Quantitative Research in High Frequency Trading for Natural Gas Futures Market

Saulius Masteika(✉) and Mantas Vaitonis

Faculty of Humanities, Vilnius University, Muitines 8, 44280 Kaunas, Lithuania
saulius.masteika@khf.vu.lt, mantas.vaitonis@khf.stud.vu.lt

Abstract. High frequency trading (HFT) in micro or milliseconds has recently drawn attention of financial researches and engineers. In nowadays algorithmic trading and HFT account for a dominant part of overall trading volume. The main objective of this research is to test statistical arbitrage strategy in HFT natural gas futures market. The arbitrage strategy attempts to profit by exploiting price differences between successive futures contracts of the same underlying asset. It takes long/short positions when the spread between the contracts widens; hoping that the prices will converge back in the near future. In this study high frequency bid/ask and last trade records were collected from NYMEX exchange. The strategy was back tested applying MatLab software of technical computing. Statistical arbitrage and HFT has given positive results and refuted the efficient market hypothesis. The strategy can be interesting to financial engineers, market microstructure developers or market participants implementing high frequency trading strategies.

Keywords: High frequency trading · Algorithmic trading · Futures market · Statistical arbitrage

1 Introduction

One of the least discussed aspects on the financial science base is the high frequency trading (HFT). HFT execute thousands of orders a second and alter strategies in a matter of milliseconds. High frequency trading typically refers to trading activity that employs extremely fast automated programs for generating, routing, canceling, and executing orders in electronic markets [1]. HF algorithms trade in and out of positions in milliseconds and leave flat positions at end of the day. Recent expansion of HFT mostly relates to adoption of electronic trading and algorithmic model-based order generation in the exchanges, which exceeds 70–80 % of overall market trading activity [2]. HFT is widespread in all the most important markets of the world, like stocks, currencies, futures, options, and other derivatives [3, 4]. The dominant HFT strategies contribute to market liquidity, i.e. market making strategies [5, 6], and to price discovery and market efficiency, i.e. arbitrage strategies [7, 8]. The shortage of hard scientific evidences about the profitability of HFT algorithms was a driving force for this study. A contribution of this paper lies in testing HFT and statistical arbitrage strategy with natural gas futures

W. Abramowicz (Ed.): BIS 2015 Workshops, LNBIP 228, pp. 29–35, 2015.
DOI: 10.1007/978-3-319-26762-3_3

contracts, one of the most liquid instruments in global energy market [9, 10]. This paper presents the details and the results of the research.

2 The Basic Algorithm of Statistical Arbitrage in HFT

The roots of statistical arbitrage can be traced back to the first hedge funds, round 1950, running portfolio hedging strategies which had long and short positions that helped to reduce market risks [11]. The hedged portfolio as well as statistical arbitrage is profitable when long position earns more or loses less than the short position. Statistical arbitrage is defined as a long horizon trading opportunity that generates a riskless profit. More important, statistical arbitrage is defined without reference to any equilibrium model therefore, its existence is inconsistent with market equilibrium and, by inference, market efficiency [12]. Recent reports reveal great importance of electronic market making, as filed with the Securities and Exchange Commission, Virtu Financial, Inc. was able to perform only with one day losing out of 1238 days using electronic trading strategies [13].

Pair's trading is one of the basic strategies of statistical arbitrage and has been widely used by professional traders, institutional investors, and hedge fund managers. The strategy takes advantage of market inefficiencies based on a pair of correlated assets, e.g. stocks. The perception is to identify two stocks that move together and to take long and short positions simultaneously when they diverge abnormally [14]. It is expected that the prices of correlated financial instruments will converge back to a mean in the near future [15, 16]. Hence, the algorithm of statistical arbitrage in HFT needs to take highly correlated assets, e.g. the same commodity futures contracts of different delivery months [17]. The following chapter presents methodology of the algorithm in more detail.

