ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics
Bachelor Thesis Economics and Business Economics

Cointegration based pairs trading during COVID-19: trading thresholds and loss limitation.

Name student: Govert Bettman
Student ID number: 521628

Supervisor: Dr. S. Vogel
Second assessor: Dr. F. Core

Date final version: 5 July 2022

Abstract

Pairs trading is a form of statistical arbitrage, which exploits co-movements of stocks. I evaluate cointegration based pairs trading in the U.S. during the COVID-19 bear market. Daily data on S&P 500 stocks over 2019 are employed to identify 20 cointegrated pairs for trading, using methodology proposed by Engle and Granger (1987). Various trading thresholds based on historical standard deviations are assessed in the 2020 bear market, and the separate implementation of a stop-loss condition is reviewed. Contrary to prior literature, I find no evidence for the profitability of pairs trading after subtracting transaction costs. Neither varying trading thresholds nor implementing a loss limitation condition results in reliable profits. Although these findings question if pairs trading is still relevant, opportunities for improvement appear plenty and pairs trading should not be written off. I primarily direct to optimizing trading thresholds and loss limitation rules.

Keywords: Statistical arbitrage; Pairs trading; Cointegration; Trading thresholds; Loss limitation.

Table of Contents

1. Introduction

In their continuous search for returns, investors have always been interested in the concept of arbitrage. This follows from the characteristics of arbitrage: profits which, in theory, require no capital and involve no risk (Shleifer & Vishny, 1997). Statistical arbitrage is quantitatively based and pairs trading is one of its most prominent applications. The first practical implementation is often attributed to Nunzio Tartaglia, who led a trading group at Morgan Stanley in the 1980s.

Standard pairs trading is simple and intuitive, as described by Gatev, Goetzmann and Rouwenhorst (1999, 2006). It involves two securities whose prices have moved together historically, implying a long run equilibrium relationship. The spread between these security prices can diverge in the short-term due to a variety of reasons, but the securities are tied in the long run and are thus expected to converge back to their long-term equilibrium spread. If the short-term spread widens beyond a certain threshold, we should short the relative ‘winner’ and buy the relative ‘loser’. These positions are reversed when the prices converge back to their long-term equilibrium spread. Successful execution implies profits without net investment or risk.

The mechanism underlying pairs trading is related to various concepts described in the literature. The efficient market hypothesis by Fama (1970) rules out systematic excess returns. However, many challenge this and find systematically profitable strategies, refer to Grossman and Stiglitz (1980), Bondt and Thaler (1985), and Jegadeesh and Titman (1993). Also related, are contrarian investment strategies, refer to Lakonishok, Shleifer and Vishny (1994). Applications of contrarian strategies rely on mechanisms like mean reversion, as used by Poterba and Summers (1988), and overreaction, as applied by Bondt and Thaler (1985). Likewise related to pairs trading is the Law of One Price, refer to Lamont and Thaler (2003) for a definition. Furthermore, cointegration is linked with pairs trading due to the implied long-run relationship. Important literature on cointegration is written by Engle and Granger (1987) and Johansen (1988). Finally, pairs trading is related to statistical arbitrage and its characteristics (Avellaneda & Lee, 2010).

In the empirical literature employing pairs trading strategies, five streams can be identified (Krauss, 2017). The time-series approach is led by Elliot, Van der Hoek, and Malcolm (2005) and the stochastic control approach by Jurek and Yang (2007). The category of ‘other approaches’ contains machine learning, outlined by Huck (2009, 2010), copula approaches, led by Liew and Wu (2013) and Xie and Wu (2013), and the principal components analysis approach, presented by Avellaneda and Lee (2010). Due to various issues with these approaches, this paper focuses on the empirical findings of the distance approach, pioneered by Gatev et al. (1999, 2006), and cointegration approach, described by Vidyamurthy (2004).

Examiners of the distance and cointegration approaches observe two major phenomena. First, both approaches show decaying returns since the 1980s and 1990s, refer to Gatev et al. (1999, 2006), Do and Faff (2010, 2012), and Rad, Low and Faff (2016). Second, returns from pairs trading strategies pick up in the bear markets of 2000-2002 and 2007-2009. This is found by Do and Faff (2010, 2012), Huck and Afawubo (2015), Huck (2015), and Rad et al. (2016).

In this paper, the cointegration approach is applied because of the superior statistical foundation compared to the distance approach (Lin, McCrae & Gulati, 2006). Furthermore, the cointegration approach typically outperforms the distance approach (Huck & Afawubo, 2015; Huck, 2015; Rad et al., 2016).

The aforementioned literature finds an interesting pattern of decaying returns which pick up in times of bear markets. The 2000-2002 and 2007-2009 bear markets have been extensively studied. However, this does not apply to the most recent bear market in 2020, resulting from the COVID-19 pandemic and its turmoil. The occurrence of this bear market is a chance to test if cointegration based pairs trading strategies are still profitable. This paper aims to do exactly this, and the main research question can be stated as follows:

Will cointegration based pairs trading strategies achieve profits in the U.S. during the 2020 bear market?

To gain further insights in how profits from pairs trading strategies can be optimized, I will look at two means to alter trading rules. First, the trigger for trading can be varied. There appears to be a trade-off between the number of trades and the returns per trade. The literature often uses two or three historical standard deviations, refer to Gatev et al. (1999, 2006), Do and Faff (2010, 2012), Huck and Afawubo (2015), Huck (2015), and Rad et al. (2016). This paper uses five levels and attempts to determine which trading threshold performs best. Second, the addition of a loss limitation condition could avoid major losses, which is observed by Rad et al. (2016) and Huck and Afawubo (2015). Stop-loss mechanisms have been proposed by Nath (2003) and Caldeira and Moura (2013). I will assess if a new stop-loss condition improves pairs trading performance.

First, this study contributes to the existing literature by evaluating cointegration based pairs trading in a more recent bear market. Hence, I asses if the decay in returns over time and uptick in bear markets, as observed in the literature, continues into the 2020s. Next, I add a broader review of trading thresholds to the literature since five distinct levels are evaluated. The implementation of a new loss limitation condition adds to the scarce literature surrounding capping extreme losses in pairs trading. Furthermore, this research contributes to a larger body of literature surrounding the efficiency of financial markets, as pairs trading should not be systematically profitable if markets are efficient. For investment practitioners, the results of this paper can point to opportunities for statistical arbitrage or discourage allocating resources to these strategies.

To evaluate cointegration based pairs trading strategies in the 2020 bear market, I use daily stock data from the Center for Research in Security Prices (CRSP). Data for all stocks in the Standard & Poor (S&P) 500 are collected from 2 January 2019 until 30 June 2020. Following Vidyamurthy (2004), I use three steps in developing a cointegration based pairs trading strategy. First, I rank all pairs based on the lowest sum of squared deviations (SSD) between the normalized price series over 2019. Second, pairs are tested for cointegration in order of lowest SSD, using the Engle and Granger (1987) methodology, until 20 pairs have been identified for trading. Third, trading rules are designed based on the number of historical standard deviations the spread deviates from its historical mean. I evaluate 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations. Using these rules, the cointegrated pairs are traded over the first half of 2020. Net profits or losses are found after trading costs. To evaluate loss limitation, I repeat the methodology including a stop-loss condition when the spread deviates 6.0 historical standard deviations from its historical mean.

This paper finds no statistically significant evidence for profits from cointegration based pairs trading in the U.S. during the 2020 bear market when no loss limitation is included. Varying the trading threshold also provides no profitable strategy in this case. Implementing the stop-loss condition appears to improve performance, yet no statistically significant evidence for profitability is brought forward. Nevertheless, the potential of loss limitation is evident. For investors, the results of this paper might appear worrisome for the future of pairs trading. However, opportunities in the subjects of optimal trading thresholds and loss limitation could mitigate these concerns. Regarding academia, excess returns in this paper reject the efficient market hypothesis.

The remainder of this paper is organized as follows. The second chapter displays a review of the literature on underlying concepts, empirical results, trading thresholds and loss limitation. This framework is the basis for the hypotheses. Next, in the third chapter, the data, sample selection and methodology used for the trading strategy is described. The fourth chapter displays and discusses the performance of pairs trading. Finally, the fifth chapter will present conclusions.

2. Literature review

This chapter reviews the literature surrounding pairs trading. The first section discusses underlying concepts and their link with pairs trading. The second section reviews the empirical literature on pairs trading, including methodological streams and empirical patterns. Next, the third section discusses trading thresholds and the fourth section loss limitation. Finally, the fifth section clarifies this papers contribution to the literature and the sixth section shows the hypotheses development.

2.1 Underlying concepts

The main underlying concept for all trading strategies is market (in)efficiency. The fundamental work of Fama (1970) summarizes prior findings and formulates the efficient market hypothesis. This hypothesis states that asset prices fully reflect all available information on the market and, consequently, nobody can earn systematic risk-adjusted excess returns. However, several papers find evidence of inefficiencies and possibilities for excess returns, refer to Grossman and Stiglitz (1980), Bondt and Thaler (1985), and Jegadeesh and Titman (1993) for examples. Some market inefficiency is a prerequisite for profitable pairs trading, otherwise excess returns are unachievable.

Closely related to pairs trading are contrarian investment strategies. The setup of contrarian strategies is buying ‘losing’ and shorting ‘winning’ securities. This strategy is widely covered in the literature and appears profitable. Notably, Lakonishok et al. (1994) find evidence that abnormal returns from contrarian strategies are caused by mistakes of individual and institutional investors and not by more risk. Other studies attribute profits to mean reversion and overreaction.

Poterba and Summers (1988) find mean reversion in stock prices in the United States from 1871 until 1986. This indicates opportunities for contrarian strategies, as buying ‘losers’ and shorting ‘winners’ would be profitable due to expected mean reversion of both securities. Gatev et al. (1999, 2006) split the effect of pairs trading from the effect of general mean reversion. They find that trading random pairs, relying on general mean reversion, does not appear profitable. However, trading statistically selected pairs does seem to be a lucrative strategy.

Central in the literature around overreaction is the study by Bondt and Thaler (1985), who find empirical evidence for overreaction in security markets. This causes temporary overvaluation after good news and undervaluation after bad news. Buying undervalued and shorting overvalued securities creates a profitable contrarian strategy because overreaction is usually temporary.

Another related concept is the Law of One Price. Lamont and Thaler (2003) provide a simple definition: identical securities should have identical prices. Close substitutes should, therefore, have stable long-term relative pricing. Short-term deviations from this pricing are expected to disappear. The short-term relative ‘winner’ should be shorted, and the relative ‘loser’ should be bought (Gatev et al., 1999, 2006). If the Law of One Price holds, long-term relative pricing will be restored, and profits are made.

Pairs trading is also related to cointegration. Key methods for identification are presented by Engle and Granger (1987) and Johansen (1988). Cointegration implies that a spread between stocks eventually converges back to its long-run equilibrium. Alexander (1999) stresses that cointegration and correlation should be separated. Cointegration identifies long-run relationships, but correlation does not necessarily imply this relationship. Alexander (1999) is one of the first to show that long-short positions in cointegrated portfolios exploit long-run co-movements.

