneva, 1205, Geneva, Switzerland*

Abstract

We investigate the mean reversion of the naphtha crack after large price moves on daily data over 2014-2024. Our non-parametric estimation of the dynamics of daily changes assuming a univariate diffusion process shows that the reversion strength increases non-linearly after daily moves exceeding a certain threshold. We perform Monte Carlo simulations to study the duration for which the reversion is likely to remain active. We then backtest corresponding trading strategies. We calibrate parameters of the strategy using grid search while controlling for multiple testing. On average the tested strategies deliver positive returns after transaction costs. We are able to select a subset of outperforming strategies generating robust positive net returns. The existence of positive returns can be explained by differences in liquidity, execution speed, and categories of participants in the naphtha and Brent markets constituting the two legs of the naphtha crack.

Keywords: oil derivatives, naphtha crack, statistical arbitrage, mean reversion
JEL Classification: G13, G14, G15, G17, G18

1. Introduction

The financial investment world considers oil mainly as an alternative asset class that can be used as a hedge against inflation or geopolitical risks, and a constituent of wider commodity indexes. However, financial crude oil and oil product contracts are also important tools for physical players such as oil majors, physical traders, or refiners to hedge against flat price fluctuations or lock economics such as geographical

Email addresses: turquet18.briac@gmail.com (Briac Turquet), pierre.bajgrowicz@gmail.com (Pierre Bajgrowicz), olivier.scaillet@unige.ch (Olivier Scaillet)

arbitrage or refinery margins. The available CME and ICE[1] futures contracts are only a subset of the financial instruments used by the physical oil industry. A large number of forward contracts called 'swaps' in the industry have until not long ago existed only as pure over-the-counter (OTC) contracts and continue to transact mostly as block trades through broker markets. The academic literature (see Section 2 below) has looked at quantitative trading strategies in petroleum contracts. However, existing research has focused on exchange-traded contracts as the understanding of how OTC oil derivatives are being used remains often understood by specialised players, and data is harder to obtain. Studies of OTC markets exist, but have tackled mainly market design from a theoretical perspective and the empirical investigations have looked at other markets such as interest rates or foreign exchange.

In this article, we investigate precisely this relatively unexplored area by studying the European naphtha crack contract that trades virtually only off-exchange and is used mainly by participants involved in physical oil activities. The naphtha crack is defined as the difference between the price of the intermediary petroleum product called naphtha and the price of Brent crude oil. Our study uses daily data over the period 15/05/2014 to 14/02/2024. We scale the raw data to account for periods of changing volatility. Similarly to Prigent et al. (2001), we estimate non-parametrically the drift and diffusion coefficients of the naphtha crack daily changes assuming the corresponding process follows an arithmetic Brownian motion. The results indicate the presence of significant mean reversion, with naphtha crack prices reverting more strongly in a non-linear fashion after exceeding a certain threshold. Next we perform Monte Carlo simulations to get insights into the duration of the reversion effect. The results suggest that most of the reversion takes place the day following a large price move. In light of the above findings, we design and backtest trading strategies to check if the naphtha crack mean reversion property can be exploited to generate robust positive returns. We account for multiple testing when assessing different combinations of strategy parameters using the False Discovery Rate (FDR) developed by Benjamini and Hochberg (1995) and applied to selecting technical trading strategies on equity data in Bajgrowicz and Scaillet (2012). We use k-fold cross-validation in our backtest to assess the robustness of our results regarding different strategy types and standardisation methods of the raw price changes. Transaction costs derived from actual broker and exchange fees as well as typical bid-ask spreads and worst case scenario slippage are included in the strategy returns computations. We are able to identify a number of strategies generating significant positive returns. We also show that our trading strategies outperform benchmark dummy strategies.

[1] CME and ICE are the two main petroleum futures exchanges.

Our backtest therefore shows that the insights from non-parametric estimation can be exploited in real life. Further confirming the robustness of our findings, the simple average of the performance of the strategies in the entire universe we consider is positive after accounting for transaction costs.

The naphtha crack being the difference between the two outright contracts of respectively naphtha and Brent, our work is related to Gatev et al. (2006) who study pairs trading strategies in stocks. As argued by Gatev et al. (2006), the existence of successful short term mean reversion strategies points towards inefficiencies in the market. By trading the mean reversion, the arbitrageur reduces these inefficiencies and gets rewarded for enforcing the Law Of One Price. One candidate explanation for these temporary market inefficiencies is the difference in liquidity and execution speed between the two legs of the naphtha crack. Brent futures prices react quickly to news as they are traded by a variety of players including from the financial investment community and algorithmic trading, and trade execution happens principally on electronic platforms. On the other hand, naphtha prices need more time to adjust given contracts transact through brokers connecting a much smaller number of actors consisting mainly of physical players. As a result, liquidity is more limited in naphtha versus Brent. The low liquidity alone can also explain large price moves and the subsequent mean reversion. Finally, naphtha often trades as a spread to other products like propane or gasoline. Sharp movements in those markets can cause the naphtha crack to temporarily deviate from its fundamental value.

The reminder of the paper is organised as follows. Section 2 reviews the existing literature on the OTC markets, mean reversion arbitrage strategies and continuous-time models based on the Brownian motion (diffusion processes). Section 3 describes the naphtha crack contract and discusses differences between the Brent and naphtha markets, and explains our methodology to account for periods of changing volatility and the expiry roll. Section 4 presents results of the non-parametric estimation of the drift and dispersion coefficients, and insights from the Monte Carlo simulations. Section 5 describes results from the trading strategies backtest, and discusses potential causes for the observed temporary market inefficiencies. Section 6 provides concluding remarks and future research ideas.

2. OTC markets and commodities trading strategies literature

2.1. Over-the-Counter markets

The Dodd-Frank Wall Street Reform and Consumer Protection Act enacted in 2010 in the wake of the global financial crisis and corresponding European regulations enacted in 2012 have resulted in new regulations of OTC derivatives (EU Regulation

No 648/2012, 2012; EU Regulation No 151/2013, 2013). The resulting changes have been studied in a number of academic articles. Bellia et al. (2024) investigate whether the regulatory changes following the 2008 crisis have indeed favored clearing of OTC derivative contracts. The study is based on regulator data of sovereign credit default swap (CDS) transactions, which are cleared on a voluntary basis. Babus and Kondor (2018) propose a model for OTC trading and price information diffusion where each dealer strategy is represented as a quantity-price schedule. Glode and Opp (2019) compare the efficiency of OTC markets to that of centralised limit-order markets, assuming that trader expertise is endogenous and assuming asymmetric information among traders. Glode and Opp (2019) find that the OTC market is less competitive and more strongly benefits well-connected core traders. Lee and Wang (2018) build an OTC market model and finds that closing the OTC market can improve welfare by mitigating inefficiencies caused by search frictions and dealers' market power, particularly for the most heavily OTC-traded assets. Cartea et al. (2022) build an OTC market model and examine the potential for artificial intelligence algorithms to facilitate collusion in OTC financial markets but find that collusion is unlikely. Ehlers and Hardy (2019) look at the growing importance of OTC trading of interest rates derivatives. Ehlers and Hardy (2019) argue that regulatory initiatives promoting central clearing and electronic trading alongside advances in compression services have driven a shift from exchanges to OTC markets. Huang and Martin (2018) investigate pairs trading strategies on a variety of assets including three OTC traded currencies. The above shows that OTC trading is primarily examined from the perspective of theoretical market models and the few empirical studies do not focus on commodities markets, in particular oil contracts that we target in this paper.

2.2. Pairs trading and trading strategies in commodities markets

Our study of trading strategies on the naphtha crack which is the difference between the price of naphtha and Brent contracts can be related to the concept of pairs trading. The fundamental paper of Gatev et al. (2006) identifies stocks trading in pairs and investigates a trading rule exploiting the mean reversion of the spread between the two stocks forming a pair. Gatev et al. (2006) are able to identify profitable pairs, but note that the profitability declines over time likely due to increased hedge funds activity. The study argues that the excess returns generated by pairs trading is a compensation for arbitrageurs enforcing the Law of One Price. Advances in machine learning have enabled more sophisticated approaches to selecting pairs trading rules, with for instance Kim and Kim (2019) who use deep reinforcement learning to optimise thresholds and find that it improves performance relative to constant thresholds methods. Credit markets are also prone to pairs trading.

Prigent et al. (2001) study the spread between U.S. corporate and Treasury bonds by means of an empirical study with both non-parametric and parametric estimations of a functional arithmetic Brownian motion model. The authors also propose their own model for credit spread indices, building on the work of Stanton (1997). Another sector that can be studied under the prism of pairs trading is commodities markets where a multitude of product, quality and geographical spreads can be traded. Those spreads typically mean revert as physical traders exploit open arbitrage opportunities, and fundamental supply and demand factors adjust to balance physical markets. Holmes et al. (2013) study gasoline market integration in the US, observing different prices at state level and studying their convergence towards a long run equilibrium, demonstrating the Law of One Price in a physical setting. Ohana (2010) explores pairs trading between US heating oil and natural gas futures. Ohana (2010) decomposes daily price movements into seasonal, short-term, and long-term components, and finds that heating oil prices are a leading indicator in the short term and that both commodities influence long term forward prices symmetrically. Murat and Tokat (2009) attempts to forecast crude oil price movements using the 3-2-1 crack spread which is a very simple refinery margin indicator computed as two barrels of gasoline plus one barrel of heating oil minus three barrels of WTI crude oil. The study finds that after 2003 there is Granger-causality between the simple margin proxy and future crude oil prices. Shen et al. (2020) study mean reversion on various non-petroleum Chinese commodities futures pairs. Similarly, He et al. (2023) investigates 47 commodities traded in the Chinese commodity futures market including three oil and petroleum products contracts using intraday data, but their algorithm is not able to beat the the benchmark of holding the Wenhua Commodity Index. Vaitonis and Masteika (2017) study high frequency data on futures contracts for natural gas, Brent crude oil, WTI crude oil, New York Harbor ULSD, and New York Harbor gasoline over the month of May 2015. Vaitonis and Masteika (2017) identify pairs and apply a mean reversion trading strategy that is able to generate positive returns. However, transaction costs are not taken into account. Lubnau and Todorova (2015) use a similar approach but focus on time spreads of WTI crude oil, natural gas, heating oil and gasoline futures rather than cross-commodity pairs. Using Bollinger bands, the authors are able to identify strategies yielding positive profits. The academic literature has studied trading strategies in commodities markets extensively, but mainly looking only at exchange-traded futures contracts. It leaves OTC-traded energy contracts as an unexplored area.

2.3. Stochastic modelling and commodities prices

Stochastic processes have been widely used in finance. The Black-Scholes model for option pricing has become a cornerstone of financial theory and practice, and is based on the assumption that the price of the underlying follows a geometric Brownian motion. However, a geometric Brownian motion is not ideally suited to commodities, as these often do not follow exponential growth and can exhibit negative prices, especially with spread contracts. Stanton (1997) develops a non-parametric continuous-time diffusion model for processes that are observed at discrete intervals and applied it to Treasury Bills. Schwartz (1997) studies three Ornstein-Uhlenbeck based models for copper, oil, and gold, and finds strong evidence of mean reversion. Brooks and Prokopczuk (2013) investigate six commodities including crude oil and gasoline using variations of the arithmetic Brownian motion with jumps and compare them against equity markets. There is an ongoing debate about the best approach to modelling spreads: whether to model the two legs separately, or use a model directly for the spread itself. Prigent et al. (2001) argue that the former can lead to adding up errors if the residuals are positively correlated. Furthermore, they point out that a spread can be influenced by other factors such as relative liquidity differences between the two legs, a key behaviour that is better captured by modelling the spread directly. Mahringer and Prokopczuk (2015) investigate this fundamental question in the context of pricing heating oil and gasoline crack spread options by comparing the performance of univariate and bivariate models. They find that modelling the crack spread directly yields lower pricing errors than using a more complex bivariate model, this being true in-sample and most importantly out-of-sample. In this paper, we use the univariate approach to model the naphtha crack spread directly.