2.1 Methodology for Pairs Trading Strategy

The algorithm of pairs trading observes the pair and waits till the spread between the prices of the contracts widens. When the opportunity occurs the algorithm bets for the contracts to return to their previous spread.

Pairs trading algorithm consists of the following parts:

- Selection of contracts
- HF data normalization
- Setting triggers for a long/short position

Futures trade in series of short-lived contracts that are only active for a few months [18]. The most active contracts of the same underlying asset were selected for statistical arbitrage. The contracts chosen were the closest to delivery months in order to trade in the most liquid market.

HF data normalization was done as follows, for each futures contract price $P(i,t)$ we calculated empirical mean $\mu(i,t)$ and standard deviation $\sigma(i,t)$ and then applied the following equation [10, 16]:

$$p(i,t) = \frac{P(i,t) - \mu(i,t)}{\sigma(i,t)} \tag{1}$$

The value *p(i,t)* is the normalized price of asset *i* at time *t*. The normalization is necessary in order to implement the pairs trading method. The same normalization method was implemented by M.Perlin using normalized unit data to create sell/buy signals [10, 16]. The method presented by M.Perlin creates a trading signal every time the absolute value of the spread of the normalized data is higher than a given prespecified threshold *d* [16]. Therefore the parameter *d* must be tested carefully, because it cannot be too high, in this case only few trading signals would be created, and cannot be too low, because to many trading signals would be created, what would result in poor signal's quality. In this research a bid price of one contract, i.e. *Natural Gas October* contract, and ask price of the other, i.e. *Natural Gas November* contract, were collected. Thus, whenever the distance of the given *d* is reached the algorithm shorts the commodity contract with the bid price and takes a long position of the contract of the ask price. Limit orders on market prices were used in HFT, which lowers the risk of order execution slippage to a minimum. The logic behind this HFT algorithm is that when the distance between the two contracts is too large, there is a higher probability that such prices are going to converge back in the future, and such situation can be put for profit purpose. Hence, this study tries to find out if pair's trading can be profitable with correlated futures contracts in HFT environment and whether higher *d* values increase the quality of the arbitrage and generates better results.

3 Experimental Setup

HFT statistical arbitrage calculations were implemented with the MathWorks language of technical computing Matlab. High frequency trading data, i.e. Henry Hub Natural Gas Futures, October - *NGV12* and November - *NGX12* contracts, was taken from New York Mercantile Exchange during the period from 04/09/12 till 20/09/12. The frequency of data was in milliseconds. An average number of normalized records were round 350000 per day, and reaching nearly 5 million data records in total. Time stamp sequences of correlated contracts were synchronized with no data loss or data distortions. The prices of correlated contracts were recalculated to a normalized unit, Eq. (1); moving data window for price normalization, i.e. μ and σ, was set to 100. A fragment of raw data stream is presented in the following figure:

In Fig. 1 *X* axis represents high frequency records of NGV and NGX contracts measured in milliseconds. *Y* axis shows price changes of each contract during the period of time.

The most active periods of the day were chosen for HFT statistical arbitrage. Trading activity was calculated as an average every 15 min during a period of one month. The most active trading periods during a day are presented in Fig. 2:

