Universidad
Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Optimal Portfolio Strategies of Cointegrated Assets

Autor:

Tao Tang

Director/es:

Isabel Figuerola-Ferretti Garrigues

Pedro José Serrano Jimenez

DEPARTAMENTO/INSTITUTO

Economía de la Empresa

Getafe, Mayo de 2017

Universidad Carlos III de Madrid
www.uc3m.es

(a entregar en la Oficina de Posgrado, una vez nombrado el Tribunal evaluador , para preparar el documento para la defensa de la tesis)

TESIS DOCTORAL

Optimal Portfolio Strategies of Cointegrated Assets

Autor: ***Tao Tang***

Director/es: Isabel Figuerola-Ferretti Garrigues

Pedro José Serrano Jimenez

Firma del Tribunal Calificador:

Firma

Presidente: (Nombre y apellidos)

Vocal: (Nombre y apellidos)

Secretario: (Nombre y apellidos)

Calificación:

Leganés/Getafe, de de

Contents

List of Figures

List of Tables

Acknowledgements

Looking back to the journey of my Ph.D. study, I am lucky to have many kind people around me and grateful to their help, inspiration, patience and encouragement.

First of all, my deepest thanks go to my advisor, Isabel Figuerola-Ferretti, for her excellent guidance, continuous support, and insightful comments that have certainly raised the standard of my work. She has been the leader during the whole journey, providing me with plenty of courage to go through difficulties and making the thesis fruitful. She was involved in the three chapters of my thesis. My deepest gratitude also goes to my co-advisor, Pedro Serrano, for his caring, countless help, and useful advice. He has taken this role since the fifth year of my Ph.D., sharing his academic research experience for helping me to build confidence. The third chapter was co-authored with him. I am also grateful to my co-author, Ioannis Paraskevopoulos, for his guidance and efforts on the first and the second chapters. Thank you for his deep involvement in my work whenever I need comments and discussions.

I extend my special thanks to the faculty in the Business Department at UC3M, Josep A. Tribó and Mikel Tapia, for their support during the job market process. I would like to express my gratitude to my colleagues, Jiali Wu, Yannson Wong, and Yanlei Zhang who has helped me a lot at different times over the past few years. They made my stay in Madrid funny and memorable. Thanks to my friends, Daisy Lam, Percy Wong, Yuebin Chen and Xueming Feng, for accompanying me on the road pursuing my academic career.

I would like to thank my family for their endless love and dedication during my studies. They supported me in all my pursuits without asking for returns. It is their strong support and encouragement that made me dedicated to my work and made this thesis possible.

Abstract

Statistical arbitrage, as a quantitative method of speculation, has been increasingly prevalent along with the evolution of computational finance. One of the most popular statistical arbitrage strategies is called pairs trading, which is widely used by hedge funds and investment banks since the mid-1980s. Pairs trading strategy exploits price spread between paired assets by taking long-short positions. If price spread is temporary according to past price information, a trading opportunity arises and profits can be made from price correction process. To capture these opportunities, we focus on assets sharing cointegration relations. This long-term relationship implies that paired assets are exposed to common fundamentals, and hence it guarantees price convergence to the equilibrium level. Therefore, this thesis applies cointegration technique to capture short-term market anomalies and exploits these inefficiencies using pairs trading in order to build optimal portfolio strategies.

The thesis consists of three chapters. The first chapter presents an equilibrium framework based on equity commonality explicitly adapted to describe the dynamics of pairs trading. Our methodology, built on the price discovery model of Figuerola-Ferretti and Gonzalo (Journal of Econometrics 2010) exploits price leadership for portfolio replication purposes and shows how pairs trading profitability is linked to the speed of equilibrium reversion. A persistence-dependent trading trigger is introduced to impose higher thresholds on pairs with slower mean reversion. Our model demonstrates that equilibrium price convergence guarantees positive abnormal profitability. Applied to STOXX Europe 600 traded equities our strategy delivers Sharpe ratios that outperform benchmark rules used in the literature. Portfolio performance is enhanced after firm fundamental factor restrictions are imposed.