3. Petroleum derivatives markets and standardisation of naphtha crack price changes

3.1. Naphtha versus Brent markets and participants

Except for the 5 futures contract trading on the Chicago Mercantile Exchange (CME) and Intercontinental Exchange (ICE), other petroleum derivatives have historically traded as pure OTC forward contracts referred to as 'swaps' in the industry. In the wake of U.S. Dodd-Frank and European MiFID regulations, swaps started being cleared on exchanges and that is now the case for virtually 100% of the transactions. A crack or crack spread is a term used in the energy markets for the difference between the price of a petroleum product such as heating oil or gasoline, and the price of crude oil. The term crack is derived from the refinery cracking process, and trading crack spreads allows refiners to hedge their price risk. In the present article, we focus on the European naphtha crack contract cleared on the ICE. Although it can technically be traded on the exchange electronic platform WebICE, virtually all transactions are executed through brokers. Block trades, defined as privately negotiated transactions via brokers, made up more than 99% of the daily traded volume over the first six months of 2024. The naphtha crack trades by month and is a cash settled contract based on the difference between the Platts daily assessment price for Naphtha CIF NWE Cargoes and the ICE daily settlement price for Brent 1st Line Swap Future. It means that after expiry, the final settlement is computed as the average of the difference between the above two prices over the month corresponding to the maturity of the contract. Together with Argus, Platts is one of the main Price Reporting Agencies (PRAs) whose role is to publish official assessments of physical transactions used either in physical deals or to price financial derivatives. At the time of writing, the daily Platts assessment for Naphtha CIF NWE Cargoes reflects the mean value of naphtha for physical delivery 10-25 days forward from the date of publication. Contrary to the naphtha contracts, the ICE Brent Crude futures contract is widely traded electronically. Holding Brent futures can result in physical delivery with an option to cash settle.

We have just seen that the two legs of the naphtha crack typically transact on fundamentally different types of markets, i.e., broker market versus electronic trading on exchanges. The differences extend to the players active in the respective markets. Naphtha is a liquid unfinished petroleum product between the lighter gases (i.e., propane and butane) and the heavier jet/kerosene. It is primarily produced in refineries from the distillation of crude oil, but can also be obtained in refinery secondary units, or directly from field production after being processed at a natural gas liquids plant. Naphtha is used for producing gasoline, either via direct blending or after being processed in a reformer. The other main uses are as a feedstock in the petrochemical industry for producing olefins in steam crackers or aromatics after going through a reforming process, or as diluent for heavy crude (e.g. from Canada or Venezuela) to reduce the viscosity and facilitate transport. The wide range of naphtha demand outlets results in a diversity of players with different utility functions being active in naphtha financial derivatives. We have mentioned refiners hedging their economics, such as geographical arbitrage or refinery margins, by trading cracks. Naphtha outright swaps are also the primary instruments to hedge physical cargoes pricing of the Platts naphtha quote against flat price movements. The following simplified example explains the concept. A company buys a cargo of naphtha and agrees to pay the average of the Platts naphtha quotation minus a discount over the 5 days around the load date. To protect itself from flat price moves that are most of the times unrelated to naphtha fundamentals, the company sells naphtha swaps for one fifth of the volume on each of the 5 pricing days. The company then finds a buyer in another region and agrees to sell the cargo at the average of the Platts naphtha quotation plus a premium over the 5 days around the discharge date. The company buys back its hedges on each day of the pricing window of the sale. Thanks to this mechanism, the company has insulated itself from flat price moves between the time of the purchase and the time of the sale. For example, if flat price drops between loading and discharge, the company will sell its cargo for less than it bought it, but it will make an offsetting gain on its hedges. Due to its use for making gasoline, or the competition between naphtha and propane or butane as a feedstock for the petrochemical industry, naphtha is also traded as a spread to gasoline or propane swaps. A further instance where European naphtha swaps trade in connection to another contract is geographical arbitrage. Given Asia needs to structurally import naphtha from e.g. Europe and the Arab Gulf, market players use a Europe versus Japan naphtha contract to lock economics. Also, refiners or traders owning storage capacity buy and sell naphtha time spreads again to lock their economics. The consequence of this variety of players interacting with naphtha means that naphtha prices can be more affected by drivers specific to naphtha, gasoline, or petrochemical markets, than by what is driving crude prices. The majority of naphtha derivatives transactions are carried out by players involved in a physical activity, e.g. oil majors, trading houses, refiners, terminal operators. Some hedge funds with no intention to touch physical oil are setup to trade naphtha contracts, but it represents a minority. High costs to get registered with authorities and the clearing entity, and go through the approval process with brokers represent a significant barrier to entry. Retail investors are therefore absent from the market. The next paragraph provides numbers on the resulting liquidity difference between naphtha forward contracts and Brent futures.

Brent futures on the other hand are traded by a much larger variety of players. It obviously constitutes one of the main hedging instrument for the physical petroleum industry. Brent is used not only in relation to North Sea crude grades but is one of the main benchmarks for pricing crude globally. While mostly only naphtha players get involved in trading naphtha swaps, Brent is used by players from the entire oil market spectrum from gasoline, to diesel and fuel oil. In addition to the physical world, the financialisation of crude futures markets has increased the influence of financial investors such as commodity index traders or hedge funds on the oil futures price. ICE Brent futures are a major component of benchmarks such as the S&P GSCI or the Bloomberg Commodity Index. The indexes are designed to be investable by including the most liquid commodity futures, and are used by mutual funds or exchange-trade funds (ETFs). Investors use them for portfolio diversification purposes, as a hedge against inflation or geopolitical risks. Investment money can also flow into Brent funds to collect the roll when crude oil is in backwardation. Retail traders (i.e., non-professional traders) can directly trade Brent futures on their accounts. Contrary to naphtha, there is no need to involve brokers. Most trading happens electronically (97% of the daily volume on average during the first half of 2024), including through algorithmic trading execution. Finally, retail investors can also trade Brent futures on electronic platforms. As a result, Brent futures are much more liquid market than naphtha swaps. For example during the first half of 2024, average daily traded volume of prompt-month naphtha swap was 1.4 Mio bbls compared to 111 Mio bbls for the prompt Brent futures contract.

3.2. Contract roll and volatility standardisation

The data in the current paper is obtained from ICE for the European naphtha crack referred to as Naphtha CIF NWE Cargoes (Platts) vs Brent 1st Line Future. We use historical daily end-of-day settlement prices for the 15/05/2014 to 14/02/2024 period. Naphtha swaps are settled at 19:30 London time to match Brent futures settlements. The historical data available from ICE starts in 2014, which corresponds to when naphtha swaps started being cleared on the exchange. Pre-2014 naphtha crack prices could probably be obtained from brokers. However, the older data would need to be cleaned as there was no regulatory obligation to report it, and few people had realised the value of data at the time. In the 1-2 days preceding expiry, naphtha prices can be subject to strong volatility caused by squeezes in the physical naphtha cargoes market, and become disconnected from crude oil prices. Lower liquidity as a number of players square off their positions ahead of expiry can also result in sharp naphtha price moves. In order to not be affected by this expiry noise, and to avoid having to deal with rolling the position from the prompt contract to the next around expiry when we backtest trading strategies later in the paper, our naphtha crack time series transitions to the next-month contract 5 days ahead of expiry. The naphtha crack contract expires on the last day of the month prior to the contract maturity month. Therefore, our time series corresponds to prompt-month naphtha crack prices from the beginning of the month to 5 exchange business days before expiry, and to the second-month contract over the last 5 days of the month.

The outright naphtha crack price is not stationary. Its level is driven by market drivers including other product crack spreads or the relative price of crude oil versus natural gas liquids (NGLs) such as ethane or propane. The present paper studies daily changes in the naphtha crack, which display much better stationary properties. As shown in Figure 1, naphtha crack daily price changes are subject to periods of changing volatility. Volatility increased in 2020 due to the COVID-19 pandemic lockdowns that caused demand for transportation fuels to drop sharply while demand for petrochemical products remained healthy on a global basis. Diesel and gasoline cracks came off while the naphtha was supported. The opposite phenomena took place during the reopenings from lockdowns phases when refiners had to resume producing more transportation fuels again. Volatility increased again from 2022 on the back of the conflict in Ukraine, and as global flows of naphtha had to be redirected following the European Union ban on imports of petroleum products from Russia, and the G7 group price cap sanctions. In order to account for the changing volatility, we standardise the raw price changes using different volatility measures. Gatev et al. (2006); He et al. (2023); Lubnau and Todorova (2015); Shen et al. (2020); Vaitonis and Masteika (2017) use a similar approach to account for periods of different volatility levels. It allows to make sure that the trading strategies we investigate in Section 5 do not generate a signal only during the periods of higher volatility.

We use the following notation for the raw daily naphtha crack change:

$$\Delta_t^{Raw} = C_t - C_{t-1}, \tag{1}$$

where C is the price of the naphtha crack in \$/bbl. Throughout this paper, we investigate three different volatility standardisation methods. The first is a reactive 20-day rolling historical standard deviation. The second is a less reactive 60-day rolling historical standard deviation. The third is a predictive volatility measure, obtained from a daily one-day ahead GARCH(1,1) forecast (Bollerslev, 1986). We need a volatility measure that is reactive but also not jittery so that the dynamics of the raw signal around large daily changes are preserved. To help smooth each normalising volatility time series, we apply an Exponentially Weighted Moving Average (EWMA) with $\alpha = 0.2$. Finally, we re-scale the standardised price changes in order to put them on the same scale and be able to compare reversion threshold levels across the three volatility estimation methods. It is done by dividing the standardised daily price changes series by its historical standard deviation (single standard deviation value), and multiplying by the raw series historical standard deviation (single standard deviation value). The process of standardisation is described as follows:

$$\Delta_t = \frac{\Delta_t^{Raw}}{\mathrm{Vol}(R_t)} \times \frac{\mathrm{std}(R_t)}{\mathrm{std}(S_{t-1})}, \tag{2}$$

where Δ_t is the standardised Delta at time t, Δ_t^{Raw} is the raw Delta at time t. std(.) is the population standard deviation, and Vol(.) is the chosen volatility measure for the standardisation (20-day rolling historical standard deviation, 60-day rolling historical standard deviation, or one-day ahead GARCH(1,1) forecast). The subset

R_t is defined as $R_t = \{\Delta_i^{\text{Raw}} \mid i = 0, \ldots, t\}$. S_{t-1} is defined as $S_{t-1} = \{\Delta_i \mid i = 0, \ldots, t-1\}$ so as to avoid any look-ahead bias.