Finally, pairs trading can be perceived as a type of statistical arbitrage. Following Avellaneda and Lee (2010), all types of statistical arbitrage have three common features. First, trading signals are based on systematic rules, in contrast to trading based on fundamentals. Second, the trading strategies are market neutral and, therefore, have no systematic risk. Lastly, the mechanism driving the returns is statistically based. Pairs trading clearly shares these features.

2.2 Empirical literature on pairs trading

In the past two decades, a rich body of empirical literature regarding pairs trading has been developed. Attention for this strategy appeared after the study by Gatev et al. (1999), which is the most cited paper in this field. Gatev et al. (1999) find positive returns in the U.S. with simple pairs trading rules, using daily data from 1962 until 1997. They implement a two-stage strategy. First, co-moving pairs are identified during a 12-month formation period, which is parametrized by selecting pairs that minimize the SSD between the price series. This is known as the distance approach in literature. Second, during the 6-month trading period, long-short positions in pairs are opened if the prices diverge by more than two historical standard deviations, which are calculated over the formation period. The positions are closed if the price series cross. This strategy achieves average annualized excess returns of 12 percent for the top performing pairs, exceeding their estimation of transaction cost. As mentioned, Gatev et al. (1999) split the results of pairs trading from general mean reversion. Next, they examine the development of pairs trading performance over time and find worse results in the 1990s compared to 1970s and 1980s. Later, the authors extend their research to the period 1962-2002 and confirm their prior findings (Gatev et al., 2006).

After the studies by Gatev et al. (1999, 2006), the literature expanded quickly, and several streams developed. Krauss (2017) reviews the developments from this point in time and identifies five streams: the time-series approach, stochastic control approach, distance approach, cointegration approach, and ‘other’ approaches. These streams will now be examined separately.

Elliot et al. (2005) lead the time-series approach and suggest to describe the spread between price series, observed in Gaussian noise, using a mean reverting Gaussian Markov chain model. On the stochastic control approach, the prominent study by Jurek and Yang (2007) derives the optimal asset allocation for nonmyopic arbitrageurs between pairs trading and a risk-free asset. A crucial note is that the time-series and stochastic control approaches assume that co-moving pairs have already been identified and, therefore, ignore the formation period. This contrasts with the aim of my paper, which is to find statically sound pairs and assess the profitability from trading. Hence, the time-series and stochastic control approaches will not be considered further.

The category of ‘other’ approaches contains several sub streams. First, the machine learning and combined forecasts approach is based on three steps: forecasting, outranking, and trading (Huck, 2009, 2010). Next, the copula approach is split into the returns-based method, refer to Liew and Wu (2013), and the level-based method, refer to Xie and Wu (2013). Finally, the principal components analysis approach is presented by Avellaneda and Lee (2010). This category of ‘other’ approaches has limited support in prior literature and often suffers from computational complexity. For these reasons, they will not be considered in further detail here. I will now focus on the distance approach, the cointegration approach and comparisons between these two.

As mentioned, Gatev et al. (1999, 2006) present the premier application of the distance approach. Recall the pairs selection based on SSD and opening of long-short positions if prices diverge by two standard deviations, resulting in positive returns with a declining trend over time. Do and Faff (2010) replicate the methodology of Gatev et al. (1999, 2006) to assess profitability in an extended sample. They observe the decline in returns found in Gatev et al. (1999, 2006) and continuation into the 2000s. However, excess returns remain positive and significant. The authors attribute the decline partly to increased competition in the hedge fund industry and market efficiency, but mainly to worsening arbitrage risks. Interestingly, they find improved results in the

2000-2002 and 2007-2009 bear markets. Next, Do and Faff (2010) find higher returns when they combine SSD with homogeneity metrics and number of historic spread reversals.

According to Do and Faff (2012), the distance approach-based papers by Gatev et al. (1999, 2006) and Do and Faff (2010) do not consider all types of trading costs. Therefore, the authors re-evaluate these papers, now considering commissions, market impact and short selling fees. They find that the returns found in Gatev et al. (1999, 2006) and Do and Faff (2010) are seriously hurt, emphasizing the importance of trading costs in pairs trading. In line with Gatev et al. (1999, 2006) and Do and Faff (2010), the authors find a decrease in returns over time. Noteworthy, the returns picked up during the 2000-2002 and 2007-2009 bear markets.

The publication by Vidyamurthy (2004) is the most cited concerning the cointegration approach. His paper lacks empirical application but provides a framework, consisting of three steps, for cointegration based pairs trading strategies in later studies. First, pairs should be selected based on statistical or fundamental similarity. Next, these pairs need to be tested for cointegration as proposed by Engle and Granger (1987). Finally, trading thresholds are designed.

Lin et al. (2006) advocate using cointegration in pairs trading, emphasizing the importance of statistical foundation for pairs trading decisions. Next, they introduce integrating a minimum profit threshold in their trading model. Using the cointegration approach based on Engle and Granger (1987), they find steady positive excess returns from 2001 until 2002 in Australia.

Several studies employ both the distance approach and cointegration approach to compare results. Huck and Afawubo (2015) evaluate both approaches using the S&P 500 stocks over the period 2000-2011. After controlling for trading costs, risk factors and data dredging, the distance approach shows insignificant excess returns. By contrast, the cointegration approach, based on Johansen (1988), obtains significant excess returns of 1.38% per month. In line with the papers by Do and Faff (2010, 2012), the authors find high performance in the 2007-2009 bear market.

Next, Huck (2015) compares the distance and cointegration approaches over the period 2003 until 2013 for both S&P 500 and Nikkei 225 stocks. The distance approach usually records insignificant excess returns, while the cointegration approach, based on Johansen (1988), achieves significant excess returns of at least 1.5% per month. Further, the author finds no effect of volatility timing on returns from the cointegration approach. Notably high performance in the 2007-2009 bear market is consistent with Do and Faff (2010, 2012) and Huck and Afawubo (2015).

Finally, Rad et al. (2016) compare the two approaches in the U.S. during the period 1962-2014. The authors employ the method by Engle and Granger (1987) for the cointegration approach. They find significant and similar excess returns after transaction costs for the cointegration and distance approaches, 0.33% and 0.38% on average per month, respectively. In line with Gatev et al. (1999, 2006) and Do and Faff (2010, 2012), they find decreasing returns over time and continuation into the 2010s. Consistent with Do and Faff (2010, 2012), Huck and Afawubo (2015), and Huck (2015), pairs trading excels in bear markets. Lastly, the authors find superior excess returns per month for the cointegration approach during crises compared to the distance approach.

2.3 Trading thresholds

As mentioned, I focus on the distance and cointegration approaches onwards, which usually employ standard deviations of the historical spread as base for trading rules. If the observed spread diverges by more than several historical standard deviations, long-short positions are opened. Intuitively, a larger number of standard deviations leads to higher triggers and less trades (Huck & Afawubo, 2015). However, if the spread then converges to its equilibrium, returns are higher. This leads to a trade-off between more trades and more returns per trade (Vidyamurthy, 2004).

The literature employs various levels of triggers. Gatev et al. (1999, 2006) and Do and Faff (2010, 2012) use two standard deviations as a trigger in their distance approaches. Comparing the distance and cointegration approaches, Rad et al. (2016) use two historical standard deviations too. Huck and Afawubo (2015) and Huck (2015) compare both approaches using two as well as three historical standard deviations. Both studies find equal returns for the distance approach, however, the cointegration approach shows better performance using three historical standard deviations.

2.4 Loss limitation

Profits from pairs trading follow from either of two situations (Nath, 2003). First, after positions are opened, the spread converges, and the positions are closed. Second, opened positions do not fully converge and are closed because the end of the trading period is reached. Note that in both situations returns should exceed transaction costs to record profits. As Rad et al. (2016) mention, extreme profits do not occur because these are capped by the size of the opening trigger, often several historical standard deviations. Losses are recorded because of two reasons. First, returns of fully, or partially at the end of the trading period, converged pairs can be lower than transactions costs. Rad et al. (2016) remark that less volatile spreads, with low standard deviations, are an important reason for convergence with a loss. Note that these losses are capped by the size of transaction costs. Second, losses can occur because of non-convergence after the first opening or after multiple openings. Therefore, these positions are closed at the end of the trading period. Note that these losses are not capped and can become extreme. Rad et al. (2016) and Huck and Afawubo (2015) observe extreme losses and recommend using a stop-loss mechanism in future research.

Applications of stop-loss mechanisms to limit extreme losses are scarce in the literature. Nath (2003) uses a stop-loss trigger whenever the spread widens between 5^{th} and 95^{th} percentiles in a distance approach-based trading strategy. Caldeira and Moura (2013) use a stop-loss where positions are closed if the losses exceed 7%. They identify premature closings, due to crossing the stop-loss boundary, of eventually convergent pairs as clear disadvantage of a stop-loss. Besides these papers, stop-loss applications are scarce and no consensus about the best methodology exists.

2.5 Contribution

The literature surrounding the distance and cointegration approaches is most relevant and exhibits two trends. Gatev et al. (1999, 2006), Do and Faff (2010, 2012), and Rad et al. (2016) find that both approaches became far less profitable since the 1980s and 1990s and this persists into the 2010s. Next, several papers find an exception in this trend during bear markets like 2000-2002 and 2007-2009, refer to Do and Faff (2010, 2012), Huck and Afawubo (2015), Huck (2015), and Rad et al. (2016). Thus, distance and cointegration based pairs trading appears to thrive in previously studied bear markets. The most recent bear market was caused by the COVID-19 pandemic during the first half of 2020 and marks a new chance to test the profitability of pairs trading. By assessing the 2020 bear market, I add a new measurement point of pairs trading performance to the literature.

Next, this paper contributes to the literature by evaluating a broader range of trading triggers. Gatev et al. (1999, 2006), Do and Faff (2010, 2012), and Rad et al. (2016) only assess two historical standard deviations. Huck and Afawubo (2015) and Huck (2015) evaluate both two and three historical standard deviations. This paper will try to conclude on the trade-off between more trades and more return per trade by documenting the performance of trading rules based on 1.0, 1.5, 2.0, 2.5, or 3.0 historical standard deviations.

Finally, this paper will add to the scarce literature on loss limitation in pairs trading. Extreme losses due to non-convergence are observed by both Rad et al. (2016) and Huck and Afawubo (2015). Nath (2003) and Caldeira and Moura (2013) propose different stop-loss mechanisms to solve this problem but no consensus on the best method is reached in the literature. This paper evaluates if a new and simpler stop-loss method based on historical standard deviations prevents extreme losses and contributes to a more profitable pairs trading strategy.