The second chapter proposes a VECM representation for cointegrated assets in the continuous time framework. This model implies a simple method to check for cointegration based on the speed of equilibrium reversion. A pair of cointegrated assets is then identified to derive a dynamically optimal pairs trading portfolio with a risk-free bond. This involves maximizing the portfolio value at terminal time without the requirement of a functional form for investor´s preferences. To this end, we connect the derived optimal portfolio with European-type spread options and in consequence the optimal investment policies can be modeled using the spread option's resulting delta hedging strategies. Our framework is tested empirically using pairs identified from the Dow Jones Industrial Average. This analysis requires maximum likelihood estimates on continuous VECM parameters, compared to the benchmark Johansen methodology. We find that the proposed optimal strategy delivers consistent profitability in terms of Sharpe ratio and cumulative returns. This supports the usefulness of introducing spread option's deltas as the optimal investment policies for pairs portfolio construction. In addition, our model-implied selection algorithm outperforms the Johansen (1991) methodology commonly applied in the previous literature.

Finally, the third chapter examines the performance of pair trading portfolios when sorted by the level of cointegration of their constituents. The supercointegrated portfolio, that is formed by pairs at 1% confidence level of cointegration tests, exhibits a superior out-of-sample performance than simple buy-and-hold and passive investments in terms of Sharpe ratio. We find that the degree of performance of pairs strategy is positively related to the level of cointegration among pairs. These evidence are also documented in an international context, from the analysis on the European stock market. The time-varying risk of the pairs strategy is linked to aggregate market volatility. A positive risk-return relationship of the strategy is also found.

Chapter 1

Pairs Trading and Spread Persistence in the European Stock Market

1.1 Introduction

Short-term price discrepancies are common across assets that are imperfectly integrated. Pairs trading strategies are designed to earn profits from relative mispricing of closely related assets. This paper extends the results from an equilibrium demand and supply framework to show how pairs trading can be used to exploit the mean reversion property of cointegration errors. Pairs trading belongs to the family of convergence trade strategies. It relies on a well-known trading rule for cointegrated price series based on simultaneous long-short positions that are closed when prices revert to long run equilibrium. When an investor opens a position he shorts the out-performer and longs the under-performer, until the mispricing is eliminated. We extend the Figuerola-Ferretti and Gonzalo (2010) (FFG thereafter) demand and supply equilibrium framework to describe price dynamics in two distinct but cointegrated assets and show how market participants exploit temporary mispricings performing pairs trading strategies. The setup requires a finite elasticity of arbitrage services and equilibrium error persistence. It evolves around the speed by which arbitrageurs restore equilibrium allowing measurement of price discovery for portfolio replication purposes and arbitrage profit determination. A market is regarded as dominant in this framework if it concentrates a higher number of participants. Cointegration therefore guarantees equilibrium price convergence that is represented in terms of a stationary error correction term. The trading trigger in this context linked to the degree of persistence of the cointegration error so that higher stationarity requires a lower trading trigger.

This paper is related to Gatev et al. (2006) (GGR thereafter), which examines the performance of pairs trading using daily U.S. stock return data. GGR perform pairs selection using the minimum distance algorithm. They find economically and statistically significant excess returns of around 11% per annum. Following GGR, Andrade et al. (2005), Broussard and Vaihekoski (2012) and Bowen and Hutchinson (2014) apply the algorithm to Asian and European equity markets. A common drawback from these studies is that they essentially apply an ad hoc trading trigger. Vidyamurthy (2004) sheds light to this problem by searching for trading trigger optimality by maximizing a profit function under the assumption of Gaussian errors.

Another strand of the literature models the cointegration spread under dynamic settings. Elliott et al. (2005) and Avellaneda and Lee (2008) consider an Ornstain-Uhlenbeck process to model the cointegration error allowing spread estimation and setting the framework for determination of the threshold value. While Avellaneda and Lee (2008) empirically determine cutoff values based on the process assumed for the cointegration error, Elliott et al. (2005) link the trading trigger to the degree of mean reversion. This paper contributes to the literature by proposing a trading trigger that is determined by the speed of convergence to equilibrium arising from VECM estimates. Our model-based trading rule is therefore related to Elliott et al. (2005) in that the trading trigger is defined as a function of the speed of mean reversion.

Our empirical application is based on an out-of-sample analysis and uses STOXX Europe 600 traded equities whose prices are quoted in the euro currency to identify cointegration relationships with a sample ranging from 2000 to 2017. Common factor and industry restrictions are imposed to illustrate the existence of the long-run equilibrium. This justifies the use of a model with equity shared fundamentals that drive prices to parity. We analyze the profitability of pairs strategies at the portfolio level and compare their performance with benchmark pairs trading methodologies used in the literature. We find that the proposed pairs strategies outperform the seminal strategy of Gatev et al. (2006), as evidenced by significant abnormal returns and higher Sharpe ratios. The documented outperformance is enhanced once we control for common firm fundamentals as well as the industry effect.