We test the standardised price change (Delta) stationarity using the augmented Dickey-Fuller test with no trend or mean component, therefore testing the unit root hypothesis on $\Delta_t = \Delta_{t-1} + u_t$. We cannot reject the null hypothesis for all three standardisation methods. The stationarity of the series justifies the use of the k-fold cross-validation procedure in Section 5.1. Table 1 presents summary statistics for the different raw and standardised time series. The mean of the four time series is close to 0. The standard deviations are all similar which was expected after the re-scaling process. The skewness is approximately twice as large for the slow standardisation method, and negative contrary to the other two normalisation methods. It can be explained by the structural downward trend in the outright naphtha crack in 2022, and the fact that the slower volatility measures do not filter out the initial bursts of strong negative daily price changes. We can observe the relatively lower kurtosis of the GARCH(1,1) standardisation. It confirms that the predictive abilities of the GARCH(1,1) model enables a more uniform standardisation across the range of daily changes values, as the 20- and 60-day historical rolling standard deviations introduce a time lag which causes the initial large price changes of volatility bursts to undergo lesser standardisation. In the remaining of the paper, we sometimes omit the "standardised" qualifier, and refer to the standardised time series simply as daily naphtha crack changes (Deltas).

	Raw daily changes	Daily changes standardised using:		
		20-day std. dev.	60-day std. dev.	GARCH(1,1) vol. forecast
Mean [$/bbl]	-0.001	0.002	0.000	0.000
Std. dev [$/bbl]	0.58	0.58	0.58	0.58
Min [$/bbl]	-3.90	-2.60	-3.50	-2.58
Max [$/bbl]	5.72	5.00	4.23	3.59
Skewness [-]	0.13	0.08	-0.23	-0.11
Kurtosis [-]	10.14	3.91	3.11	1.24
aDF stat [-]	-10.70	-30.14	-22.86	-34.16

Table 1: Summary statistics of the time series. Raw statistics refer to the original naphtha crack daily Delta time series in $/bbl units. 20-day standard deviation, GARCH(1,1), and 60-day standard deviation refer to the standardised daily Delta time series. aDF stands for the augmented Dickey-Fuller test.

3.3. Fee structure and trading costs

The bid-ask spread on the naphtha crack typically ranges between $0.05/bbl and $0.10/bbl during most of the day. The market is most active in the period of

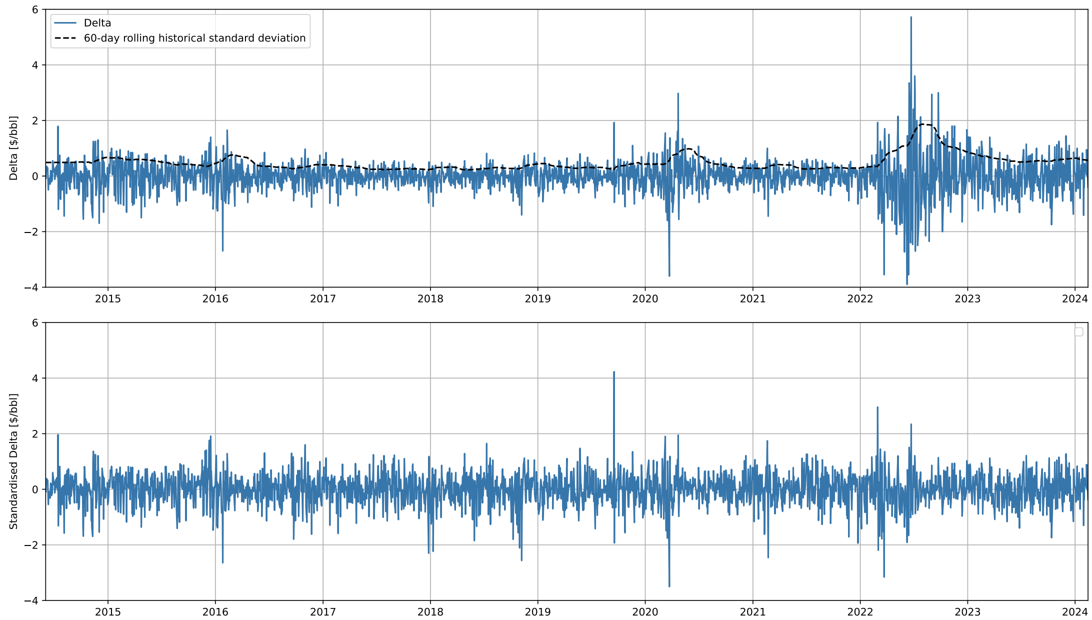


Figure 1: Naphtha crack daily price changes (Deltas) time series (top) and standardised price changes (Deltas) time series (bottom) using 60-day rolling historical standard deviation.

time before the end of the Platts Market-On-Close (MOC) price assessment process (16:30 London time). Liquidity increases during that time as discussed in Frino et al. (2018). In addition to the bid-ask spread, broker and exchange clearing fees must be accounted for in the transaction costs. Finally, there are negligible fixed costs like exchange subscription fees. Broker fees may vary by broker and company but a typical value for trading naphtha cracks is \$0.01/bbl one-way. Clearing fees for the Naphtha CIF NWE Cargoes (Platts) vs Brent 1st Line Future contract cleared on ICE amount to \$0.001/bbl one-way. Based on the above, we assume total one-way transaction costs of \$0.10/bbl when evaluating trading strategies in Section 5. It is a realistic number for covering the spread between value and bid or ask, as well as broker and clearing fees that occur in dynamic trading of the naphta crack. In case a stop-loss is triggered, we add an additional \$0.20/bbl for extra slippage.

4. Non-parametric estimation

Before designing a trading strategy, we apply a non-parametric approach to extract information about the dynamics of the naphtha crack, and in particular the presence of a potential mean reversion. We assume the standardised naphtha crack

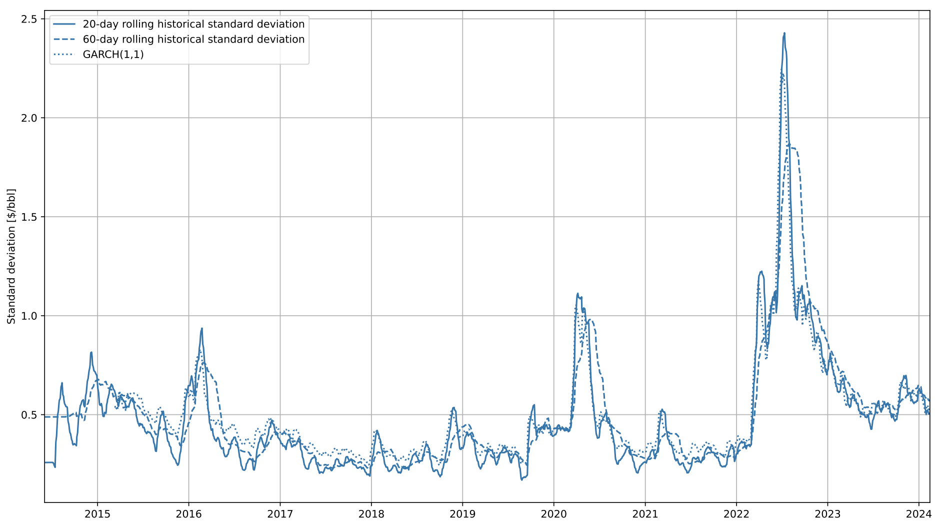


Figure 2: Time series used for standardisation. The dotted line is the standard deviation of the GARCH(1,1) one-day forecast, the dashed line represents the 60-day rolling historical standard deviation, the full line stands for the 20-day rolling historical standard deviation.

changes follow the stochastic differential equation (arithmetic Brownian motion)

$$d\Delta_t = \mu(\Delta_t)dt + \sigma(\Delta_t)dW_t, \tag{3}$$

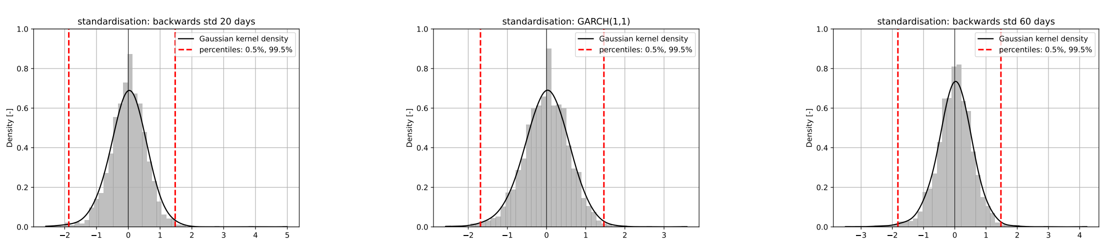
where W_t is a standard Brownian motion, and Δ_t is the standardised naphtha crack price change. We use the estimations for μ and σ proposed by Stanton (1997), and applied by Prigent et al. (2001) to investigate the dynamics of credit spread indices. Next we perform Monte Carlo simulations to study the duration of mean reversion effects in the naphtha crack.

4.1. Density estimation

Similarly to Prigent et al. (2001), we estimate the densities of the standardised daily naphtha crack price changes by means of a Gaussian kernel given by

$$\widehat{f}_i(x) = \frac{1}{nh}\sum_{t=1}^{n}\varphi\left(\frac{x-\Delta_t}{h}\right), \tag{4}$$

where $\varphi(.)$ is the standard normal density function, n is the number of observations, and h is a smoothing parameter called the bandwidth. We estimate h via 10-fold cross-validation hyperparameter tuning. The resulting density estimations are plotted in Figure 3 for the three standardisation methods.



(a) 20-day rolling historical standard deviation

(b) GARCH(1,1) 1 day ahead standard deviation forecast

(c) 60-day rolling historical standard deviation

Figure 3: Gaussian kernel density estimation of the standardised price changes (Deltas). For each standardisation method experimented, we show a histogram of the standardised prices change (Delta) values in grey, overlaid by the kernel estimated density curve in black. 0.5 and 99.5 percentiles are depicted by vertical dashed red lines.

4.2. Drift estimation

We estimate the drift term using the Taylor series approximation presented in Stanton (1997) and use the 3^{rd} order expansion, since a higher order provides a better approximation and no discontinuities are observed over the domain for this order. The 1^{st} and 2^{nd} order Taylor series for the drift term are provided in Stanton (1997) and are available upon request. The 3^{rd} order drift term can be estimated using:

$$\hat{\mu}(x) = \frac{1}{6}\left(18\frac{\sum_{t=1}^{n-1}(\Delta_{t+1}-\Delta_t)\varphi\left(\frac{x-\Delta_t}{h}\right)}{\sum_{t=1}^{n-1}\varphi\left(\frac{x-\Delta_t}{h}\right)} - 9\frac{\sum_{t=1}^{n-2}(\Delta_{t+2}-\Delta_t)\varphi\left(\frac{x-\Delta_t}{h}\right)}{\sum_{t=1}^{n-2}\varphi\left(\frac{x-\Delta_t}{h}\right)} + 2\frac{\sum_{t=1}^{n-3}(\Delta_{t+3}-\Delta_t)\varphi\left(\frac{x-\Delta_t}{h}\right)}{\sum_{t=1}^{n-3}\varphi\left(\frac{x-\Delta_t}{h}\right)}\right). \quad (5)$$

Results shown in Figure 4 provide evidence that mean reversion is present. The charts also show that the mean reversion effect increases in a non-linear fashion with larger naphtha crack price changes, displaying a visible inflexion point after the naphtha crack change exceeds a certain threshold. The inflection happens around a \$1/bbl move in the naphtha crack, but the threshold depends on the standardisation method and the sign of the price change.