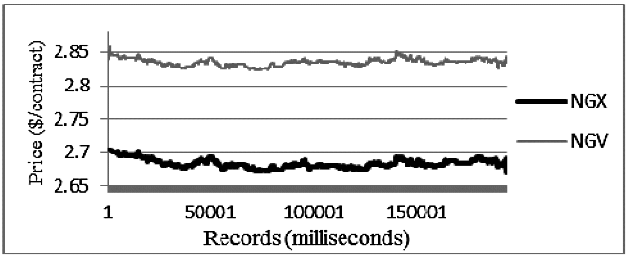

Fig. 1. Futures contract NGV and NGX high frequency data stream example

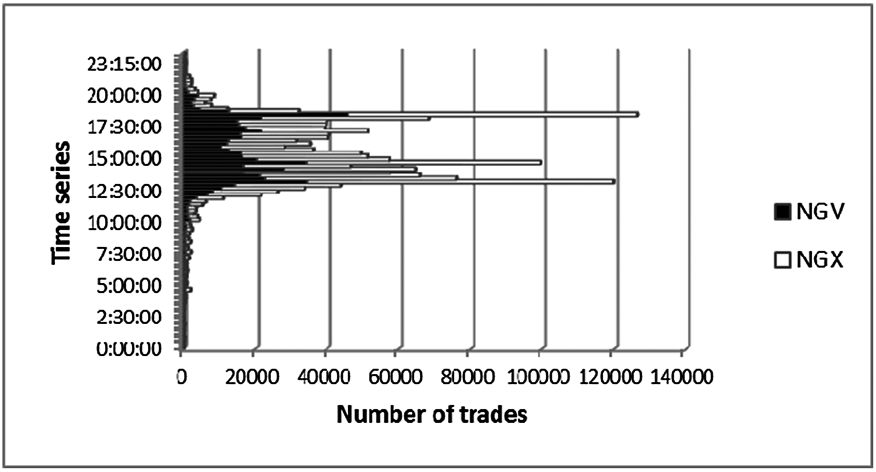

Fig. 2. HF trading activity during a day

In Fig. 2 *Y axis* represents time periods during a day and *X axis* shows the number of high frequency records every 15 min. The charts in Fig. 2 reveal that the most active period for high frequency trading lasts from 11.15 till 18.45. This period was chosen for the experimental research of statistical arbitrage strategy in high frequency trading environment.

4 Experimental Results

Experimental research has revealed that the best results were generated with the following parameters: the distance between normalized prices, i.e. parameter *d,* was calculated as 6.5, when maximal period to keep positions open was no longer than 1000 time series. In end-of-day trading and his research M. S. Perlini has noticed that lower values of the parameter *d* tend to generate better results [16]. However, we have found that in HFT higher value of parameter *d* gives better results and brings to more qualitative signals. We have found that lower threshold value *d* in HFT creates a bigger amount of trading signals and leads to overtrading. In our opinion the reason higher *d* works for us, is that we were not swapping between assets, i.e. correlated stocks, and weren't using bidirectional statistical arbitrage. In our research the *bid* prices were taken of one future's contract and *ask* of the other. The calculations revealed that the value 6.5 of parameter *d* was optimal for each day during a period of 13 days and further increase of parameter *d* stopped generating adequate positive results. Experimental results and profitability of statistical arbitrage are presented graphically in the following figure:

The chart in Fig. 3 shows that during the period from 4th till 20th of September the strategy of statistical arbitrage generates positive results each day while trading Henry Hub Natural Gas Futures. The biggest increase per contract was nearly 0.4 percent per day and the smallest- about 0.05 %/day. More statistics are presented in Table 1.

Table 1. Results of statistical arbitrage in HFT

Profit from long positions	1,14 %
Profit from sell short positions	1,28 %
Total long/short signals	3334
STD on return per trade	0,01 %
Average return per trade	0,16 %
Total profit	2,42 %

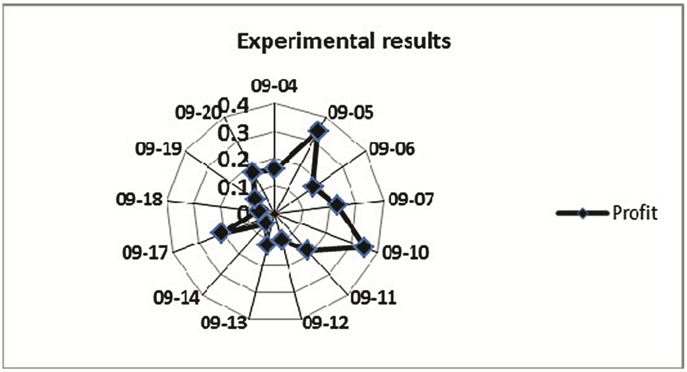

Fig. 3. Results of statistical arbitrage in HFT

Table 1 shows that higher profits were generated from sell short positions if compared to long positions. It revealed that during the period of two weeks there were up trend movement in the natural gas futures market and traders more often had overshot with buying orders, therefore statistical arbitrage and correction of the market was more profitable with short sell positions. Total profit from all trading days was 2.42 % and 3334 trades were generated.