2.6 Hypotheses development

The decreasing trend in pairs trading performance started in the 1980s and 1990s and is likely to continue due to increasing market efficiency, hedge fund competition and worsening arbitrage risks (Do & Faff, 2010). Regardless of this trend, pairs trading remains profitable in recent bear markets according to most of the research. An explanation is more mispricing in volatile markets, which is ideal to exploit for pairs trading strategies (Do & Faff, 2010, 2012; Huck, 2015). Therefore, I expect pairs trading to remain profitable in 2020 and hypothesis 1 is formulated:

H1. *Cointegration based pairs trading strategies will achieve profits in the U.S. during the 2020 bear market.*

Besides, limited research regarding the effect of different trading thresholds on pairs trading performance has been conducted. Huck and Afawubo (2015) and Huck (2015) find that, for the cointegration approach, three historical standard deviations outperform two. Related to this could be the remark of Rad et al. (2016). They note that returns might not exceed transaction costs for pairs with low historical spread volatility, and thus a small standard deviation. Consequently, I expect higher trading triggers to result in better performance and hypothesis 2 is formulated:

H2. *Higher trading thresholds will result in more profitable cointegration based pairs trading strategies.*

Finally, the research about loss limitation is scarce. Extreme losses are observed by Rad et al. (2016) and Huck and Afawubo (2015) and are caused by opened positions in pairs that become non-convergent. Contrasting, profits are limited by the size of the initial trigger. Losses can therefore dominate capped profits. Huck and Afawubo (2015) identify this as the main reason for weak performance. Implementing a stop-loss could prevent extreme losses, but premature closing of pairs may create a counter effect. Yet, I expect the stop-loss threshold, 6.0 standard deviations in this paper, to improve pairs trading performance. Thus, I formulate hypothesis 3 as follows:

H3. *Implementing the stop-loss threshold will result in more profitable cointegration based pairs trading strategies.*

3. Data and Methodology

In this chapter, the data and methodology used to create a profitable pairs trading strategy are described. The first section explains the data source and sample selection. Next, in the second section, the methodology is illustrated. Here, I will first motivate which approach is used and then illustrate the pairs selection, cointegration testing, trading rules, net profit calculation and loss limitation condition. The third section will display the descriptive statistics of the sample used.

3.1 Data and sample selection

The stock data used in this paper are retrieved from the Wharton Research Data Services (WRDS). Daily data on the stock prices of all S&P 500 companies is gathered from the CRSP database. The S&P 500 stocks are among the most liquid in the world and, therefore, transactions costs are likely to be low (Huck & Afawubo, 2015; Huck, 2015). Hence, these stocks are suitable for pairs trading. The S&P 500 constituents on 2 January 2019 are considered, which I derived from Bloomberg.

Daily data on stock prices is retrieved for the period 2 January 2019 until 30 June 2020. It is widespread practice in the literature to use a pairs formation period of one year and a subsequent trading period of 0.5 years, refer to Gatev et al. (1999, 2006), Do and Faff (2010, 2012) and Rad et al. (2016). In this study, the formation period ranges from 2 January 2019 until 31 December 2019 and the trading period from 2 January 2020 until 30 June 2020. To account for stock splits and dividends payments in this period, the daily stock prices are adjusted using the Cumulative Factor to Adjust Price (CFACPR) variable from the CRSP database.

Note that some S&P 500 constituents lack data for parts of the period 2 January 2019 until 30 June 2020. Often, this is due to mergers and acquisitions. If stocks lack data, they are dropped and, hence, the sample contains 479 stocks. When I normalize the stock price time series, the prices on 2 January 2019 will be equal to one dollar for all stocks. Besides, the logarithmic stock price time series is created by taking the logarithm, with base 10, of the non-normalized stock prices.

3.2 Methodology

This section will illustrate the employed methodology. First, I motivate which approach is most suitable. Next, the specifics of the pairs selection, cointegration testing, trading rules and net profit calculation are explained. Finally, I exhibit the loss limitation condition.

3.2.1 Choosing an approach

The literature review identifies five main approaches. The time-series and stochastic control approaches ignore the formation period and are thus not suited for this study's aim. The 'other' approaches are not chosen since support in the literature is scarce and they are computationally complex. Hence, the distance and cointegration approaches remain. The cointegration approach provides stronger statistical foundation for the existence of long run equilibrium relationships, compared to the distance approach (Lin et al., 2006). In addition, the cointegration approach achieves superior returns in recent studies, refer to Huck and Afawubo (2015) and Huck (2015). Furthermore, Rad et al. (2016) conclude that the cointegration approach is superior in bear markets. Since my paper concerns the 2020 bear market, this finding is highly relevant. Consequently, the cointegration approach emerges as the most attractive in the context of this paper.

To employ the cointegration approach, I follow the framework outlined by Vidyamurthy (2004). First, he selects pairs based on statistical or fundamental similarity. Next, these pairs are tested for cointegration using the method of Engle and Granger (1987). Last, systematic trading rules should be determined. Now, I will explain these steps and show the profits calculation after.

3.2.2 Pairs selection

First, pairs are preselected based on statistical or fundamental similarity (Vidyamurthy, 2004). Since cointegration testing is time consuming and the sample contains 479 stocks, resulting in 114,481 unique pairs, preselection is necessary. For this purpose, Rad et al. (2016) provide the following approach. Suppose that a pair consists of two stocks, i and j. Both i and j have a normalized stock price time series, given by $P_{i,t}$ and $P_{j,t}$, respectively. Time series $P_{i,t}$ contains observations $P_{i,0}, \dots, P_{i,T}$ and time series $P_{j,t}$ observations $P_{j,0}, \dots, P_{j,T}$. Here, $t = 0$ denotes the first day of the formation period and $t = T$ denotes the last day. For all possible pairs i and j the SSD is calculated over the formation period, using equation 1.

$$SSD_{i,j} = \sum_{t=0}^{T} \left(P_{i,t} - P_{j,t}\right)^2 \tag{1}$$

Next, all possible pairs are ranked based on their SSD (lowest first). Starting form the first ranked pair, I will test all pairs for cointegration in order of their SSD ranking until 20 cointegrated pairs have been identified. In this regard, I employ the distance approach for pairs preselection before testing for cointegration.

3.2.3 Testing for cointegration

Vidyamurthy (2004) prescribes to test for cointegration using the approach offered by Engle and Granger (1987). While Rad et al. (2016) and Lin et al. (2006) use this method, Huck and Afawubo (2015) and Huck (2015) employ the approach of Johansen (1988). Brooks (2019) states that the technique of Engle and Granger (1987) suffers from single equation issues, while the method by Johansen (1988) overcomes this. However, Engle and Granger (1987) provide a more intuitive and easier to implement method. Therefore, sticking with the framework of Vidyamurthy (2004), I employ the approach suggested by Engle and Granger (1987). Before I show the application of this method, it is crucial to understand the concept of cointegration and its function in pairs trading.

Following Engle and Granger (1987), time series variables which are stationary after differencing one time are integrated of order one, also called $I(1)$ or containing a unit root. Series which do not require differencing to be stationary are integrated of order zero and called $I(0)$ or not containing a unit root. Engle and Granger (1987) define cointegration in the case of two $I(1)$ variables as follows: if a linear combination of two $I(1)$ variables result in an error term that is $I(0)$, these variables are cointegrated. In this case, the error term is stationary and thus able to drift away from its mean but will often cross this mean, implying that a long run equilibrium state will occasionally arise. Drawing a parallel with pairs trading strategies, cointegration occurs when two $I(1)$ price series have a linear combination resulting in a spread between them which is $I(0)$. Therefore, the long run equilibrium spread between the price series will occasionally occur.

For each pair, the cointegration test is applied as follows. I use logarithmic stock price time series for easier returns calculations. To test if both logarithmic stock price series in the pairs are $I(1)$ over the formation period, I apply a DF test with a constant but no trend (Dickey & Fuller, 1979, 1981). This DF test is validated by the nature of stock returns, which are typically not serially correlated, heteroskedastic, and on average slightly positive (Cont, 2001). To check if stock price series are $I(1)$ and not integrated of a higher order, I apply the same DF test on returns since these are the first difference of stock prices. The returns should be $I(0)$ for the stock prices to be $I(1)$. If both stock price series in a pair are $I(1)$, the method of Engle and Granger (1987) can be applied.

Suppose that $p_{i,t}$ and $p_{j,t}$ are $I(1)$ and denote the logarithmic price time series of stocks i and j, respectively. Series $p_{i,t}$ contains observations $p_{i,0}, \dots, p_{i,T}$ and series $p_{j,t}$ contains $p_{j,0}, \dots, p_{j,T}$. Again, $t = 0$ denotes the first day and $t = T$ the last day of the formation period. A linear combination between $p_{i,t}$ and $p_{j,t}$ will be obtained using Ordinary Least Squares (OLS), where treatment as dependent or independent variable is assigned randomly. This results in equation 2 and, after rewriting, in equation 3.

$$p_{i,t} = \beta_0 + \beta_1 p_{j,t} + e_t \tag{2}$$

$$p_{i,t} - \beta_1 p_{j,t} - \beta_0 = e_t \tag{3}$$

The components of equations 2 and 3 can be interpreted as follows. β_1 is the cointegration coefficient and determines the proportion in which positions should be opened, more follows in section 3.2.4. β_0 is a constant and measures the equilibrium level spread between the logarithmic price series. Therefore, e_t can be interpreted as the deviation from the equilibrium level spread, with (almost) zero mean by definition. For cointegration between $p_{i,t}$ and $p_{j,t}$ to occur, e_t must be $I(0)$. Important to recognize is that e_t is estimated using OLS and consequently does not contain raw data. According to Engle and Granger (1987), the critical values of DF tests on e_t are thus different from those recorded by Dickey and Fuller (1979, 1981). To test if e_t is $I(0)$, I use a DF test without constant or trend and apply the critical values as reported by MacKinnon (1990, 2010).

If e_t is $I(0)$, stocks i and j form a cointegrated pair and the deviation from the equilibrium level spread is stationary. This implies that the spread between the logarithmic stock prices can drift from its equilibrium level in the short-term but is expected to come back to this level in the long-term. This is exactly what pairs trading will try to exploit. After selecting 20 cointegrated pairs, which thus show statistical foundation for a long run relationship, I will trade these pairs.

3.2.4 Trading rules

Following Vidyamurthy (2004), the last step is defining systematic trading thresholds. A common approach in the literature is using multiples of historical standard deviations as trading thresholds, refer to Gatev et al. (1999, 2006), Do and Faff (2010, 2012), Huck and Afawubo (2015), Huck (2015), and Rad et al. (2016). Because of the consensus in the literature, I will also use this method.

For each cointegrated pair, I retrieve the estimated cointegration coefficient ($\hat{\beta}_1$), equilibrium level spread ($\hat{\beta}_0$), and deviation from the equilibrium level spread ($\hat{e}_t$) in the formation period from equation 2. I calculate the mean ($\mu_{\hat{e}}$) and standard deviation ($\sigma_{\hat{e}}$) of $\hat{e}_t$ over the formation period. Note that $\mu_{\hat{e}}$ will be (almost) equal to zero since OLS is used in equation 2.