4.3. Diffusion estimation

For the estimation of the diffusion term, we investigate several approximations based on the derivations of Stanton (1997), namely Expectation-based and Variance-based estimations of 1^{st}, 2^{nd} and 3^{rd} order. Higher order Taylor expansions are problematic for higher values given the low number of data points, and are also subject to discontinuities. The highest orders without discontinuities are the 2^{nd} and 1^{st} for the Expectation-based and Variance-based estimations. Using an iterative formulation of Equation 4, we generate 10 years of data using each of the two candidate approximations for the diffusion term. We find that the 2^{nd} order Expectation-based approximation leads to unrealistic extrema (in the +/- \$15/bbl range), whereas the Variance-based approximation produces extrema that are in line with the historical observations (in the +/- \$4/bbl range). Based on those simulations and on comments in Stanton (1997), we select the 1^{st} order Variance-based estimation for the remainder of the present paper given by the following equation:

$$\hat{\sigma}(x) = \left(\frac{\sum_{t=1}^{n-1}(\Delta_{t+1}-\Delta_t-\hat{\mu}(x))^2\varphi\left(\frac{x-\Delta_t}{h}\right)}{\sum_{t=1}^{n-1}\varphi\left(\frac{x-\Delta_t}{h}\right)}\right)^{\frac{1}{2}}. \quad (6)$$

Figure 4 shows that the diffusion term increases with larger daily naphtha crack moves. Combined with the shape of the drift function, it hints to a fast return to the mean after large daily changes.

4.4. Monte Carlo simulations

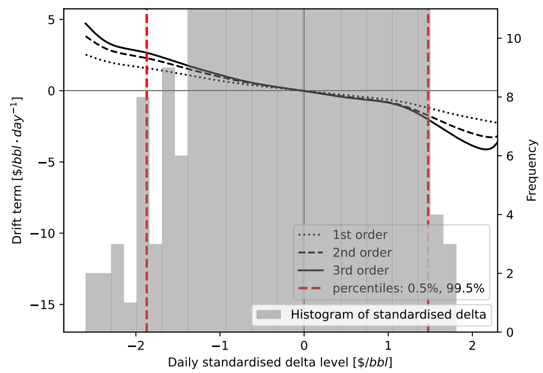
Results of Section 4.1 have confirmed that the naphtha crack is subject to mean reversion, especially after large moves. In order to study the duration of the reversion effect, we perform Monte Carlo simulations. We sample values above a given threshold from the Gaussian kernel estimation of the density from Section 4.2. For simplicity, we use a unique threshold set to \$1.6/bbl for both positive and negative naphtha crack changes. A visualisation of this process is provided in Figure 5. We then generate 5 days of standardised naphtha crack changes using the iterative form below of Equation 4 as described in Glasserman (2003):

$$\Delta_{t+1} = \Delta_t + \mu(\Delta_t)dt + \sigma(\Delta_t)\sqrt{dt}\ N[0,1], \tag{7}$$

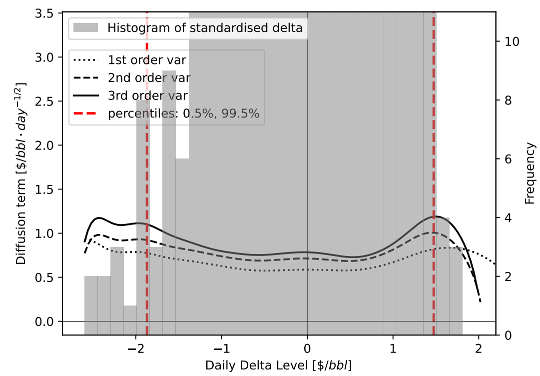
where dt is equal to 1 day, $N[0,1]$ is a standard normally distributed random variable and Δ_0 is a random sample from the Kernel density estimation that exceeds the predefined threshold set at \$1.6/bbl. For μ and σ, we use the estimates presented above. Given that Equation (7) is a model for the standardised daily naphtha crack changes, we need to map the simulated trajectories back to "unstandardised" daily changes. In order to achieve this reverse transformation, we use an OLS regression and model the residuals using the distribution presented by Johnson (1949). This process is further detailed in Appendix B. We perform 10,000 Monte Carlo iterations. The results of these simulations are shown in Figure 6. Most of the reversion takes place on the first day for all three standardisation methods. Holding the position beyond the first day decreases profits for two out the three standardisation methods.

5. Trading algorithm

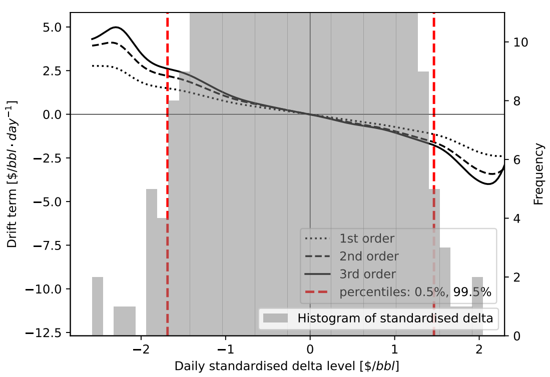
Building on the insights provided by the non-parametric estimation and the Monte Carlo simulations, we devise trading strategies that exploit the mean reversion of the naphtha crack after a large daily move. We do so by backtesting a series of trading rules, varying in terms of standardisation method of the raw daily naphtha crack price changes, entry rule, performance measure, trigger level, and holding period. In the selection process, we account for multiple testing and test the robustness of the performance using a cross-validation setup. We assess the performance of the selected strategies against two benchmark trading strategies. Finally, we discuss potential causes of the temporary market inefficiencies allowing our trading strategies to generate positive profits.



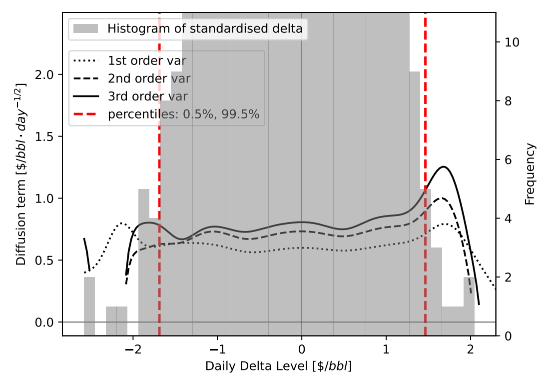
(a) Drift: 20-day historical standard deviation



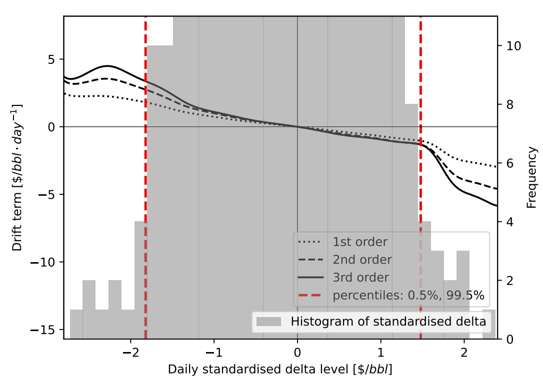
(b) Diffusion: 20-day historical standard deviation



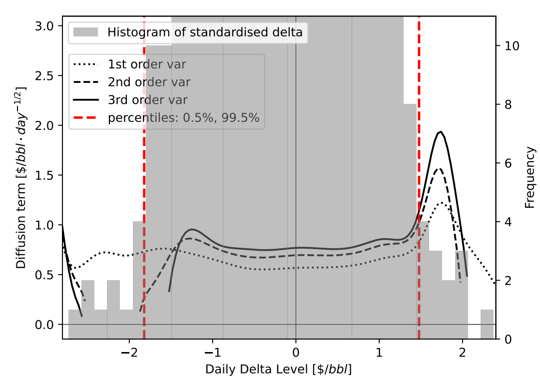
(c) Drift: GARCH(1,1) day ahead volatility forecast



(d) Diffusion: GARCH(1,1) day ahead volatility forecast



(e) Drift: 60-day historical standard deviation



(f) Diffusion: 60-day historical standard deviation

Figure 4: Non-parametric drift and diffusion estimations for each of the three tested standardisation methods. The dotted, dashed, and solid lines correspond to respectively first, second, and third order Taylor expansions. The histogram of the corresponding standardised price changes (Deltas) series is shown in grey, the 0.5 and 99.5 percentiles are depicted by vertical dashed red lines.

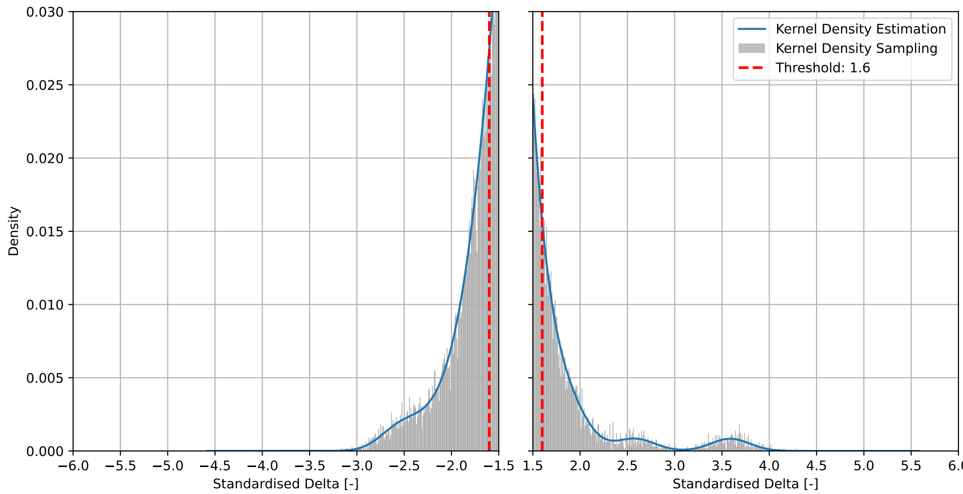
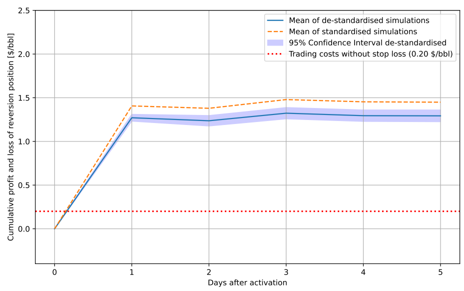
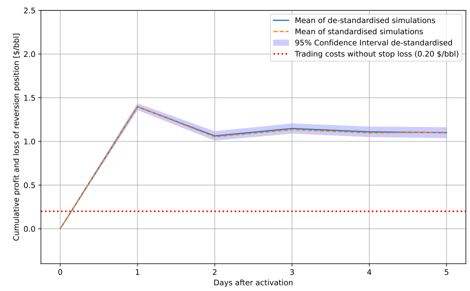


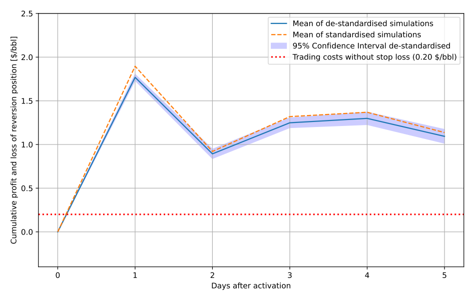
Figure 5: Sampling from the Kernel Density Estimation. The sampled values are shown in the form of a grey histogram. The Gaussian kernel estimated density function from which the samples are drawn is shown as the blue solid line. The defined threshold magnitude for negative and positive standardised Delta is shown at -1.6 \$/bbl and 1.6 \$/bbl as vertical red dashed lines. The standardisation series used in this example is the 60-day historical rolling standard deviation.