The rest of the paper proceeds as follows. In Section 2 we relate the VECM dynamics to the construction of pairs trading strategies. This requires a description of preliminaries and main results of the FFG model applied to two distinct but cointegrated assets. Section 3 presents the data sample and empirical results on cointegration and price discovery. In section 4 we conduct the pairs trading performance analysis with a number of robustness tests. Section 5 provides conclusions.

1.2 The theoretical model

1.2.1 Model set-up

The aim of this section is to introduce pairs trading strategies in a demand and supply equilibrium framework. In this context, arbitrage takes place through pairs trading strategies that exploit mean reversion of the pricing error. The model is built on the presumption that price correction of two cointegrated assets departing from equilibrium relationships depends on the average speed of convergence in each market. This determines the degree of persistence of the cointegration error and becomes an important factor for designing the trading trigger.

In this section we present the joint dynamics between two cointegrated assets within a demand and supply equilibrium framework. Investors either take single asset positions or trade two assets that share common fundamentals simultaneously via the use of pairs trading strategies. Mean reversion of the spread is of critical importance to arbitrageurs who, will exploit short-lived deviations from equilibrium in search of benefits from pairs strategies.

Under imperfect integration, there is a finite elasticity of demand for pairs trading strategies (H),[1] and relative prices may differ between markets for short intervals of time by more than transaction costs. The speed by which such price discrepancies are eliminated depends on the degree of persistence in the error term z_t. The speed of mean reversion is determined by market imperfections that represent impediments to arbitrage such as liquidity or potential funding constraints.[2]

In what follows, we extend the FFG model to describe an equilibrium framework for imperfectly integrated markets. Let y_t and x_t be the price of paired assets in time t, respectively. In order to find the non-arbitrage equilibrium condition, the following set of standard assumptions applies in this section:

1. No limitations on borrowing.
2. No cost other than arbitrage transaction cost.
3. No limitations on short-sale.
4. Arbitrage opportunities that generate a random price difference between paired assets are determined by the stationary process z_t. These arise as a result of market imperfections that impede arbitrage opportunities between markets and lead to a finite elasticity of demand for pairs trading strategies ($H > 0$) and a stationary cointegration error. Higher market imperfections are translated into more persistent errors and lower elasticity of demand for pairs trading. In the limit, when arbitrage opportunities are exploited instantaneously there is an infinite elasticity of demand for pairs trading strategies ($H \to \infty$). This leads to immediate price adjustments to divergences between the two cointegrated markets implying that market frictions are eliminated and temporary mispricings disappear and $z_t = 0$. Under these circumstances there is no cointegration and potential profits from pairs strategies become zero.[3]

[1] This elasticity measures the proportional change in demand for arbitrage strategies in the form of "pairs trading" for a given change in the quantity of arbitrage services.

[2] See for instance Shleifer and Vishny (1997), Xiong (2001), Gromb and Vayanos (2002) and Kondor (2009) for a detailed discussion on the limits to arbitrage.

[3] Market imperfections lead to stationary cointegration errors and finite elasticity of arbitrage services. This differs from other frameworks in the literature that have considered non linearities in the cointegration error to identify the presence of non

5. The equity price series y_t and x_t are $I(1)$, implying that their mean and auto-covariance are different for every realization of t.

By the above assumptions 1-5, long-run equilibrium conditions imply:

$$y_t = \gamma_0 + \gamma_1 x_t + z_t \tag{1.1}$$

where γ_0 is the (constant) cash amount invested (or borrowed) to buy γ_1 units of asset x_t (required to replicate prices of asset y_t). Therefore γ_1 is the hedge ratio as it reflects the size of position that has to be taken in the portfolio with asset x_t to replicate the prices of asset y_t. Equation (1.1), implies that y_t and x_t are cointegrated suggesting (imperfect) market integration. Pairs trading strategies are triggered when temporary mispricings arise from the long-run cointegration relationship. When the spread between y_t and x_t widens by an amount higher than a given threshold, there is a positive profit potential that can be exploited by an arbitrageur who shorts the winner and buys the loser. If the long and short components measure common fundamentals, the prices will restore equilibrium providing positive average (and cumulative) profits. This framework is consistent with Brennan et al. (1993) in that it is market imperfections that give rise to the arbitrage opportunity.