Recall that the trading period ranges from 2 January 2020 to 30 June 2020. Suppose that $p_{i,t}$ and $p_{j,t}$ denote the logarithmic price time series of stocks i and j, respectively. Series $p_{i,t}$

contains observations $p_{i,0}, \dots, p_{i,T}$ and series $p_{j,t}$ contains $p_{j,0}, \dots, p_{j,T}$. Now, $t = 0$ denotes the first day of the trading period and $t = T$ the last day. In the trading period, I will calculate the daily deviation from the equilibrium level spread ($\tilde{e}_t$) as follows. The spread on day t is calculated using the historical $\hat{\beta}_1$, as estimated over the formation period. After subtracting the historical $\hat{\beta}_0$, estimated over the formation period, $\tilde{e}_t$ in the trading period is found. This is shown in equation 4. Next, $\tilde{e}_t$ is standardized in equation 5 by $\mu_{\hat{e}}$ and $\sigma_{\hat{e}}$, which are calculated over the formation period.

$$p_{i,t} - \hat{\beta}_1 p_{j,t} - \hat{\beta}_0 = \tilde{e}_t \tag{4}$$

$$n_t = \frac{\tilde{e}_t - \mu_{\hat{e}}}{\sigma_{\hat{e}}} \tag{5}$$

The variable n_t denotes how many historical standard deviations $\tilde{e}_t$ differs from its historical mean. n_t is compared to certain trading thresholds, which are denoted by n^*. I will evaluate 1.0, 1.5, 2.0, 2.5 and 3.0 as values of n^* to find the optimal trading threshold.

During the trading period, long-short positions are opened in two situations. If $n_t > n^*$, I buy 10 dollars of stock j and short sell $10^{\frac{1}{\hat{\beta}_1}}$ dollars of stock i. In contrast, if $n_t < -n^*$, I buy 10 dollars of stock i and short sell $10^{\hat{\beta}_1}$ dollars of stock j. The positions are closed when $n_t = 0$, indicating that $\tilde{e}_t$ is equal to its historical mean and that the spread returned to its long-term equilibrium level. All open positions on the last day of the trading period are also closed.

3.2.5 Net profit calculation

To assess the trading results of the cointegrated pairs, the net profits need to be calculated. The first step is to calculate the returns. Note that returns represent the returns on committed capital in this paper. Therefore, returns equal to zero will be reported for pairs that do not trade during the trading period. Since $p_{i,t}$ and $p_{j,t}$ denote logarithmic price series, the returns are given by the difference in $\tilde{e}_t$ over time. The returns of a pair on day t are computed by equation 6.

$$Returns_t = |\tilde{e}_{t-1}| - |\tilde{e}_t| \tag{6}$$

The absolute values of $\tilde{e}_{t-1}$ and $\tilde{e}_t$ are used to account for both positive and negative deviations from the equilibrium level spread at the time positions are opened. Note that returns can only occur after positions are opened. There are no returns after the positions are closed until they are opened again. Note that the first day of the trading period cannot report returns. To calculate the returns of a pair over the trading period, the daily returns are summed, as shown in equation 7.

$$Total\ Return = \sum_{t=1}^{T} Returns_t \tag{7}$$

Because the returns are calculated over self-financing long-short positions, the returns following from equation 7 can be viewed as excess returns (Gatev et al., 1999, 2006). However, it should be noted that the cointegration approach results in different values for the long and short positions and, therefore, individual long-short positions are not always completely self-financing. Nevertheless, following Rad et al. (2016), I assume self-financing long-short positions on average.

Finally, we adjust the excess returns for trading costs. I follow Do and Faff (2012), who lead the literature and consider commissions, market impact and short selling fees. According to their research, commissions are 9 basis points (bps), and the market impact is 20 bps for both one-way positions in the most recent years (2007-2009). Following Rad et al. (2016), I do not consider short selling cost explicitly. This is validated by the liquid S&P 500 stocks considered in this paper. Transaction costs (29 bps) are doubled when a pairs trade is completed since this entails two round-trips. If a pair completes multiple pairs trades, the transaction costs are also multiplied by this number. After subtracting the total transaction cost from the excess returns, we find the net profit.

3.2.6 Loss limitation condition

As mentioned, losses due to non-converging pairs can become extreme and hurt pairs trading performance (Rad et al., 2016; Huck & Afawubo, 2015). In the above 'standard' methodology, maximum losses are not capped and can hamper profits. However, this prevents the premature closing of ultimately converging pairs (Caldeira & Moura, 2013).

To assess if capping extreme losses improves the profitability of pairs trading, I introduce a stop-loss mechanism. Since optimizing stop-loss mechanisms is not the goal of this paper, only one level will be assessed. Different from the 5^{th} and 95^{th} percentile spread approach of Nath (2003) and the 7% loss threshold from Caldeira and Moura (2013), I will base the stop-loss on historical standard deviations. This serves simple implementation of the method in practice. Further, required standard deviations are known because of their adoption in the opening triggers.

To assess the influence of a stop-loss threshold, I repeat the above 'standard' sample selection and methodology. Consequently, the same pairs will be traded over the same period with the same opening thresholds. However, now, positions will not only be closed when $n_t = 0$ or the last trading day is reached, but also if $n_t > 6$ or $n_t < -6$. This value doubles the highest trading threshold and is chosen arbitrarily. Recall that n_t denotes how many historical standard deviations $\tilde{e}_t$ differs from its historical mean. If either of these stop-loss thresholds are crossed once, the pair cannot be traded from that point on during the trading period. The underlying intuition is that if the deviation from the historical equilibrium spread is this large, this must imply that the cointegration relationship does not hold anymore. Net profit calculations are the same as before.

3.3 Descriptive statistics

The sample used to find profitable cointegration based pairs trading strategies consists of 479 stocks. Table 1 reports the descriptive statistics of this sample. As shown in Panel A, the average annualized returns over the formation period are positive and substantial. However, the average annualized returns over the trading period are negative and the standard deviation is larger than in the formation period. This points out the 2020 bear market, which is documented in this paper.

Table 1
Descriptive Statistics of Full Sample

This table shows the descriptive statistics for the full sample. The full sample contains all S&P 500 stocks on 2 January 2019, minus the stocks that lacked data. Therefore, the sample consists of 479 stocks. Panel A shows descriptive statistics on the average returns of the stocks over different periods. The formation period ranges from 2 January 2019 until 31 December 2019. The following trading period ranges from 2 January 2020 until 30 June 2020. Therefore, the entire period ranges from the first day of the formation period, 2 January 2019, until the last day of the trading period, 30 June 2020. Note that the reported returns over these periods are annualized for easier comparison. Panel B displays the proportion of different industries in the sample. The categories are based on the Standard Industrial Classification (SIC) code.

Panel A: Annualized returns over parts of the researched period						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Formation period	25.96	26.43	28.67	-86.28	308.82	479
Trading period	-20.38	-23.88	40.80	-92.20	179.03	479
Entire period	5.40	5.52	28.52	-87.15	220.41	479

Panel B: Distribution sample over industries		
SIC Code	Industry	Share of sample (%)
01-09	Agriculture, Forestry, Fishing	0.21
10-14	Mining	4.62
15-17	Construction	1.04
20-39	Manufacturing	35.43
40-49	Transportation & Public Utilities	12.90
50-51	Wholesale Trade	3.65
52-59	Retail Trade	7.23
60-67	Finance, Insurance, Real Estate	18.85
70-89	Services	13.95
91-99	Public Administration	2.14

Furthermore, the full sample includes all major industry categories. Manufacturing is represented the most with 35.43% and Agriculture, Forestry, and Fishing the least with 0.21%.

After testing pairs from the sample in order of lowest SSD, we find 20 cointegrated pairs including 35 individual stocks. Table A.1 in the Appendix provides an overview of the cointegrated pairs and their stocks. The descriptive statistics of these stocks are displayed in Table 2. The same discrepancy in average annualized returns between the formation period and trading period as in Table 1 is visible in Panel A. The stock prices of the cointegrated pairs appear to be hurt by the 2020 bear market. Besides, Panel B shows that the stocks in the cointegrated pairs are less diversified across industries than in the full sample. Most stocks are from the Transportation & Public Utilities (37.14%), Finance, Insurance, and Real Estate (28.57%), and Manufacturing (20.00%) industries. The relationship between cointegration and stocks from certain industries is an interesting future avenue but will not be further considered in this study.

Table 2
Descriptive Statistics of Stocks in Cointegrated Pairs

This table shows the descriptive statistics for the stocks in cointegrated pairs. The 20 cointegrated pairs include 35 individual stocks. Panel A shows descriptive statistics on the average returns of the stocks over different periods. The formation period ranges from 2 January 2019 until 31 December 2019. The following trading period ranges from 2 January 2020 until 30 June 2020. Therefore, the entire period ranges from the first day of the formation period, 2 January 2019, until the last day of the trading period, 30 June 2020. Note that the reported returns over these periods are annualized. Panel B displays the proportion of different industries in this sample. The categories are based on the Standard Industrial Classification (SIC) code.

Panel A: Annualized returns over parts of the researched period

	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Formation period	31.34	32.22	7.45	16.48	45.70	35
Trading period	-20.33	-20.43	25.62	-61.64	39.52	35
Entire period	9.57	10.56	13.17	-14.06	35.63	35

Panel B: Distribution sample over industries

SIC Code	Industry	Share of sample (%)
01-09	Agriculture, Forestry, Fishing	0.00
10-14	Mining	0.00
15-17	Construction	0.00
20-39	Manufacturing	20.00
40-49	Transportation & Public Utilities	37.14
50-51	Wholesale Trade	0.00
52-59	Retail Trade	0.00
60-67	Finance, Insurance, Real Estate	28.57
70-89	Services	8.57
91-99	Public Administration	5.71

4. Results and Discussion

This chapter displays and discusses the empirical results following from the methodology used in this paper. First, the profitability of the standard case will be assessed. Second, the best trading threshold will be evaluated. Finally, the impact of the loss limitation condition is investigated.

4.1 Profitability standard pairs trading strategy

The performances of pairs trading portfolios without the stop-loss are exhibited in Table 3. Note that the portfolios are assumed to be constituted by equally weighted pairs and, hence, the reported means display the returns and profits. Corresponding trading statistics are reported in Table A.2 in the Appendix. The pooled portfolio, shown in Panel A of Table 3, reports excess returns of 0.89% but after transaction costs the strategy loses 0.81%. Likewise, the portfolios based on trading thresholds in Panels B, C, D, E, and F of Table 3 all report positive excess returns, but negative profits after considering trading costs. Note that all results in Table 3 are statistically insignificant.

Therefore, in this study, cointegration based pairs trading without a stop-loss condition does not achieve statistically, or economically, significant profits in the U.S. during the 2020 bear market. Consequently, I do not find enough evidence to support hypothesis 1, when no loss limitation condition is included. This contrasts with the literature, since Do and Faff (2010, 2012), Huck and Afawubo (2015), Huck (2015), and Rad et al. (2016) all found statistically and economically significant profits from pairs trading in previous bear markets.

First, I will try to explain this discrepancy with context and time related factors. Contrary to my results, pairs trading strategies are expected to thrive in bear markets due to more security mispricing (Do & Faff, 2010, 2012; Huck, 2015). One interpretation of vanished profits in this paper could be the observed negative trend in performance since the 1980s and 1990s, which persisted into the 2010s (Gatev et al., 1999, 2006; Do & Faff, 2010, 2012; Rad et al., 2016). Partial causes for continuation of this decay in pairs trading profitability towards the 2020s could be increasing market efficiency and tougher competition amongst hedge funds (Do & Faff, 2010).