(a) Monte Carlo simulations for the 20-day historical rolling standard deviation standardisation.



(b) Monte Carlo simulations for the GARCH(1,1) standard deviation standardisation.



(c) Monte Carlo simulations for the 60-day historical rolling standard deviation standardisation.

Figure 6: Monte Carlo simulations results for each of the three tested standardisation methods. The average of standardised Delta obtained through the 10,000 Monte Carlo simulations is shown as the orange dashed line. The blue solid line represents the average of the de-standardised Delta. The blue area represents the interval containing 95% of the de-standardised Delta values. The red horizontal dotted line represents the costs related to trading, estimated at \$0.20/bbl.

5.1. Backtesting framework, entry strategy, standardisation method, and performance measurement

To assess the robustness of the trading strategies performance, we use a k-fold cross-validation procedure. Similarly to He et al. (2023), we choose a 80% calibration and 20% validation split, thus leading to a 5-fold cross-validation. The use of k-fold as opposed to a rolling cross-validation is motivated by the stationary nature of the standardised daily Delta time series, as shown in Section 3.2 with the augmented Dickey-Fuller test and by the short-term nature of the trading strategies. Because we test holding periods up to 5 days after the initial trigger, we ensure that no position is initiated during the 5 days preceding the end of the calibration or validation period.

We investigate two trading strategies each having a different entry rule. The first strategy consists in entering a reverting position on the day that the standardised daily change exceeds a pre-determined threshold, and holding the position for a pre-determined number of days. The second strategy consists in waiting one day before entering the reverting position, and also holds the position for a pre-determined number of days. We call the two strategies respectively "immediate entry" and "delayed entry". In both cases, we apply a stop-loss in case an open position exceeds a \$1.5/bbl loss, and a take profit rule when an open position exceeds a \$3/bbl gain. The insights from the non-parametric estimation of Section 4 indicate immediate mean reversion on average over the full sample. We still test the "delayed entry" strategy to check if there are periods where the reversal does not occur immediately, i.e., when a large move in the naphtha crack carries some momentum the following day before correcting later. As in the non-parametric estimation, we explore three different standardisation methods. The first is a reactive 20-day rolling historical standard deviation. The second is a less reactive 60-day rolling historical standard deviation. The third is a predictive volatility measure, obtained from a daily one-day ahead GARCH(1,1) forecast.

We test two different performance measures as the criterion to select the best combination of entry strategy, threshold, holding period, and standardisation method. The first measure is the total profit over the sample period Π^{total} computed as the sum of the raw daily naphtha crack changes when the strategy is generating a long or short investment signal. The second measure is the average profit per trade $\Pi^{per\,trade}$, which is computed as the total profit over the entire sample divided by the number of trades. Both performance measures are defined in Equations 8 and 9 below. Since we include stop-loss and take-profit rules, we do not use volatility-dependent performance measures such as the Sharpe ratio since the downside is strictly capped. Each trading rule k, defined by a combination of performance measure, standardisation method, entry strategy, threshold, and holding period, generates an investment signal $s_{k,t-1}$ for each prediction period t, $L \leq t \leq T$. $s_{k,t-1}$ equals 1 for a long position, 0 for no position, and -1 for a short position. T is the number of days in the sample, and L is the number of days needed to initialise the GARCH(1,1) model. We use $L = 200$.

$$\Pi_k^{total} = \sum_{t=L}^{T} s_{k,t-1} \Delta_t^{Raw}, \tag{8}$$

$$\Pi_k^{per\,trade} = \frac{1}{n} \sum_{t=L}^{T} s_{k,t-1} \Delta_t^{Raw}, \tag{9}$$

where n is the number of trades triggered over the sample period. In addition, we add the transaction costs discussed in Section 3.3. As a reminder, these amount to \$0.10/bbl one-way, and an additional \$0.20/bbl for slippage in case the stop loss is triggered. The trading signal is generated using standardised daily naphtha crack changes. However, the performance needs to be measured on raw price changes. For ease of comparison between the calibration and the testing samples, we report the annualised total profit when discussing the results.

5.2. Controlling for multiple testing during threshold and holding period selection

The optimal threshold and holding period are defined during the calibration phase by testing different combination for both parameters. We test thresholds ranging from 0.8 to 2.0 by 0.1 increments, amounting to 13 values, and holding periods between 1 and respectively 4 or 5 days depending on whether the entry strategy is delayed entry or immediate entry. It is because we do not want to deal with rolling an open position from one contract month to the next. With the immediate entry strategy, it results in a universe of 65 combinations of parameters. With the delayed entry strategy, there are 52 parameter combinations to test. For a given choice of standardisation method, entry strategy and performance measure, the following procedure is applied to select the optimal threshold and holding period over the calibration sample. The trading signal is generated using standardised daily naphtha crack changes, and the performance is computed on the raw price changes. Before selecting the best performing combination of parameters, we test if the average of the obtained returns are significantly higher than 0 using a one sided z-test at size level $\alpha = 0.05$. Due to the large number of threshold and holding period parameter combinations, it is necessary to control for multiple testing. To do so, we use the False Discovery Rate (FDR) introduced by Benjamini and Hochberg (1995), and applied to technical trading rules by Bajgrowicz and Scaillet (2012). As in Bajgrowicz and Scaillet (2012), we assume block dependence of the tests to satisfy the weak dependence condition. Indeed, for close threshold and holding period combinations the trading rules can possibly generate similar trading signals. However, for more distant thresholds or holding periods the trading signals become independent. We set the FDR at 10%. Among the strategies that pass the one sided z-test and the multiple testing procedure, we select the threshold and holding period combination corresponding to the strategy with the best performance (with respect to the chosen performance measure). We then assess the robustness of the selected trading rules performance in the validation phase.

5.3. Backtest results

We run cross-validations for 24 combinations of performance measure, scaling type, and entry strategy. The results are shown in their entirety in Appendix A. Out of these results, we select the strategies that have a higher than 50% win rate, a positive net profit, and a share of successful cross-validations exceeding or equal to 80%. A cross-validation is considered successful if a statistically significant combination of threshold and holding period is found and trades are triggered in validation. Using this process, we are able to select 7 strategies generating robust positive performance. Table 2 shows the performance numbers together with a number of statistics such as the percentage of winning signals or the typical threshold levels. All profits reported include transaction costs and are thus referred to as net profits. Across these 7 best performing strategies, the average net profit per trade over the validation samples ranges between \$0.25/bbl and \$0.78/bbl. The average total yearly net profit lies between of \$0.54/bbl and \$1.03/bbl, and the average win rate between 52.1% and 65.6%.

The threshold and holding period parameters selected during calibration depend on the chosen performance measure. When using the total profit as the performance measure, the median of selected thresholds lies between \$1.45/bbl and \$1.7/bbl, and the median of the holding periods is 4 days. Using the average profit per trade criteria results in selecting higher thresholds, ranging from \$1.95/bbl and \$2.0/bbl for the median values. The median holding period is also slightly longer, reaching 5 days in one occurrence. We also observe a relationship between the performance measure and the standardisation method. The best strategies according to the total profit measure use a reactive standardisation method (20-day historical rolling standard deviation and GARCH(1,1)), whereas the profit per trade measure results in using a less reactive standardisation (60-day historical rolling standard deviation) in most cases. The average profit per trade measure is more selective, as one single trade can deliver good performance according to this metric. On the other hand, the total profit measure incentivises having more frequent trading signals. The slower 60-day historical rolling standard deviation implies that the sudden extreme daily price changes are less smoothed, as shown in Figure 2. It is reasonable that a more selective strategy favours a standardisation method that does not smooth outliers from a sudden volatility burst too much, allowing for a stronger discrimination but less frequent opportunities. The average profit per trade measure results in an average number of trades per year ranging between 1.1 and 1.8. With the total profit criteria, the average number of trades per year increases to between 2 and 3.6. Interestingly, the average profit per trade does not translate into a higher win rate in validation despite being more selective. The maximisation of the profit per trade implies a greater gain per trade on the winning trades, but not a higher win rate, and a lower trading frequency when compared to maximising the total profit.

Strategy	performance measure	Total Profit	Total Profit	Total Profit	Total Profit	Profit Per Trade	Profit Per Trade	Profit Per Trade	benchmark strategy reversion	benchmark strategy buy&hold
	scaling type	20-day std	20-day std	garch 1,1	garch 1,1	60-day std	60-day std	garch 1,1		
	EWMA alpha	0.15	0.15	0.15	0.2	0.15	0.2	0.2		
	core strategy	1	2	1	1	1	1	1		
Calibration	average yearly number of trades	4.33	2.74	4.57	4.53	2.14	1.84	0.97	-	-
	median yearly number of trades	4.11	2.62	5.33	5.40	1.74	1.74	1.04	-	-
	average p_value in multiple testing	0.00078	0.00542	0.00032	0.00053	0.00090	0.00084	0.00073	-	-
	median p_value in multiple testing	0.00032	0.00408	0.00030	0.00049	0.00112	0.00085	0.00059	-	-
	average net profit per trade [$/bbl]	0.53	0.49	0.61	0.64	0.92	1.01	1.46	-	-
	median net profit per trade [$/bbl]	0.59	0.55	0.53	0.48	0.95	0.95	1.41	-	-
	average yearly total net profit [$/bbl]	2.09	1.37	2.41	2.27	1.74	1.48	1.42	-	-
	median yearly total net profit [$/bbl]	2.31	1.43	2.65	2.58	1.54	1.54	1.60	-	-
	average winrate	60%	67%	60%	60%	60%	63%	89%	-	-
	median winrate	58%	65%	59%	59%	63%	64%	94%	-	-
Calibration results	average threshold [$/bbl]	1.45	1.57	1.45	1.45	1.90	1.94	1.88	-	-
	median threshold [$/bbl]	1.60	1.70	1.45	1.45	2.00	2.00	1.95	-	-
	average holding period [days]	4.0	4.3	4.3	4.3	4.2	4.4	5.0	-	-
	median holding period [days]	4	4	4	4	4	4	5	-	-
Validation with stop loss and take profit	average yearly number of trades	3.5	2.0	3.6	3.5	1.8	1.7	1.1	3	3
	median yearly number of trades	3.0	2.0	4.2	4.2	1.5	1.0	1.0	3	3
	average net profit per trade [$/bbl]	0.25	0.36	0.73	0.74	0.78	0.78	0.68	-0.22	-0.18
	median net profit per trade [$/bbl]	0.27	0.58	0.15	0.18	0.66	0.66	0.44	-0.22	-0.18
	average yearly total net profit [$/bbl]	1.03	0.68	0.54	0.60	0.75	0.81	0.67	-2.67	-2.12
	median yearly total net profit [$/bbl]	1.19	1.31	0.76	0.87	1.20	1.20	0.70	-2.65	-2.11
	average winrate	56.7%	65.6%	61.9%	63.3%	62.5%	62.5%	52.1%	36.1%	39.6%
	median winrate	63.0%	67.0%	55.0%	58.0%	75.0%	75.0%	54.0%	34.2%	42.0%
Share of successful CV (out of 5)		100%	100%	80%	80%	100%	100%	80%	100%	100%

Table 2: Results from the 5-fold cross-validation. The most successful parameter combinations are presented in this table. The full results are displayed in Appendix A. The first 4 lines describe parameters, the following 10 describe calibration statistics, the next 4 the calibration results (threshold and holding period) and the remaining the validation statistics. A cross-validation is considered successful if a statistically significant combination of threshold and holding period is found and trades are triggered in validation.