The model developed in Appendix A describes the dynamics of agents who trade single securities and agents who trade two cointegrated assets. Market imperfections are in this context translated into a finite elasticity of demand for pursuing pairs strategies. In this framework the demand schedule for each traded asset has two components: (a) the own asset demand and (b) the speculative demand based on the long-term relationship between the two traded assets. Market clearing conditions are defined by equating aggregate market demand and aggregate supply schedules in a context where new information arrival is reflected in the difference between lagged market clearing prices and current bid prices. The resulting bivariate dynamics between y_t and x_t are represented by the following VECM:

linearities within the basis term arising under the presence of transaction costs, or interaction between sentiment and trading behavior. See McMillan and Ülkü (2009).

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = \frac{H}{d} \begin{pmatrix} -N_x \\ N_y \end{pmatrix} \begin{pmatrix} 1 & -\gamma_1 & -\gamma_0 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \\ 1 \end{pmatrix} + \begin{pmatrix} u_t^y \\ u_t^x \end{pmatrix} \tag{1.2}$$

with

$$d = (H + AN_y)\, N_x + \gamma_1 H N_y \tag{1.3}$$

where there are N_y participants in the market for asset y_t and N_x participants in the market for asset x_t and, as previously specified, the elasticity of demand for pursuing pairs strategies is noted by H. We rewrite the theoretical result in (1.2) as:

$$\Delta P = \begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} z_{t-1} + u_t \tag{1.4}$$

where u_t is a vector white noise with i.i.d shocks.[4]

In order to define the VECM well and ensure that "pairs strategies" can be applied, the following conditions need to be satisfied:

1. If α_1 and a_2 are both statistically significant, they must have opposite signs, as predicted by the theoretical result in Equation (1.2). This implies that, if there is a change in the equilibrium error, so that for instance y_t is greater than its replicating portfolio $(\gamma_0 + \gamma_1 x_t)$ by a given threshold value (TV), i.e. $(z_t > TV)$, in order to restore equilibrium y_t is required to fall in the next period while x_t is expected to increase. In this case, α_1 will be negative, $\frac{H}{d}\left(-N_x\right)$, and α_2 positive, $\frac{H}{d}\left(N_y\right)$, so pairs strategists will short y_t (outperformer) and buy x_t (underperformer) to exploit price divergences. This allows positive profits until temporary mispricing vanishes. Conversely when $(z_t < -TV)$, asset x_t is overpriced in t , which implies that α_2 will be negative $\frac{H}{d}\left(-N_y\right)$ and α_1

[4] Note that in the empirical part lags of ΔP are chosen in order to obtain white noise errors.

will be positive $\frac{H}{d}(N_x)$ to guarantee mean reversion of the error term. Note that the determination of the TV is described in Section 2.2 below.

2. If $z_t > 0$ and the asset y_t was contributing significantly to price discovery, α_2 will be positive and statistically significant as the asset x_t adjusts to incorporate new information. Similarly, if the market trading x_t is an important venue for price discovery and liquidity then α_1 would be negative and statistically significant. If both coefficients are significant then both markets contribute to price discovery. The existence of cointegration means that at least one market has to restore long-run equilibrium, indicating that the given market is under (over) priced and is short-term inefficient. Profits from pairs strategies can therefore be achieved. If the adjustment of both prices is immediate and independent of the cointegration error ($\alpha_1 = \alpha_2 = 0$), the elasticity of demand for pairs strategies is infinite ($H \to \infty$),[5] and there is no VECM, no price discovery, and no profit from pairs strategies.

3. In the VECM framework, the paired assets are modeled to converge to each other to restore equilibrium. The coefficients α_1 and a_2 are the adjustment coefficients, and measure the speed by which both assets adjust to long-run equilibrium. This is slow when the parameter is close to 0, and fast when it is close to 1. In the case where $\alpha_1 \neq 0$ and $\alpha_2 = 0$, asset x_t does not adjust to asset y_t as it is essentially the common factor or efficient price. (The reverse is true when $\alpha_1 = 0$ and $\alpha_2 \neq 0$).

4. Pairs trading strategies require stationarity of the error term. The error correction mechanism links directly the adjustment speed of paired series to the cointegration error, which follows an autoregressive (mean reverting) process specified as:

$$z_t = (1 - (-\alpha_1 + \gamma_1\alpha_2))z_{t-1} + u_t^y - \gamma_1 u_t^x \tag{1.5}$$

$$z_t = \rho z_{t-1} + u_t^y - \gamma_1 u_t^x \tag{1.6}$$

[5] In this case both markets are perfect substitutes and prices are "discovered" in both markets simultaneously. The model is not sustainable for this case.