However, Do and Faff (2010) identify worsening arbitrage risks as the main driver for the historical decay and conclude that this is largely fuelled by the increasing risk of non-convergence after first opening. As shown in Panel A of Table A.2, 18% of pairs is non-converging after the first opening in the pooled portfolio. This rate is low compared to the 32% in the 2007-2009 bear market found by Do and Faff (2010). The low rate of non-convergence in this study should enhance performance; however, I find worse and statistically insignificant results. Hence, non-convergence after first opening is presumably not mainly driving inferior performance but can still influence returns due to extreme losses, regardless of their small share in total pairs.

Table 3
Performance Pairs Trading without Loss Limitation

This table shows the performance of different pairs trading portfolios based on various trading rules, all without the loss limitation condition included. Panel A shows the pooled performance of trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, Panel A has 100 observations. Panels B, C, D, E, and F display the performance of the 20 cointegrated pairs with trading rules based on 1.0, 1.5, 2.0, 2.5, or 3.0 historical standard deviations, respectively. These panels each have 20 observations. Each panel exhibits the corresponding excess return, which is equal to the returns on committed capital because I assume self-financing long-short portfolios on average. Note that these returns are not annualized, and, therefore, are recorded over the trading period. All panels also report the net profit over the trading period, which is reached after subtracting the transaction costs from the excess returns. Assumed transaction costs are 29 bps for the one-way opening of positions and are doubled when the pairs trade is closed. If a pair records multiple complete pairs trades, total transaction costs are multiplied by this number. Statistical significance of excess returns and net profits is denoted by ***, **, and * for 1%, 5%, and 10% levels, respectively.

Panel A: Performance pooled portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.89	0.55	6.23	-10.34	15.67	100
Net profits	-0.81	-0.37	5.51	-11.50	11.84	100
Panel B: Performance n^*=1.0 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	1.47	1.15	7.05	-9.31	15.67	20
Net profits	-0.99	-1.28	5.91	-10.24	11.03	20
Panel C: Performance n^*=1.5 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.43	1.22	6.22	-10.34	10.68	20
Net profits	-1.42	-0.81	5.53	-11.50	7.78	20
Panel D: Performance n^*=2.0 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.72	0.62	5.97	-9.27	10.68	20
Net profits	-0.90	-0.60	5.36	-9.89	7.78	20
Panel E: Performance n^*=2.5 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.82	0.41	6.27	-9.68	13.26	20
Net profits	-0.51	-0.29	5.63	-10.26	10.36	20
Panel F: Performance n^*=3.0 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	1.02	0.64	6.20	-8.93	14.74	20
Net profits	-0.22	-0.15	5.60	-9.51	11.84	20

Next, performance in this paper can be hampered by non-convergence of pairs after one or more round-trip trades are completed. Panel A from Table A.2 shows that the average number of openings (2.94) is substantially higher than the number of closings (2.13), excluding forced closing at the end of the trading period. This is partially explained by non-converging pairs after first opening, but these only make up 18.00% of total pairs. Consequently, the multiple opening pairs, which make up 77.00%, explain the smaller number of closings. These pairs have recorded at least one completed round-trip trade, but their last opening is often not closed. Non-convergence after the last opening hurts profits by definition. Partial convergence can be profitable but will often hurt profits due to the illustrated convergence with a loss. A possible reason for non-convergence is the expiration of the cointegration relationship of pairs after the sudden impact of the COVID-19 pandemic in February and March 2020. This could explain successful round-trip trades at the start of the trading period followed by non-convergence.

Additionally, I will try to explain the discrepancy with the literature using methodological factors. First, the difference in performance between papers based on the approaches by Johansen (1988) and Engle and Granger (1987) is remarkable. Huck and Afawubo (2015) and Huck (2015) employ the former and find statistically significant returns of 1.38% and 1.5% per month, respectively. Like this paper, Rad et al. (2016) employ the latter and find statically significant returns of only 0.33% per month. Hence, worse performance in this study could be related to employing the approach of Engle and Granger (1987), which suffers from single equation issues (Brooks, 2019). However, singling out this effect from other factors is tricky and no clear evidence exists since both methodologies have never been compared directly.

Next, poor performance can be explained by pairs converging with a loss, which could be unintentionally reinforced by methodological choices in this study. These pairs fully, or partially, converge after opening and thus record positive excess returns by definition. However, if returns, capped by the opening trigger, are dominated by transaction costs, losses are recorded. Pairs with low spread volatility, and thus low opening triggers, are perhaps prone to this issue in my study (Rad et al., 2016). Section 4.2 will expand on the impact of various opening triggers. As discussed by Krauss (2017), convergence with a loss could be amplified by first ranking pairs based on lowest SSD. Consequently, pairs with low spread variances are chosen and, hence, the selected pairs in this paper are more likely to be vulnerable to converging with a loss. Related is the estimation of transaction costs. The employed estimation by Do and Faff (2012) in this paper can be seen as conservative, since 29 bps is based on 2007-2009. Given the declining historical trend, transaction costs are expected to be lower in the 2020 bear market. Applying these higher transaction costs fuels the chances of converging with a loss but ensures a conservative estimation.

4.2 Different trading thresholds

Table 3 reports the performances of various trading thresholds, without the loss limitation condition implemented. The corresponding trading statistics can be found in Table A.2. In both Tables, Panels B, C, D, E, and F report the results for 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations, respectively. As mentioned, all portfolios in Table 3 report positive excess returns but losses after transaction costs. The portfolio based on 3.0 historical standard deviations performs 'best' in this sample and reports a statistically insignificant loss of 0.22%. In order of profitability, the portfolios based on 2.5, 2.0, 1.0, and 1.5 historical standard deviations report statistically insignificant losses of 0.51%, 0.90%, 0.99%, and 1.42% in this sample, respectively.

In this study, the portfolios based on the highest trading thresholds (2.5 and 3.0 historical standard deviations) record losses of smaller economic significance than the other portfolios, when no loss limitation condition is included. However, these results lack statistical significance. Hence, without loss imitation, I do not find enough evidence to support hypothesis 2. This contrasts with the literature since Huck and Afawubo (2015) and Huck (2015) find better performance for larger trading thresholds, which cannot be proven statistically in this study. Nevertheless, the association between higher trading thresholds and lower losses in this sample appears similar to what Huck and Afawubo (2015) and Huck (2015) discovered.

The similarity in observed patterns will mostly be explained by the fact that Huck and Afawubo (2015), Huck (2015) and this paper all employ historical standard deviations as trading trigger. Therefore, it is interesting to unravel why this method for establishing trading triggers results in this pattern.

Important in explaining the observed association between higher trading thresholds and lower losses is the trade-off between the number of trades and the returns per trade by Vidyamurthy (2004). The highest average number of openings (4.25) and closings (3.35) is, as expected and shown in Table A.2, recorded by the portfolio based on 1.0 historical standard deviation. This portfolio records the most round-trip trades, which ensure positive returns by definition. Consequently, returns (1.47%) are the highest of all portfolios and economically significant. However, this also implies the highest trading costs, as these increase by 29 bps per one-way trade. Simultaneously, returns are capped by 1.0 standard deviation per completed pairs trade. This indicates that the portfolio presumably suffered from convergence with a loss (Rad et al., 2016). This resulted in a loss of 0.99%.

The worst performing portfolio, based on 1.5 standard deviations, suffered from a 24.71% decrease in the number of openings compared to the portfolio with 1.0 standard deviation. The share of non-convergent pairs also increased from 5% to 15%. Next, plausibly this portfolio also suffered from convergence with a loss since returns are capped to the relatively low 1.5 historical standard deviations per opening. These factors most likely affected excess returns and profits.

Both excess returns and net profits appear to increase with a rise in triggers from 2.0 to 2.5 and onwards to 3.0 historical standard deviations in Table 3. As expected, the average number of openings decreases with higher triggers and is equal to 2.80, 2.30, and 2.15 for the portfolios based on 2.0, 2.5 and 3.0 historical standard deviations, respectively. Therefore, the excess returns and profits per opening must increase with higher triggers in the sample.

Three reasons for this can be identified. First, the probability of (partial) convergence with a loss decreases with higher triggers (Huck & Afawubo, 2015). If the number of historical standard deviations increases, the maximum returns per opening increases. Therefore, returns are more likely to exceed transaction costs and recorded returns and profits per opening increase. Next, losses from non-convergence, after the first or a later opening, are lower for pairs with higher triggers. Since the deviation from the equilibrium spread required to open positions is larger, the difference with the deviation from the equilibrium spread on the last trading day is smaller. This limits the losses per opening. Finally, I suspect that higher trading thresholds took more advantage of the 2020 bear market. Large shocks and mispricing for brief periods appear more likely in this market. Pairs trading with higher triggers takes more advantage of large spreads, while still closing rapidly due to the brief nature of shocks. This could result in larger returns and profits per trade.

4.3 Loss limitation condition

The performance of the same portfolios, now with stop-loss included, is displayed in Table 4. Trading statistics can be found in Table A.3 in the Appendix. As shown in Panel A of Table 4, the pooled portfolio reports statistically insignificant excess returns of 0.58% and statistically significant losses of 0.62% on a 10% level. Panels B, D, and E show statistically insignificant positive excess returns and losses. Panel C shows statistically insignificant negative excess returns of 0.08% and a loss of 1.35%, statistically significant on a 5% level. Panel F is the first displaying positive excess returns (1.16%) and profits (0.24%), however, these are not statistically significant.

After implementing the stop-loss, the pooled portfolio records losses (0.62%) which are less economically impactful than before (0.81%). Compare Panels A of Tables 3 and 4 for this result. Next, the profitability of all thresholds improved in the sample, compare Panels B, C, D, E, and F between Tables 3 and 4. Economically insignificant changes are observed for 1.0 and 1.5 standard deviations, as losses change from 0.99% to 0.93% and from 1.42% to 1.35%, respectively. Mediocre change is observed for 2.0 and 2.5 standard deviations, where losses drop from 0.90% to 0.76% and from 0.51% to 0.28%, respectively. Economically significant change is observed for 3.0 standard deviations, where the loss of 0.22% turns into a profit of 0.24%. These comparisons are tested but all appear statistically insignificant, refer to Table A.4 in the Appendix for details.

The loss limitation condition proves to have an economically significant effect on the recorded performance of the cointegration-based pairs trading strategy in this sample. However, this effect is not statistically significant in any of the cases. Therefore, I do not find enough evidence to support hypothesis 3. Next, these findings provide no evidence to support either hypothesis 1 or hypothesis 2, when the loss limitation is applied. Huck and Afawubo (2015) expected results to improve after implementing the stop-loss condition, however, this cannot be proven statistically in this paper. Nonetheless, the economically significant effect of the stop-loss in this sample corresponds with their expectations.

Since, to my knowledge, none have directly compared performance with and without stop-loss included before, methodological or context related comparison with other studies is not possible. Hence, I will attempt to explain the effect of the stop-loss in this sample by applying elements out of prior research and my own reasoning to the observations.