Lastly, both performance measures differ in the impact of the stop-loss and take-profit on the net profits. In both cases, the use of this risk management tool reduces the profits, but it is the case to a larger extent for Total Profit strategies. It is also noteworthy that using stop-loss and take-profit decreases the standard deviation of the average winrates across the selected strategies (from 8% to 6%) and increases the winrate on average (from 57% to 61%).

The results from the backtest reinforce the findings from the non-parametric estimation. The optimal trading rule threshold is in the same range as where the mean reversion becomes stronger according to the non-parametric estimation of the drift term (Figure 4). The backtest also indicates an optimal holding period of 4 days in most cases, in line with the Monte Carlo simulations. Another testament to the robustness of the mean reversion property of naphtha crack behind our trading strategies is that averaging the performance of the entire universe of 24 strategies still yields a positive return. Averaging validation results across all 24 tested strategies, we obtain \$0.26/bbl net profit per trade, \$0.45/bbl yearly net total profit, and 53.9% winrate when using total profit as a performance measure. When using profit per trade as a performance measure, we get \$0.19/bbl net profit per trade, \$0.23/bbl yearly net total profit, and 43.6% win rate.

5.4. Benchmark trading strategies

In order to further establish the significance of the performance of the selected strategies, we compare the results to two benchmark strategies. The first is a simple reversion strategy, and the second just a buy and hold strategy where a long position is taken until the end of the holding period. As in Lubnau and Todorova (2015), the entry points are selected randomly over the validation set. The number of trades is set to 3 per year in order to be comparable to the frequency of trading of the our selected strategies. The holding period is selected randomly between 1 and 5 days, again in line with the selected strategies. For each of the 5 validations samples, we perform $10,000$ runs of the two benchmark strategies and average the results. The reversion benchmark strategy performs the worst with an average net profit per trade of \$-0.22/bbl. The buy and hold benchmark strategy performs slightly better but still at a loss, with an average net profit per trade of \$-0.18/bbl. Contrary to strategies selected in the backtest, not using the stop-loss and take-profit rules results in even worst performance. Without the stop-loss and take-profit rules, the average net profit per trade decreases to \$-0.25/bbl and \$-0.20/bbl for the reversion and the buy and hold benchmark strategies. Despite evidence of mean reversion at any level from the non-parametric estimation of Section 4.2, the reversion benchmark strategy does not deliver positive net profits even if ignoring the transaction costs. The buy and hold benchmark strategy net profits are close to 0 but only when ignoring transaction costs.

5.5. Potential causes of short term market inefficiencies

The fact that our selected strategies are able to deliver robust positive returns after transaction costs implies the presence of at least temporary market inefficiencies where the Law of One Price is not fully enforced, similarly to the discussion in Gatev et al. (2006). In the present section, we propose three causes for the presence of short term market inefficiencies affecting the naphtha crack. Firstly, one potential reason is the difference in liquidity and the different players active in the naphtha swaps and the Brent futures markets that we discussed in Section 3. Brent futures are traded mostly on electronic platforms with a large variety of market participants including algorithmic and high frequency traders. On the other hand, naphtha swaps trade virtually exclusively through brokers and between a more limited pool of specialised players. The result is that naphtha prices are slower to react to information. Intra-day liquidity patterns might further exacerbate this characteristic. The greatest liquidity for naphtha swaps is observed during the Platts MOC assessment ending at 16:30 London time, while our daily naphtha crack prices are taken at 19:30 London time when Brent futures are settled. Therefore, we analyse the crack spread data at a point in time when the naphtha leg has had low liquidity for three hours while the Brent leg is still strongly active. It can result in the naphtha price to align itself back with Brent only the day after the Brent move ends, thus causing a short-term reversion in the naphtha crack. Secondly, another possible cause for the temporary disconnect can be the sometimes low liquidity of the naphtha market alone. Large volume trades in a context of low liquidity can cause the price to slip beyond its fundament value, and revert back to the fair price shortly after. Thirdly, as discussed in Section 3, naphtha is often traded as a spread to other products such as gasoline or propane, and is also affected by the oil versus natural gas relationship. Events in gasoline or propane markets can result in exaggerated short-term moves in naphtha crack prices, and result in a subsequent reversion to the fair value.

6. Conclusion

This paper investigates short term inefficiencies in the price of the naphtha crack, a mostly OTC traded security that corresponds to a long naphtha and a short Brent exposure. First, we describe the specificities of the naphtha versus the Brent markets. We describe the time series of interest, how we roll contracts and standardize daily price changes, and present the main statistical properties. We explain what trading fees and transaction costs are incurred when trading the naphtha crack.

Next, we model the dynamics of the standardised naphtha crack changes by means of drift and diffusion terms, defined as a function of the value of the standardised daily changes and assuming that the series follows an arithmetic Brownian motion. We estimate the drift and diffusion non-parametrically, in a similar way to Stanton (1997) and Prigent et al. (2001). Studying the drift coefficient, we observe a non linear increase in the rate of reversion when daily changes exceed about \$1/bbl for all three standardisation methods (20-day trailing standard deviation, GARCH(1,1) volatility, and 60-day trailing standard deviation). We run Monte Carlo simulations to get insights into the magnitude and duration of the reversion effect, and propose a method to convert the simulation results back to de-standardized returns. Our simulations show the mean reversion effect to be stronger than transaction costs for all three standardization methods, with most of the reversion taking place on the first day.

Finally, building on top of the findings from the non-parametric estimation, we backtest simple mean reversion trading strategies. Throughout this process, transaction costs are included. The optimal threshold for trigerring a trade and the holding period are defined in the calibration phase. The thresholds obtained in calibration correspond to the inflexion point in the non-paramtric drift term estimate, at between \$1.4/bbl and \$2/bbl. The optimal holding period is 4 days. Based on win rate and cross validation performance criteria, we are able to identify 7 strategies generating robust outperformance. The selected strategies generate on average an out-of-sample net profit per trade of \$0.62/bbl, a yearly total net profit of \$0.73/bbl, and a win rate of more than 60%. Further confirming the robustness of the mean reversion property of the naphtha crack, simply taking the average of all the strategies in our universe delivers positive returns. We also show that our strategies beat dummy benchmark strategies.

We identify three possible causes for short-term market inefficiencies affecting naphtha crack prices. The first is the differentce in reaction speed to market events due to differences in trading venues (OTC versus electronic) and number of market participants. The second cause is the low naphtha markets liquidity alone that can push prices for up to a day when large transactions are executed before prices correct back. The third potential cause is the fact that naphtha often trades as a spread to gasoline or propane. Sharp moves in gasoline or propane can temporarily shift the napthha crack price away from its fair value. Our findings of positive outperformance points towards the fact that the OTC market for naphtha can be inefficient in the short term. It constitutes an opportunity for the arbitrageur to correct these short-term inefficiencies, and get rewarded for enforcing the Law of One Price (Gatev et al., 2006). One area for future research would be to decompose the naphtha price movement into its naphtha and Brent components when a trade signal is trigerred, with particular focus on their evolution between the Platts MOC and Brent settlement. Investigating concurrent fluctuations propane and gasoline swaps would also provide further insights.

Appendix A. Comprehensive backtesting results

Strategy	cross_val_id	3	4	7	8	11	12	15	16	19	20	23	24
	scaling_type	std 20 days	std 20 days	std 20 days	std 20 days	std 60 days	std 60 days	std 60 days	std 60 days	garch 1,1	garch 1,1	garch 1,1	garch 1,1
	smoothing EWMA alpha	0.15	0.15	0.2	0.2	0.15	0.15	0.2	0.2	0.15	0.15	0.2	0.2
	core strategy	1	2	1	2	1	2	1	2	1	2	1	2
Calibration	average yearly number of trades	4.33	2.74	5.80	5.03	3.61	17.32	3.59	17.27	4.57	3.05	4.53	3.67
	median yearly number of trades	4.11	2.62	5.30	4.36	3.61	22.42	3.61	22.42	5.33	3.05	5.4	3.67
	average p_value in multiple testing	0.00078	0.00542	0.00140	0.00536	0.00075	0.00492	0.00040	0.00641	0.00032	0.00179	0.00053	0.00299
	median p_value in multiple testing	0.00032	0.00408	0.00067	0.00333	0.00037	0.00088	0.00005	0.00136	0.00030	0.00179	0.00049	0.00299
	average net profit per trade [$/bbl]	0.53	0.49	0.41	0.34	0.61	0.26	0.64	0.108	0.607	0.528	0.636	0.49
	median net profit per trade [$/bbl]	0.59	0.55	0.43	0.32	0.54	0.11	0.69	0.11	0.53	0.53	0.48	0.49
	average yearly total net profit [$/bbl]	2.09	1.37	2.12	1.55	2.17	1.69	2.30	1.63	2.41	1.61	2.27	1.50
	median yearly total net profit [$/bbl]	2.31	1.43	2.39	1.90	2.17	1.73	2.58	1.39	2.65	1.61	2.58	1.50
	average winrate	60%	67%	60%	67%	56%	63%	57%	62%	60%	66%	60%	70%
	median winrate	58%	65%	59%	71%	55%	62%	57%	64%	59%	69%	59%	74%
Calibration results	average threshold [$/bbl]	1.45	1.57	1.35	1.40	1.60	1.18	1.58	1.06	1.45	1.4	1.45	1.4
	median threshold [$/bbl]	1.60	1.70	1.40	1.30	1.70	1.00	1.70	1.00	1.45	1.4	1.45	1.4
	average holding period	4.0	4.3	4.0	4.3	4.0	4.2	4.0	4	4.25	4.5	4.25	4.5
	median holding period	4	4	4	4	4	4	4	4	4	4.5	4	4.5
Validation	average yearly number of trades	3.5	2.0	5.2	4.3	3.1	15.3	3.4	16.25	4	3.0	3.49	2.74
	median yearly number of trades	3.0	2.0	4.5	3.0	3.5	20.0	3.5	20.46	4.24	3	4.24	2.74
	average net profit per trade [$/bbl]	0.46	0.48	0.24	0.32	0.70	0.32	0.80	0.10	0.88	0.30	0.90	0.33
	median net profit per trade [$/bbl]	0.47	0.58	0.02	-0.07	0.65	-0.02	0.65	-0.02	0.29	0.3	0.32	0.33
	average yearly total net profit [$/bbl]	2.04	0.85	1.27	-0.09	2.47	0.16	2.58	0.87	0.92	0.73	0.98	0.38
	median yearly total net profit [$/bbl]	1.40	1.46	-0.16	-0.66	2.27	-0.36	2.27	-0.50	1.36	0.73	1.36	0.38
	average winrate	57.0%	75.0%	45.0%	42.0%	52.0%	50.0%	56.0%	51.0%	52.0%	67%	54%	55%
	median winrate	63.0%	80.0%	47.0%	39.0%	50.0%	53.0%	50.0%	51.0%	55.0%	69%	58%	58%
Validation with stop loss and take profit	average net profit per trade [$/bbl]	0.25	0.36	0.01	-0.18	0.36	0.31	0.45	-0.01	0.73	0.20	0.74	-0.14
	median net profit per trade [$/bbl]	0.27	0.58	-0.07	-0.13	0.65	-0.02	0.65	-0.02	0.15	0.2	0.18	-0.14
	average yearly total net profit [$/bbl]	1.04	0.68	0.06	-0.73	1.43	-0.04	1.54	0.13	0.54	0.47	0.60	-0.34
	median yearly total net profit [$/bbl]	1.19	1.31	-0.49	-0.94	2.20	-0.36	2.20	-0.50	0.76	0.47	0.87	-0.34
	average winrate	56.7%	65.6%	46.3%	40.7%	46.1%	58.7%	56.1%	50.9%	61.9%	59%	63%	42%
	median winrate	63.0%	67.0%	47.0%	39.0%	50.0%	53.0%	50.0%	51.0%	55.0%	59%	58%	42%
share of successful CV (out of 5)		100%	100%	80%	60%	100%	100%	100%	100%	80%	40%	80%	40%