5 Conclusions

The experimental research did prove that the pair of futures contracts allow to exclude pair selection procedures of the algorithm, increase the speed of calculations, and can be used for HFT with higher values of the parameter d. It has revealed that HFT gives positive results and therefore can be interesting for market infrastructure developers and market participants. Nevertheless before applying the strategy in real market conditions some additional aspects must be considered in order to find the best possible application of the algorithm, e.g. bidirectional arbitrage, leverage trading, HFT infrastructure costs and rebates from brokers and exchanges for HFT liquidity provision.

References

1. Cvitanic, J., Kirilenko, A.A.: High frequency traders and asset prices. SSRN (2010). http://ssrn.com/abstract=1569067
2. Zubulake, P., Lee, S.: The High frequency game changer: how automated trading strategies have revolutionized the markets. Wiley, Aite group (2011)
3. Netherlands Authority for the Financial Markets. High frequency trading: The application of advanced trading technology in the European marketplace. http://www.afm.nl/~/media/files/rapport/2010/hft-report-engels.ashx
4. Hagströmer, B., Norden, L.: The diversity of high-frequency traders. J. Finan. Markets. **16**(4), 741–770 (2013)
5. Gomber, P., Arndt, B., Lutat, M., Uhle, T.: High-Frequency Trading. Goethe University, Deutsche Börse Group (2011)
6. Menkveld, A.J.: High frequency trading and the new market makers. J. Finan. Markets. **16**(4), 712–740 (2013)
7. Aldridge, I.: High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems. Wiley, Hoboken (2013)
8. Hanson, T.A., Hall, J.R.: Statistical arbitrage trading strategies and high frequency trading. SSRN (2012). http://ssrn.com/abstract=2147012
9. Acworth, W.: Record Volume 2010 (Annual Volume Survey). Futures Industry, 12–29 March 2011
10. Driaunys, K., Masteika, S., Sakalauskas, V., Vaitonis, M.: An algorithm-based statistical arbitrage high frequency trading system to forecast prices of natural gas futures. Transform. Bus. Econ. **13**(3), 96–109 (2014)
11. Eichengreen, J.: Barry.: Hedge funds and financial market dynamics. International Monetary Fund (1998)
12. Hogan, S., Jarrow, R., Teo, M., Warachka, M.: Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. J. Finan. Econ. **73**, 525–565 (2004)
13. Virtu Financial, Inc. FORM S-1. US securities and exchange commission. https://www.sec.gov/Archives/edgar/data/1592386/000104746914002070/a2218589zs-1.htm
14. Miao, J.: George.: High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. Int. J. Econ. Finan. **6**(3), 96–110 (2014)
15. Caldeira, J.F., Moura, G.V.: Selection of a portfolio of pairs based on cointegration: a statistical arbitrage strategy. Revista Brasileira de Financas **11**(1), 49–80 (2013)

16. Perlin, M.S.: Evaluation of Pairs-trading strategy at the Brazilian financial market. J. Deriv. Hedge Funds **15**(2), 122–136 (2009)
17. Masteika, S., Rutkauskas, A.V.: Research on futures trend trading strategy based on short term chart pattern. J. Bus. Econ. Manage. **13**(5), 915–930 (2012)
18. Masteika, S., Driaunys, K., Rutkauskas, A.V.: Historical data formation for back test and technical analysis in North American futures market. Transform. Bus. Econ. **12**(1A), 473–488 (2013)