The impact of the stop-loss on net profits in this study can be explained by a trade-off between loss prevention and premature closings (Caldeira & Moura, 2013). The main aim of the stop-loss was preventing extreme losses, which could dominate capped profits. Here the stop-loss proves successful. The distributions of net profits for the pooled portfolio without and with stop-loss are displayed in Figures A.1 and A.2 in the Appendix, respectively. These figures show major loss prevention by the stop-loss. The largest loss before applying the stop-loss was 11.50%, refer to Panel A of Table 3. Now, the largest loss is 4.89%, shown in Panel A of Table 4.

However, Figures A.1 and A.2 also show another effect of the stop-loss. A large part of the high profits is cut down due to premature closing of pairs. Trading is blocked due to a crossing with the stop-loss threshold, while these pairs would still converge and provide one or multiple profitable round-trip trades. The trading statistics also show this effect, refer to Panel A of Tables A.1 and A.2. The average number of openings for the pooled portfolio decreased from 2.94 to 2.06 with loss limitation and the average time open dropped down from 73.19% to 26.46%.

Table 4
Performance Pairs Trading with Loss Limitation

This table shows the performance of different pairs trading portfolios based on various trading rules, all with the loss limitation condition included in the trading rules. Panel A shows the pooled performance of trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, Panel A has 100 observations. Panels B, C, D, E, and F display the performance of the 20 cointegrated pairs with trading rules based on 1.0, 1.5, 2.0, 2.5, or 3.0 historical standard deviations, respectively. These panels have 20 observations. Each panel exhibits the corresponding excess return, which is equal to the returns on committed capital because I assume self-financing long-short portfolios on average. Note that these returns are not annualized, and, therefore, are recorded over the trading period. All panels also report the net profit over the trading period, which is reached after subtracting the transaction costs from the excess returns. Assumed transaction costs are 29 bps for the one-way opening of positions and are doubled when the pairs trade is closed. If a pair records multiple complete pairs trades, total transaction costs are multiplied by this number. Statistical significance of excess returns and net profits is denoted by ***, **, and * for 1%, 5%, and 10% levels, respectively.

Panel A: Performance pooled portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.58	-0.53	3.95	-4.31	15.67	100
Net profits	-0.62*	-1.60	3.25	-4.89	11.84	100
Panel B: Performance n^*=1.0 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.70	-0.85	4.80	-4.31	15.67	20
Net profits	-0.93	-2.23	3.64	-4.89	11.03	20
Panel C: Performance n^*=1.5 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	-0.08	-1.12	3.53	-3.73	10.68	20
Net profits	-1.35**	-1.97	2.87	-4.55	7.78	20
Panel D: Performance n^*=2.0 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.40	-0.55	3.49	-3.36	10.68	20
Net profits	-0.76	-1.42	2.84	-3.94	7.78	20
Panel E: Performance n^*=2.5 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	0.71	-0.01	3.84	-3.33	13.26	20
Net profits	-0.28	-0.77	3.26	-3.91	10.36	20
Panel F: Performance n^*=3.0 portfolio						
	Mean (%)	Median (%)	St. Dev. (%)	Min (%)	Max (%)	N
Excess returns	1.16	0.13	4.24	-2.79	14.74	20
Net profits	0.24	-0.63	3.66	-3.37	11.84	20

Since profitability of all portfolios improved with the stop-loss in this sample, prevention of extreme losses slightly outweighs premature closings with the stop-loss threshold at 6.0

historical standard deviations. Consequently, a trade-off appears: higher stop-loss thresholds prevent more premature closings of pairs but allow for larger losses to occur.

Besides, the prevention of extreme losses and cut down of large profits leads to a lower standard deviation of net profits, which decreased for the pooled portfolio from 5.51% in Panel A of Table 3 to 3.25% in Panel A of Table 4. This implies a less volatile strategy, which makes the implementation of a loss limitation possibly more attractive.

Furthermore, an important note surrounds the assumption of returns on committed capital. If trading of a pair is blocked after exceeding the stop-loss threshold, no returns or profits can be made from this point. However, I assume that capital is still committed to this pair. If investors move this capital to alternative pairs from this point, the assessed strategy could be more profitable.

Another observation is that larger trading thresholds are associated with more economically significant effects of the stop-loss. Note that the same pairs are traded in the portfolios based on different trading thresholds. Therefore, once a pair crosses the stop-loss threshold, this pair is closed at the same time in each of the portfolios and only the performance after this is affected. Note that the stop-loss closing implies a loss by definition but results in better average performance for all portfolios and this impact is larger when the pairs are traded with higher trading thresholds. No obvious reason for the association between larger trading thresholds and higher impact of the stop-loss appears. A partial explanation could be that the inevitable loss following from the stop-loss closing is smaller if the opening thresholds are higher. This loss is capped by the stop-loss threshold minus the opening threshold. The opening threshold varies but the stop-loss threshold is always the same. Next, this association could also partly follow from the nature of stock prices in the 2020 bear market in this sample, and hence no definitive relationship can be defined here.

5. Conclusion

This paper assesses a certain form of statistical arbitrage: pairs trading. Using long-short positions, divergence of pairs in the short term is exploited since convergence to a long-term equilibrium is expected. To find and trade pairs which possess an equilibrium relationship, the literature proposes various approaches. This paper employs the cointegration approach to detect pairs which are suited for trading. Prior literature concurs on a decay in pairs trading profits since the 1980s and 1990s. However, the 2000-2002 and 2007-2009 bear markets prove exceptions with notable profitability.

This paper endeavors to assess if cointegration based pairs trading strategies could achieve profits in the U.S. during the 2020 bear market. Evaluation of pairs trading in this market adds a new reference point to the literature on the evolution of performance over time. Next, since five trading thresholds are compared in this paper, a broader evaluation of the optimal threshold is contributed to the literature. Furthermore, I assess the impact of a loss limitation condition on performance, which adds a new method to the scarce work on extreme losses prevention.

Using daily stock data for all S&P 500 companies retrieved from CRSP through WRDS, I identified pairs over 2019 and traded them over the first half of 2020. Pairs were preselected by ranking based on lowest SSD and tested for cointegration until 20 pairs were determined. Next, trading occurred when the deviation from the equilibrium level spread diverged a certain multiple of its own historical standard deviation. This method is repeated with a stop-loss threshold of 6.0 historical standard deviations to assess the impact of loss limitation.

Without the stop-loss threshold implemented, this paper finds no statistically significant profits from cointegration based pairs trading strategies. In fact, all thresholds record statistically insignificant, but economically significant, losses after transaction costs. This contrasts with the literature since Do and Faff (2010, 2012), Huck and Afawubo (2015), Huck (2015), and Rad et al. (2016) all found profits in previous bear markets. Potential explanations could be increasing market efficiency, competition amongst hedge funds, fading cointegration relationships due to the sudden impact of COVID-19, the method to identify cointegration, and convergence with a loss.

The recorded performances in this paper show an association between higher trading thresholds and losses of lower economic significance. Since no statistical significance is found, this relation remains unproven and contrasts with the literature by Huck and Afawubo (2015) and Huck (2015). The observed association in this study can be explained by the trade-off between the number of trades and the returns per trade. Related are a lower probability of convergence with a loss, smaller losses from non-nonconvergence, and suspected capitalization on the nature of the 2020 bear market when trading thresholds are higher.

After implementing the stop-loss threshold, the performance of all trading thresholds improves. However, none of these effects proves to be statistically significant. Consequently, I cannot statistically support that the stop-loss improves profitability. This finding contrasts with the expectations formulated by Huck and Afawubo (2015). The observed economic impact in this sample can be explained by the trade-off between loss prevention and premature closings of pairs. The association between higher trading thresholds and more economically significant impacts of the stop-loss remains largely unexplained. A partial explanation could be smaller losses as a result of higher opening triggers. Furthermore, the implementation of a loss limitation condition neither results in statistically significant profitable portfolios nor supports better performance of higher trading thresholds. Noteworthy, due to lower volatility and the possibility to divest capital after trading is blocked, the loss limitation condition has untapped potential. More on this follows later.

All these findings serve to answer the previously stated main research question as follows. In this study, cointegration based pairs trading strategies do not achieve profits in the U.S during the 2020 bear market. Trading threshold optimization does not offer statistically significant profits. This holds both before and after implementation of the loss limitation condition.

The findings of this study bring up practical implications. Since cointegration based pairs trading does not appear profitable in this study, investors should be cautioned to allocate resources to this strategy. According to a large share of the literature, previous bear markets bolstered pairs trading performance compared to more quiet markets. Therefore, losses in the 2020 bear market do not bode well for performance in more calm markets in the upcoming years. However, pairs trading is not prospectless. This study found one, although not statistically significant, profitable trading strategy (3 historical standard deviations, with stop-loss included). Several traces point towards better performances after implementing the stop-loss threshold and as indicated, this study did not focus on optimizing loss limitation. Furthermore, the stop-loss decreases volatility in this study and can identify pairs where capital should be divested from. Hence, this paper could warn investors but should not move them to eliminating pairs trading strategies from their portfolios.

Another implication is for academia and surrounds the discussion on the aforementioned efficient market hypothesis. All strategies evaluated in this paper, except one (1.5 historical standard deviations, with stop-loss included), report positive excess returns over the approximately

6-month trading period. These findings are in contrast with the efficient market hypothesis, which predicts that trading stocks based on historical movements rules out excess returns. Therefore, results from this study imply a rejection of the efficient market hypothesis.

It is vital to acknowledge the limitations of this study. First, a well-known problem in time-series analysis is data snooping. Testing multiple portfolios over the same period is inherent to comparing different trading thresholds in this paper. Great portfolio performance could be attributed to successful pairs selection and trading rules but might also follow from coincidence. Therefore, results in this paper should always be interpreted with caution.

Next, the portfolios in this paper are relatively small. Although this serves computational ease, since cointegration testing is time-consuming, it also brings downsides. First, larger portfolios would strengthen the assumption of self-financing long-short positions on average. Second, small portfolios are less reliable and thus more likely to suffer from high standard errors and hence the statistical significance of returns is more demanding to prove. However, too large portfolios are impractical and could lead to unjust significance.

Moreover, treatment as either dependent or independent variable in the OLS used for cointegration testing and identifying trading proportions is assigned randomly. Since the formation period is finite and stocks within a pair are never perfectly correlated, treating stocks as dependent or independent variable matters. Hence, random treatment could affect the trading proportions and identification of cointegration in this study. This is a weakness of single equation techniques like the method of Engle and Granger (1987) and could be overcome by using the approach of Johansen (1988). However, random assignment serves the assumption of self-financing long-short positions on average since it likely results in a more uniform distribution of trading proportions in portfolios.

Next, pairs are ranked based on lowest SSD before cointegration testing, leading to the selection of pairs with a low equilibrium spread and a low variance of this spread. The first consequence is low historical standard deviations of the error from the equilibrium spread and thus small trading thresholds. Hence, pairs in this study are prone to converging with a loss. Second, uncertain is whether the more robust statistical foundation for long-term co-movements of the cointegration approach outweighs its more costly estimation, compared to the distance approach. Since this paper tests for 20 cointegrated pairs in order of SSD rank, the group of chosen pairs is like those picked when the distance approach would be applied. Hence, performance is expected to be alike and the extra costs of cointegration testing might offset higher profits due to robustness.