Table A.3: Results from the 5-fold cross-validation with performance measure: Total Profit. The first 4 lines describe parameters, the following 10 describe calibration statistics, the next 4 the calibration results (threshold and holding period) and the remaining the validation statistics. A cross-validation is considered successful if a statistically significant combination of threshold and holding period was found and trades were triggered in validation.

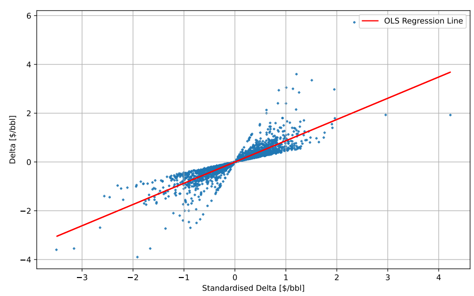
Strategy	cross_val_id	1	2	5	6	9	10	13	14	17	18	21	22
	scaling_type	std 20 days	std 20 days	std 20 days	std 20 days	std 60 days	std 60 days	std 60 days	std 60 days	garch 1,1	garch 1,1	garch 1,1	garch 1,1
	smoothing EWMA alpha	0.15	0.15	0.2	0.2	0.15	0.15	0.2	0.2	0.15	0.15	0.2	0.2
	core strategy	1	2	1	2	1	2	1	2	1	2	1	2
Calibration	average yearly number of trades	1.96	2.12	1.84	2.16	2.14	4.01	1.84	2.742	1.1075	0.835	0.97	0.76
	median yearly number of trades	2.37	2.62	1.81	2.74	1.74	2.12	1.74	2.12	1.32	0.84	1.04	0.76
	average p_value in multiple testing	0.00068	0.01267	0.00261	0.01741	0.00090	0.01751	0.00084	0.01754	0.00026	0.00400	0.00073	0.00647
	median p_value in multiple testing	0.00070	0.00408	0.00169	0.00541	0.00112	0.01976	0.00085	0.02191	0.00023	0.00400	0.00059	0.00647
	average net profit per trade [$/bbl]	0.95	0.62	0.98	0.62	0.92	0.49	1.01	0.5742	1.53175	1.1845	1.46075	1.092
	median net profit per trade [$/bbl]	0.86	0.59	0.92	0.58	0.95	0.46	0.95	0.49	1.55	1.18	1.41	1.09
	average yearly total net profit [$/bbl]	1.64	1.24	1.46	1.20	1.74	0.90	1.48	0.958	1.6425	0.97	1.4175	0.86
	median yearly total net profit [$/bbl]	1.90	1.43	1.39	1.37	1.54	0.82	1.54	0.98	1.84	0.97	1.6	0.86
	average winrate	67%	65%	63%	65%	60%	64%	63%	65%	91%	92%	89%	91%
	median winrate	71%	65%	67%	67%	63%	64%	64%	64%	95%	94%	94%	93%
Calibration results	average threshold [$/bbl]	1.80	1.77	1.85	1.80	1.90	1.76	1.94	1.82	1.95	1.95	1.875	1.8
	median threshold [$/bbl]	1.80	1.70	1.85	1.70	2.00	1.90	2.00	2	1.95	1.95	1.95	1.8
	average holding period	4.8	4.3	4.5	4.3	4.2	4.4	4.4	4.4	5	5	5	5
	median holding period	5	4	5	4	4	4	4	4	5	5	5	5
Validation	average yearly number of trades	1.6	1.5	1.6	1.5	1.8	3.5	1.7	2.792	1	0.25	1.1225	0.5
	median yearly number of trades	2.0	2.0	2.0	2.0	1.5	1.0	1.0	1	1	0.25	1	0.5
	average net profit per trade [$/bbl]	-0.06	0.07	0.07	0.07	0.84	0.33	0.84	0.34	-0.01	-0.743	0.772	0.122
	median net profit per trade [$/bbl]	-0.14	0.00	-0.05	0.00	0.66	0.20	0.66	0.20	0.00	-0.74	0.47	0.12
	average yearly total net profit [$/bbl]	0.02	0.24	0.22	0.24	0.78	0.04	0.84	0.23	0.35	-0.37	0.74	0.06
	median yearly total net profit [$/bbl]	-0.35	0.00	-0.10	0.00	1.20	0.20	1.20	0.20	-0.01	-0.37	0.78	0.06
	average winrate	39.0%	67.0%	46.0%	67.0%	50.0%	49.0%	53.0%	54.0%	50.0%	0%	56%	50%
	median winrate	10.0%	50.0%	25.0%	50.0%	75.0%	50.0%	75.0%	50.0%	17.0%	0%	54%	50%
Validation with stop loss and take profit	average net profit per trade [$/bbl]	-0.10	0.07	0.07	0.07	0.78	0.32	0.78	0.32	-0.02	-0.743	0.6765	0.122
	median net profit per trade [$/bbl]	-0.23	0.00	-0.05	0.00	0.66	0.20	0.66	0.20	0.00	-0.74	0.44	0.12
	average yearly total net profit [$/bbl]	-0.09	0.24	0.22	0.24	0.75	-0.11	0.81	0.07	0.32	-0.37	0.67	0.06
	median yearly total net profit [$/bbl]	-0.50	0.00	-0.10	0.00	1.20	0.20	1.20	0.20	-0.01	-0.37	0.7	0.06
	average winrate	25.0%	43.3%	32.5%	43.3%	62.5%	62.1%	62.5%	63.0%	27.1%	0%	52%	50%
	median winrate	10.0%	50.0%	25.0%	50.0%	75.0%	50.0%	75.0%	50.0%	17.0%	0%	54%	50%
share of successful CV (out of 5)		80%	60%	60%	40%	100%	100%	100%	100%	60%	20%	80%	40%

Table A.4: Results from the 5-fold cross-validation with performance measure: Profit per Trade. The first 4 lines describe parameters, the following 10 describe calibration statistics, the next 4 the calibration results (threshold and holding period) and the remaining the validation statistics. A cross-validation is considered successful if a statistically significant combination of threshold and holding period was found and trades were triggered in validation.

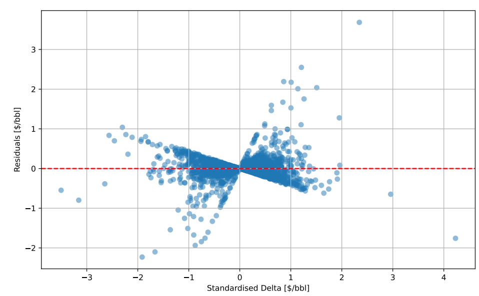
Appendix B. Converting simulated standardized changes back to raw changes

Dividing the time series of standardised daily naphtha crack price changes by its volatility enhances stationarity by mitigating the impact of heteroscedasticity. This transformation stabilizes the statistical properties of the series over time, making it more amenable to analysis and modeling. Given that the aim of such an analysis is to develop profitable trading strategies, and that these strategies are evaluated on the grounds of their performance in relation to the raw Delta, there is a need to de-standardise the standardised Delta. To do so, we employ an Ordinary Least Squares (OLS) regression and generate stochastic errors based on a functional Johnson's S_U distribution (Johnson, 1949) fitted to the residuals. From this point until the end of this section, we use the Delta standardised by the 60-day rolling historical standard deviation with an $\alpha = 0.2$ EWMA smoothing as an example for the procedure used.

We then divide the range of standardised Delta into 20 buckets, each containing the corresponding residuals. For each bucket containing more than 30 values, we fit a Johnson's S_U distribution to model the residuals. The 4 parameters of the distribution are plotted as a function of the bucket midpoint in Figure C.9. For a given standardised Delta, we estimate a de-standardised Delta by first computing an OLS prediction of the de-standardised Delta and by adding an error term drawn from a Johnson S_U distribution, which parameters are defined as a function of the standardised Delta via linear fit or interpolation. The definition of this distribution can be found in Appendix C. Figure B.8 shows the de-standardised Delta obtained through this procedure for the 60-day rolling historical standard deviation case.



(a) Scatter plot of the Delta as a function of the standardised Delta



(b) Scatter plot of the OLS residuals as a function of the standardised Delta

Figure B.7: Standardised Delta to Delta OLS regression and residuals visualisation. For the scatter plot of the Delta as a function of the standardised Delta, an OLS model is fitted and shown as a red line. For the scatter plot of the OLS residuals as a function of the standardised Delta. The red dashed line is the 0-line, referring to the OLS model. The standardisation series used in this example is the 60-day historical rolling standard deviation.

Appendix C. The Johnson's S_U distribution

The probability density function (PDF) of the Johnson's S_U distribution is given by:

$$f(x;\gamma,\delta,\xi,\lambda) = \frac{\delta}{\lambda\sqrt{2\pi}} \frac{1}{\sqrt{1+\left(\frac{x-\xi}{\lambda}\right)^2}} \exp\left(-\frac{1}{2}\left[\gamma + \delta \sinh^{-1}\left(\frac{x-\xi}{\lambda}\right)\right]^2\right),$$

where γ is the shape parameter controlling skewness, δ is the shape parameter controlling kurtosis, ξ is the location parameter, and λ is the scale parameter.