Furthermore, the estimation of transaction costs is imperfect in this study. This is of concern since these costs hurt excess returns considerably, as stated in chapter 4. The employed estimation by Do and Faff (2012) might be outdated and, hence, used institutional commissions and market impacts are possibly not representative anymore. Next, institutional commissions are based on NYSE data in their study, while this paper also considers pairs traded on other exchanges. Short selling constraints are also not accounted for in this paper while these might have an impact.

Finally, I will point out avenues for future research. Of course, pairs trading could be evaluated in different periods or countries. This would lead to insights on the development of pairs trading performance over time and differences in results between exchanges, maybe depending on the efficiency of exchanges. However, I would like to direct attention to three unpaved avenues.

First, the influence of varying trading thresholds on pairs trading performance should be investigated thoroughly. Existing literature often employs two historical standard deviations as trigger in cointegration and distance approach-based methods. This threshold is chosen arbitrarily, and other levels are barely researched, an exception is three historical standard deviations by Huck and Afawubo (2015) and Huck (2015). This paper tries to broaden the scope by evaluating five levels, yet also these do not reveal the optimal threshold since they jump by steps of half a standard deviation. An approach which attempts to approximate a continuous range of triggers could contribute to optimal trading threshold determination. Alternatively, a chance for future research is looking further than just historical standard deviations as basis for trading thresholds.

A second gap for future research is in the limitation of extreme losses. Only a handful of researchers have evaluated possible stop-loss conditions. This paper finds economically significant impact of a loss limitation rule, while it was chosen arbitrarily. Next, as mentioned, stop-loss thresholds lower volatility of portfolios and identify pairs from which an investor should divest. These indicators show the potential of optimizing extreme loss limitation in pairs trading. Whilst simple stop-loss thresholds, for example based on standard deviations as in this paper, facilitate practical implementation, more complex rules could be more impactful.

Third and lastly, more attention could be shifted to the estimation of transaction costs in pairs trading strategies. As mentioned, current popular estimates are outdated and imperfect. Although a perfect estimate will never exist, increased attention could advance the search for reliable net profits in pairs trading literature.

6. References

Alexander, C. (1999). Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *357*(1758), 2039-2058.

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

Brooks, C. (2019). *Introductory Econometrics for Finance (4th edition)*. Cambridge: Cambridge University Press.

Caldeira, J. & Moura, G. V. (2013). Selection of a portfolio of pairs based on cointegration: A statistical arbitrage strategy. *Available at SSRN 2196391*.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative finance*, *1*(2), 223.

De Bondt, W. F. & Thaler, R. (1985). Does the stock market overreact?. *The Journal of finance*, *40*(3), 793-805.

Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, *74*(366a), 427-431.

Dickey, D. A. & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: journal of the Econometric Society*, 1057-1072.

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

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

Elliott, R. J., Van der Hoek, J. & Malcolm, W. P. (2005). Pairs trading. *Quantitative Finance*, *5*(3), 271-276.

Engle, R. F. & Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383-417.

Gatev, E., Goetzmann, W. N. & Rouwenhorst, K. G. (1999). Pairs Trading: Performance of a Relative Value Arbitrage Rule.

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

Grossman, S. J. & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), 393-408.

Huck, N. (2009). Pairs selection and outranking: An application to the S&P 100 index. *European Journal of Operational Research*, *196*(2), 819-825.

Huck, N. (2010). Pairs trading and outranking: The multi-step-ahead forecasting case. *European Journal of Operational Research*, *207*(3), 1702-1716.

Huck, N. (2015). Pairs trading: does volatility timing matter?. *Applied economics*, *47*(57), 6239 -6256.

Huck, N. & Afawubo, K. (2015). Pairs trading and selection methods: is cointegration superior?. *Applied Economics*, *47*(6), 599-613.

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

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, *12*(2-3), 231-254.

Jurek, J. W. & Yang, H. (2007). Dynamic portfolio selection in arbitrage. In *EFA 2006 Meetings Paper*.

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

Lakonishok, J., Shleifer, A. & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, *49*(5), 1541-1578.

Lamont, O. A. & Thaler, R. H. (2003). Anomalies: The law of one price in financial markets. *Journal of Economic Perspectives*, *17*(4), 191-202.

Liew, R. Q. & Wu, Y. (2013). Pairs trading: A copula approach. *Journal of Derivatives & Hedge Funds*, *19*(1), 12-30.

Lin, Y. X., McCrae, M. & Gulati, C. (2006). Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Advances in Decision Sciences*, *2006*.

MacKinnon, J. G. (1990). *Critical values for cointegration tests* (No. 1227). Queen's Economics Department Working Paper.

MacKinnon, J. G. (2010). *Critical values for cointegration tests* (No. 1227). Queen's Economics Department Working Paper.

Nath, P. (2003). High frequency pairs trading with US treasury securities: Risks and rewards for hedge funds. *Available at SSRN 565441*.

Poterba, J. M. & Summers, L. H. (1988). Mean reversion in stock prices: Evidence and implications. *Journal of financial economics*, *22*(1), 27-59.

Rad, H., Low, R. K. Y. & Faff, R. (2016). The profitability of pairs trading strategies: distance, cointegration and copula methods. *Quantitative Finance*, *16*(10), 1541-1558.

Shleifer, A. & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of finance*, *52*(1), 35 -55.

Vidyamurthy, G. (2004). *Pairs Trading: quantitative methods and analysis* (Vol. 217). John Wiley & Sons.

Xie, W. & Wu, Y. (2013). Copula-based pairs trading strategy. *Asian Finance Association (AsFA)*, *10*.

7. Appendix

Table A.1
Overview of Cointegrated Pairs

This table shows an overview of the 20 cointegrated pairs and their stocks. Below, the SSD rank of the cointegrated pairs is shown in order of lowest SSD of all possible pairs. Next, for all stocks in the cointegrated pairs, the corresponding tickers, company names, SIC codes and returns over the 1-year formation period are exhibited.

SSD ranking	Ticker	Company name	SIC code	Returns (%)
1	GOOG	Alphabet Inc (Class C)	7375	27.84
	GOOGL	Alphabet Inc (Class A)	7375	26.99
2	NWS	News Corp (Class B)	9999	24.98
	NWSA	News Corp (Class A)	9999	24.14
4	DISCA	Discovery Inc (Series A)	4841	26.75
	DISCK	Discovery Inc (Series C)	4841	27.95
8	LNT	Alliant Energy Corp	4931	32.98
	FE	FirstEnergy Corp	4911	32.35
11	FE	FirstEnergy Corp	4911	32.35
	ES	Eversource Energy	4911	33.99
13	NEE	Nextera Energy Inc	4911	42.59
	ETR	Entergy Corp	4911	42.62
14	CBOE	Cboe Global Markets Inc	6231	25.14
	ICE	Intercontinental Exchange Inc	6231	22.86
15	PNC	PNC Financial Services Group Inc	6021	34.36
	JPM	JPMorgan Chase & Co	6021	40.37
20	KO	The Coca-Cola Company	2086	17.94
	EVRG	Evergy Inc	4931	16.48
22	PPG	PPG Industries	2851	31.27
	PNC	PNC Financial Services Group Inc	6021	34.36
23	CMI	Cummins Inc	3519	34.17
	PCAR	Paccar Inc	3711	36.71
24	PG	Procter & Gamble	2841	36.83
	NEE	Nextera Energy Inc	4911	42.59
25	FITB	Fifth Third Bancorp	6711	27.02
	CFG	Citizens Financial Group Inc	6022	33.10
29	SO	Southern Co	4911	45.70
	MAA	Mid-America Apartment Communities Inc	6798	42.32
35	CB	Chubb Ltd	6331	22.29
	KMB	Kimberly-Clark Corp	2844	23.04
40	CCI	Crown Castle International Corp	4812	34.17
	XEL	Xcel Energy	4911	31.67
41	RSG	Republic Services Inc	4953	25.30
	KMB	Kimberly-Clark Corp	2844	23.04
42	AJG	Arthur J. Gallagher & Co	6411	32.47
	NDAQ	NASDAQ Inc	6200	32.22
45	V	Visa International	7374	41.36
	SRE	Sempra Energy	4612	41.89
49	KMB	Kimberly-Clark Corp	2844	23.04
	PEP	PepsiCo Inc	2086	25.06

Table A.2
Trading Statistics of Pairs without Loss Limitation

This table displays the trading statistics of different pairs trading portfolios based on various trading rules, all without loss limitation condition included. Panel A shows the pooled trading statistics of pairs trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, Panel A has 100 observations. Panels B, C, D, E, and F display the trading statistics of the 20 cointegrated pairs with trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations, respectively. These panels have 20 observations. Each panel displays the shares of non-traded pairs, non-converged pairs, single round trip pairs, and multiple opening pairs. Next, all panels show descriptive statistics on the number of openings and closings per pair throughout the trading period. Note that the number of closings does not contain the forced close at the end of the trading period. Finally, each panel exhibits descriptive statistics on the share of time positions per pair are open during the trading period. The trading period includes 125 trading days.

Panel A: Performance pooled portfolio

	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	100
Non-convergent pairs (%)	18.00	-	-	-	-	100
Single round trip pairs (%)	5.00	-	-	-	-	100
Multiple opening pairs (%)	77.00	-	-	-	-	100
Number of openings	2.94	2.00	1.87	1.00	10.00	100
Number of closings	2.13	1.00	1.93	0.00	9.00	100
Time open (%)	73.19	77.60	20.35	17.60	99.20	100

Panel B: Performance n^*=1.0 portfolio

	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non-convergent pairs (%)	5.00	-	-	-	-	20
Single round trip pairs (%)	0.00	-	-	-	-	20
Multiple opening pairs (%)	95.00	-	-	-	-	20
Number of openings	4.25	4.00	2.49	1.00	10.00	20
Number of closings	3.35	3.00	2.58	0.00	9.00	20
Time open (%)	87.12	91.20	10.75	64.80	99.20	20

Panel C: Performance n^*=1.5 portfolio

	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non-convergent pairs (%)	15.00	-	-	-	-	20
Single round trip pairs (%)	0.00	-	-	-	-	20
Multiple opening pairs (%)	85.00	-	-	-	-	20
Number of openings	3.20	3.00	1.85	1.00	8.00	20
Number of closings	2.35	2.00	1.93	0.00	7.00	20
Time open (%)	79.36	83.20	14.31	54.40	99.20	20

Panel D: Performance n^*=2.0 portfolio

	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non-convergent pairs (%)	15.00	-	-	-	-	20
Single round trip pairs (%)	5.00	-	-	-	-	20
Multiple opening pairs (%)	80.00	-	-	-	-	20
Number of openings	2.80	2.00	1.51	1.00	6.00	20
Number of closings	2.00	1.00	1.62	0.00	5.00	20
Time open (%)	73.72	78.80	18.25	37.60	98.40	20

Continuation Table A.2

Panel E: Performance n^*=2.5 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non convergent pairs (%)	25.00	-	-	-	-	20
Single round trip pairs (%)	10.00	-	-	-	-	20
Multiple opening pairs (%)	65.00	-	-	-	-	20
Number of openings	2.30	2.00	1.34	1.00	5.00	20
Number of closings	1.55	1.00	1.43	0.00	4.00	20
Time open (%)	64.96	73.20	23.05	23.20	92.80	20

Panel F: Performance n^*=3.0 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non convergent pairs (%)	30.00	-	-	-	-	20
Single round trip pairs (%)	10.00	-	-	-	-	20
Multiple opening pairs (%)	60.00	-	-	-	-	20
Number of openings	2.15	2.00	1.27	1.00	5.00	20
Number of closings	1.40	1.00	1.35	0.00	4.00	20
Time open (%)	60.80	70.00	22.27	17.60	86.40	20

Table A.3
Trading Statistics of Pairs with Loss Limitation

This table displays the trading statistics of different pairs trading portfolios based on various trading rules, all with the loss limitation condition included in the trading rules. Panel A shows the pooled trading statistics of pairs trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, Panel A has 100 observations. Panels B, C, D, E, and F display the trading statistics of the 20 cointegrated pairs with trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations, respectively. These panels have 20 observations. Each panel displays the shares of non-traded pairs, non-converged pairs, single round trip pairs, and multiple opening pairs. Next, all panels show descriptive statistics on the number of openings and closings per pair throughout the trading period. Note that the number of closings again does not contain the forced close at the end of the trading period. However, the close following from the loss limitation condition is included. Finally, each panel exhibits descriptive statistics on the share of time positions per pair are open during the trading period. The trading period includes 125 trading days.

Panel A: Performance pooled portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	100
Non-convergent pairs (%)	0.00	-	-	-	-	100
Single round trip pairs (%)	50.00	-	-	-	-	100
Multiple opening pairs (%)	50.00	-	-	-	-	100
Number of openings	2.06	1.50	1.68	1.00	10.00	100
Number of closings	1.91	1.00	1.44	1.00	9.00	100
Time open (%)	26.46	23.60	18.53	0.80	84.00	100
Panel B: Performance n^*=1.0 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non-convergent pairs (%)	0.00	-	-	-	-	20
Single round trip pairs (%)	30.00	-	-	-	-	20
Multiple opening pairs (%)	70.00	-	-	-	-	20
Number of openings	2.80	2.00	2.38	1.00	10.00	20
Number of closings	2.65	2.00	2.08	1.00	9.00	20
Time open (%)	34.04	30.40	18.63	8.80	84.00	20
Panel C: Performance n^*=1.5 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non-convergent pairs (%)	0.00	-	-	-	-	20
Single round trip pairs (%)	45.00	-	-	-	-	20
Multiple opening pairs (%)	55.00	-	-	-	-	20
Number of openings	2.20	2.00	1.77	1.00	8.00	20
Number of closings	2.05	1.50	1.54	1.00	7.00	20
Time open (%)	31.92	27.20	17.38	10.40	79.20	20
Panel D: Performance n^*=2.0 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non-convergent pairs (%)	0.00	-	-	-	-	20
Single round trip pairs (%)	45.00	-	-	-	-	20
Multiple opening pairs (%)	55.00	-	-	-	-	20
Number of openings	2.00	2.00	1.41	1.00	6.00	20
Number of closings	1.85	1.50	1.18	1.00	5.00	20
Time open (%)	27.72	24.40	17.66	9.60	79.20	20

Continuation Table A.3

Panel E: Performance n^*=2.5 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non convergent pairs (%)	0.00	-	-	-	-	20
Single round trip pairs (%)	60.00	-	-	-	-	20
Multiple opening pairs (%)	40.00	-	-	-	-	20
Number of openings	1.70	1.00	1.22	1.00	5.00	20
Number of closings	1.55	1.00	0.94	1.00	4.00	20
Time open (%)	21.16	19.20	17.48	0.80	76.00	20
Panel F: Performance n^*=3.0 portfolio						
	Mean	Median	St. Dev.	Min	Max	N
Non-traded pairs (%)	0.00	-	-	-	-	20
Non convergent pairs (%)	0.00	-	-	-	-	20
Single round trip pairs (%)	70.00	-	-	-	-	20
Multiple opening pairs (%)	30.00	-	-	-	-	20
Number of openings	1.60	1.00	1.23	1.00	5.00	20
Number of closings	1.45	1.00	0.94	1.00	4.00	20
Time open (%)	17.48	12.40	17.72	0.80	74.40	20

Table A.4
Difference in Net Profits from Implementing Loss Limitation Condition

This table shows the difference in net profits between the same portfolios with and without loss limitation condition. The pooled portfolio is based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, this portfolio has 100 observations. The portfolios based on 1.0, 1.5, 2.0, 2.5, or 3.0 historical standard deviations have 20 observations each. The net profits without loss limitation for all portfolios are the same as reported in Table 3, refer to section 4.1. The net profits with loss limitation for all portfolios are the same as reported in Table 4, refer to section 4.3. The difference is calculated by subtracting the net profits without loss limitation from the net profits with loss limitation. A two-sample t-test with the assumption of unequal variances is conducted, after confirming unequal variances for all comparisons using F-tests. Statistical significance of net profits and the differences is denoted by ***, **, and * for 1%, 5%, and 10% levels, respectively.

	Without loss limitation	With loss limitation	Difference	N
Pooled	-0.81	-0.62*	0.19	100
n^*=1.0	-0.99	-0.93	0.06	20
n^*=1.5	-1.42	-1.35**	0.07	20
n^*=2.0	-0.90	-0.76	0.14	20
n^*=2.5	-0.51	-0.28	0.23	20
n^*=3.0	-0.22	0.24	0.46	20

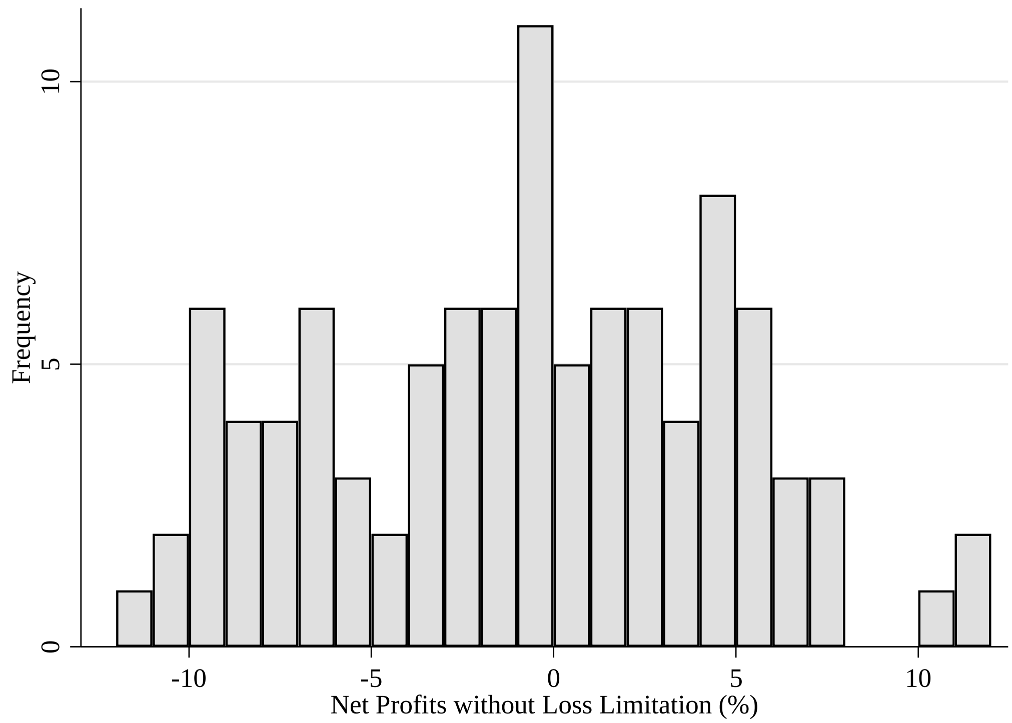

Figure A.1. Histogram of Net Profits without Loss Limitation.

This figure shows a histogram of the net profits resulting from the pooled portfolio without the loss limitation condition implemented. The pooled portfolio consists of trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, this portfolio has 100 observations. The bins in the histogram have a width of 1 percentage point. The lower limit of the first bin is set to -12%. The upper limit of the last bin is set to 12%. The frequency denotes how many pairs fall within a certain bin.

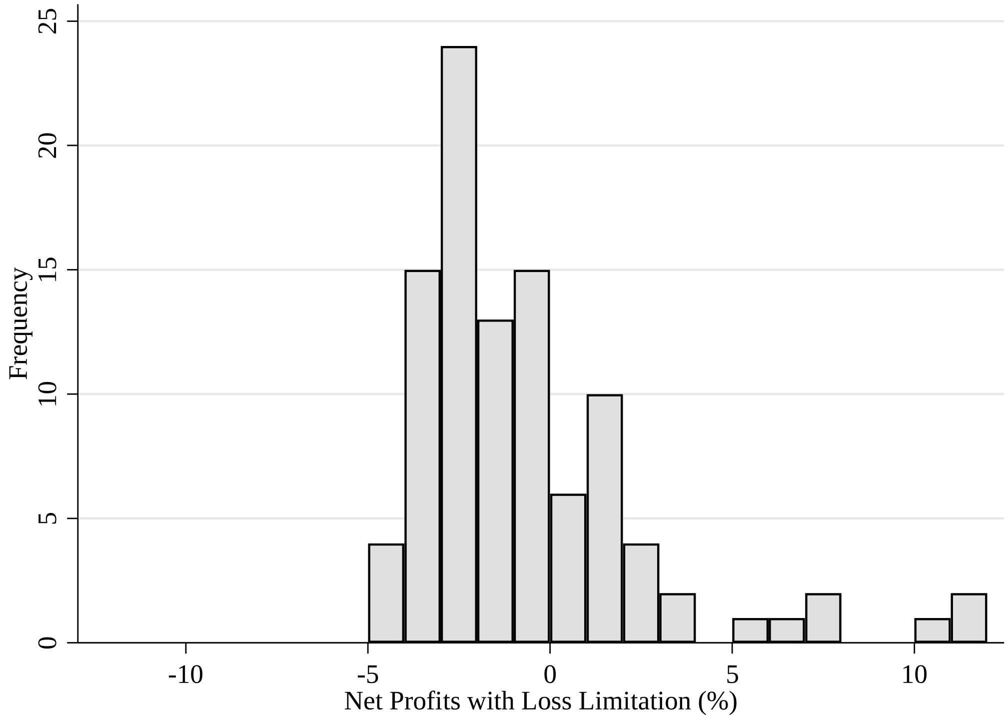

Figure A.2. Histogram of Net Profits with Loss Limitation.

This figure shows a histogram of the net profits resulting from the pooled portfolio with the loss limitation condition implemented. The pooled portfolio consists of trading rules based on 1.0, 1.5, 2.0, 2.5, and 3.0 historical standard deviations for each of the 20 cointegrated pairs. Hence, this portfolio has 100 observations. The bins in the histogram have a width of 1 percentage point. The lower limit of the first bin is set to -12%. The upper limit of the last bin is set to 12%. The frequency denotes how many pairs fall within a certain bin.