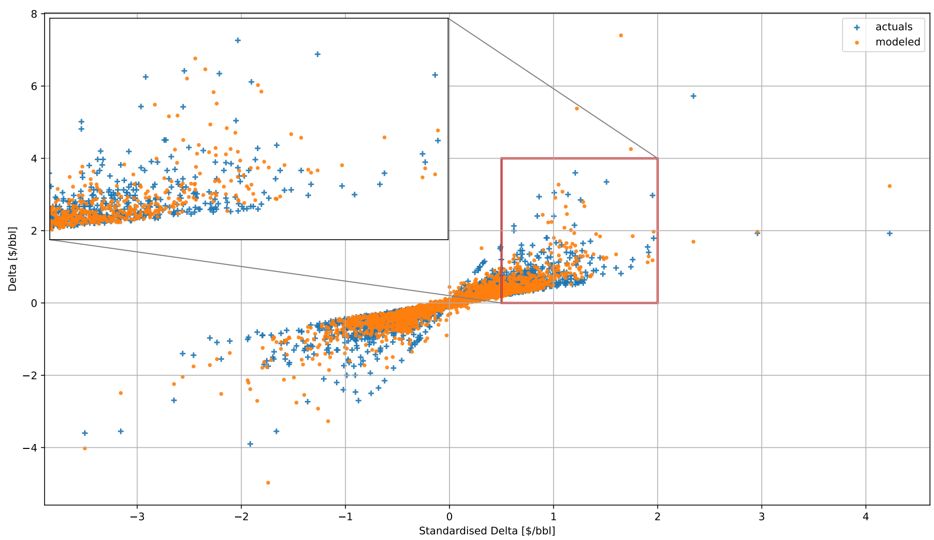
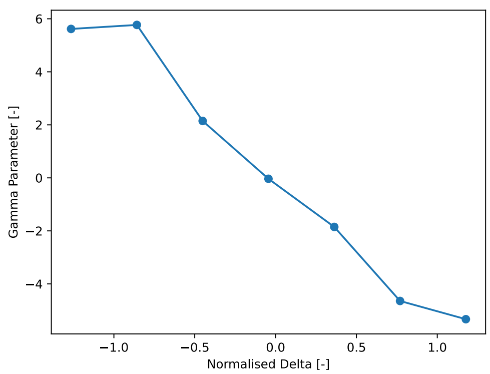
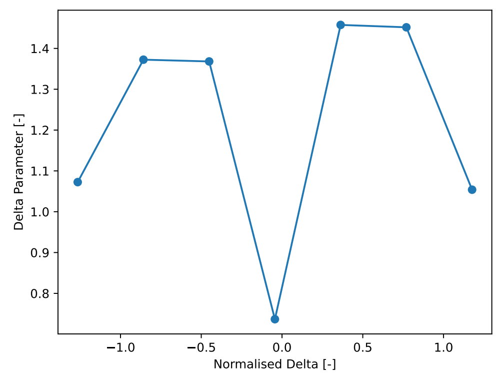


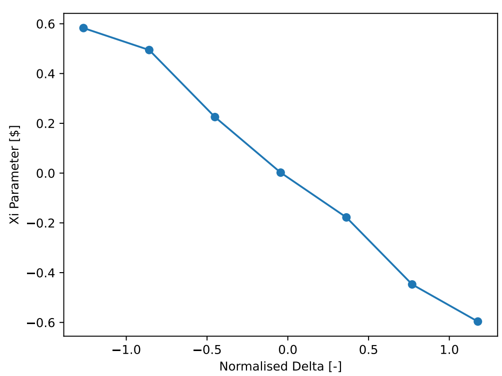
Figure B.8: De-standardisation of the standardised Delta. The blue + markers represent the true association of the standardised Delta to Delta. The orange • markers represent the stochastically generated (via fitted OLS and Johnson's S_U sampled residuals) association. The top left frame is an inset plot showing the same data with a greater resolution. The standardisation series used in this example is the 60-day historical rolling standard deviation.



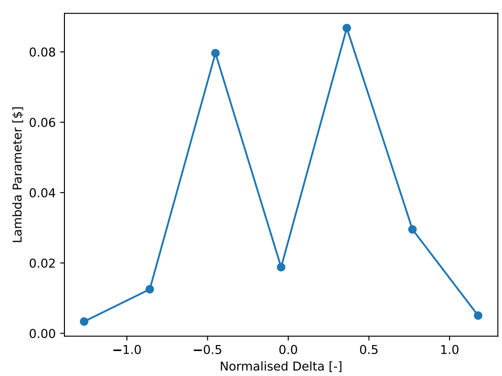
(a) Gamma parameter



(b) Delta parameter



(c) Xi parameter



(d) Lambda parameter

Figure C.9: Parameters of the fitted Johnson's S_U distributions as a function of standardised Delta.

References

Babus, A. and Kondor, P. (2018). Trading and Information Diffusion in Over-the-Counter Markets, *Econometrica*, 86(5), pp. 1727–1769. doi: https://doi.org/10.3982/ecta12043.

Bajgrowicz, P. and Scaillet, O. (2012) Technical trading revisited: False discoveries, persistence tests, and transaction costs, *Journal of Financial Economics*, 106(3), pp. 473–491. doi: https://doi.org/10.1016/j.jfineco.2012.06.001.

Bellia, M., Girardi, G., Panzica, R., Pelizzon, L. and Peltonen, T. (2024) The demand for central clearing: To clear or not to clear, that is the question!, *Journal of Financial Stability*, 72(6), pp. 101247–101247. doi: https://doi.org/10.1016/j.jfs.2024.101247.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), pp. 289–300. doi: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 31(3), pp. 307–327. doi: https://doi.org/10.1016/0304-4076(86)90063-1.

Brooks, C. and Prokopczuk, M. (2013) The dynamics of commodity prices, *Quantitative Finance*, 13(4), pp. 527–542. doi: https://doi.org/10.1080/14697688.2013.769689.

Cartea, Á., Chang, P., Mroczka, M. and Oomen, R. (2022). AI-driven liquidity provision in OTC financial markets. *Quantitative Finance*, pp.1–34. doi: https://doi.org/10.1080/14697688.2022.2130087.

Ehlers, T. and Hardy, B. (2019). The evolution of OTC interest rate derivatives markets, *BIS Quarterly Review*, December 2019, pp. 69–82. Available at: https://www.bis.org/publ/qtrpdf/r_qt1912i.htm [Accessed 7 Aug. 2024].

The European Commission. *Consolidated text: Commission Delegated Regulation (EU) No 151/2013 of 19 December 2012 supplementing Regulation (EU) No 648/2012 of the European Parliament and of the Council on OTC derivatives, central counterparties and trade repositories, with regard to regulatory technical standards specifying the data to be published and made available by trade repositories and operational standards for aggregating, comparing and accessing the data.*

2024. [Online] Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02013R0151-20240429 [Accessed 5 Aug. 2024].

The European Parliament and The Council. *Consolidated text: Regulation (EU) No 648/2012 of the European Parliament and of the Council of 4 July 2012 on OTC derivatives, central counterparties and trade repositories*. 2022. [Online] Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02012R0648-20220812 [Accessed 5 Aug. 2024].

Frino, A., Ibikunle, G., Mollica, V. and Steffen, T. (2018). The impact of commodity benchmarks on derivatives markets: The case of the dated Brent assessment and Brent futures, *Journal of Banking and Finance*, 95, pp. 27–43. doi: https://doi.org/10.1016/j.jbankfin.2017.08.017.

Gatev, E., Goetzmann, W.N. and Rouwenhorst, K.G. (2006). Pairs Trading: Performance of a Relative-Value Arbitrage Rule. *Review of Financial Studies*, [online] 19(3), pp.797–827. doi: https://doi.org/10.1093/rfs/hhj020.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*. New York: Springer, p. 81. Available at: https://link.springer.com/book/10.1007/978-0-387-21617-1 [Accessed 6 Aug. 2024].

Glode, V. and Opp, C.C. (2019). Over-the-Counter versus Limit-Order Markets: The Role of Traders' Expertise. *The Review of Financial Studies*, 33(2), pp.866–915. doi: https://doi.org/10.1093/rfs/hhz061.

He, C., Wang, T., Liu, X. and Huang, K. (2023). An innovative high-frequency statistical arbitrage in Chinese futures market. *Journal of innovation and knowledge*, 8(4), pp.100429–100429. doi: https://doi.org/10.1016/j.jik.2023.100429.

Holmes, M.J., Otero, J. and Panagiotidis, T. (2013). On the dynamics of gasoline market integration in the United States: Evidence from a pair-wise approach. *Energy economics*, 36(3), pp.503–510. doi: https://doi.org/10.1016/j.eneco.2012.10.008.

Huang, Z. and Martin, F. (2018). Pairs trading strategies in a cointegration framework: back-tested on CFD and optimized by profit factor, *Applied Economics*, 51(22), pp. 2436–2452. doi: https://doi.org/10.1080/00036846.2018.1545080.

Johnson, N.L. (1949). Systems of Frequency Curves Generated by Methods of Translation, *Biometrika*, vol. 36, no. 1-2, pp. 149-176. doi: https://doi.org/10.1093/biomet/36.1-2.149. [Accessed 8 Aug. 2024].

Kim, T. and Kim, H.Y. (2019). Optimizing the Pairs-Trading Strategy Using Deep Reinforcement Learning with Trading and Stop-Loss Boundaries. *Complexity*, 2019(1), pp.1–20. doi: https://doi.org/10.1155/2019/3582516.

Lee, T. and Wang, C. (2018). Why Trade Over-the-Counter? When Investors Want Price Discrimination, *SSRN Electronic Journal*. doi: https://doi.org/10.2139/ssrn.3192484.

Lubnau, T. and Todorova, N. (2015). Trading on mean-reversion in energy futures markets. *Energy Economics*, 51(9), pp.312–319. doi: https://doi.org/10.1016/j.eneco.2015.06.018.

Mahringer, S. and Prokopczuk, M. (2015). An empirical model comparison for valuing crack spread options. *Energy Economics*, 51, pp.177–187. doi: https://doi.org/10.1016/j.eneco.2015.06.015.

Murat, A. and Tokat, E. (2009). Forecasting oil price movements with crack spread futures. *Energy Economics*, 31(1), pp.85–90. doi: https://doi.org/10.1016/j.eneco.2008.07.008.

Ohana, S. (2010). Modeling global and local dependence in a pair of commodity forward curves with an application to the US natural gas and heating oil markets. *Energy Economics*, 32(2), pp.373–388. doi: https://doi.org/10.1016/j.eneco.2009.08.015.

Prigent, J.-L., Renault, O. and Scaillet, O. (2001) An empirical investigation into credit spread indices, *The Journal of Risk*, 3(3), pp. 27–55. doi: https://doi.org/10.21314/jor.2001.047.

Schwartz, E.S. (1997). The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging, *The Journal of Finance*, 52(3), pp. 923–973. doi: https://doi.org/10.1111/j.1540-6261.1997.tb02721.x.

Shen, L., Shen, K., Yi, C. and Chen, Y. (2020). An Evaluation of Pairs Trading in Commodity Futures Markets, in *2020 IEEE International Conference on Big Data (Big Data)*, pp. 5457-5462. doi: https://doi.org/10.1109/BigData50022.2020.9377766.

Shin, D.H. (2021). Optimal Pairs Trading Strategy under Geometric Brownian Motion and its Application to the US stocks, *International Journal for Innovation Education and Research*, 9(5), pp. 550–560. doi: https://doi.org/10.31686/ijier.vol9.iss5.3125.

Stanton, R. (1997). A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk. *The Journal of Finance*, 52(5), pp.1973–2002. doi: https://doi.org/10.1111/j.1540-6261.1997.tb02748.x.

Vaitonis, M. and Masteika, S. (2017) Statistical Arbitrage Trading Strategy in Commodity Futures Market with the Use of Nanoseconds Historical Data, *Communications in Computer and Information Science*, 756, pp. 303–313. doi: https://doi.org/10.1007/978-3-319-67642-5_25.