In this context, the sum of the absolute values of α_1 and α_2 determines the persistence of the cointegration error. In the limit, when $\alpha_1 = \alpha_2 = 0$, z_t is $I(1)$, there is no cointegration and it is not possible to benefit from trading paired assets[6]. When α_1 and/or α_2 are statistically different from zero and correctly signed (see point 1), the error term will be mean reverting and pairs trading will be profitable. Note that incorrect estimated signs for α_1 and α_2 signals explosive behavior in the error term ($\rho > 1$).

In order to describe profits from pairs strategies we define the cointegration error as:

$$z_t = y_t - \gamma_0 - \gamma_1 x_t \tag{1.7}$$

Whenever the cointegration error reaches the model trigger so that y_t on the previous period is above its equilibrium level, there will be an arbitrage opportunity which requires that the investor shorts y_t in the same amount as the replicating portfolio (constructed with asset x_t) in order to profit from pairs strategies. Profits from this strategy may be defined as:

$$\Pi_t = M\left(-\Delta y_t + \gamma_1 \Delta x_t\right) = -M \Delta z_t \tag{1.8}$$

where Π_t are measured in €, y_t and x_t represent equity prices in €, and M is the amount invested (in €). Portfolio replication is defined in terms of price levels (and not returns) and the delta or hedge ratio for a short position in asset y_t will be γ_1 implying that γ_1 units of asset x_t are acquired to replicate the value of asset y_t. Portfolio allocations are therefore determined according to the regression coefficients of the cointegration relationship. Substituting the result in Equation (1.4), we get :

$$\begin{aligned}
\Delta y_t =& \alpha_1 z_{t-1} \\
\Delta x_t =& \alpha_2 z_{t-1} \\
\Pi_t =& M\left(-\alpha_1 + \gamma_1 \alpha_2\right) z_{t-1}
\end{aligned} \tag{1.9}$$

[6] The absolute value of the estimated adjustment coefficients has to lie between 0 and 1. A proof of this statement can be provided upon request

From Equation (1.6) we can write :

$$\Pi_t = M\,(1-\rho)\,z_{t-1} \tag{1.10}$$

Profits are therefore negatively related to error persistency so that the more stationary is the error term, the higher is pairs trading profitability. When $\rho > 1$, the cointegration spread is explosive and profits become negative. Profits are therefore increasing in the degree of stationarity and the size of arbitrage opportunities as reflected in the cointegration error.

When $z_{t-1} > 0$, asset y_t is overpriced in period $t-1$. This indicates that in time t as previously specified, under VECM dynamics, α_1 must be negative, and α_2 positive as they move to restore equilibrium.

1.2.2 Threshold design for the cointegration error

The trading algorithm dictates that arbitrage opportunities will be exploited when z_t exceeds a given threshold value TV. Under these circumstances the general principle is applied. This requires placing a new trade when the error deviates from equilibrium and unwinding the trade when equilibrium is restored. In order to optimally design the trading trigger, one has to specify what would qualify as a sufficient divergence of the cointegration error from its long-term equilibrium. The literature does not offer a closed form solution to this question. Instead, it demonstrates that the actual implementation of the trading algorithm is wide and varied (see Park and Switzer (1996), Avellaneda and Lee (2008), Elliott et al. (2005) and Vidyamurthy (2004) for description of threshold possibilities). Park and Switzer (1996) perform basis trading in the fixed-income market and define the trading trigger in terms of a combination of a moving average and a standard deviation calibrated with a tolerance parameter. Avellaneda and Lee (2008) estimate trading cutoffs based on mean reversion of a dimensionless variable while Vidyamurthy (2004) proposes various band designs for different assumptions regarding the spread process. Our approach is consistent with Vidyamurthy (2004) and Elliott et al. (2005) in that we define the threshold value in terms of the amount of volatility away from the mean. However we exploit the model-based result relating lower persistency to higher profitability as outlined in Equation (1.10) to define the calibrating parameter of the trading trigger. Using the results underlying Equation (1.5) we propose $\rho\sigma = (1 + \alpha_1 - \gamma_1\alpha_2)\,\sigma$ standard deviations of z_t as the model threshold. In this way we link the model threshold to the persistence of the error term so that more persistent errors require a higher threshold. Note that this is in line with Elliott et al. (2005), who inversely relate the trigger level to the speed of mean reversion in the spread. In our framework a trade will be triggered when $\|z_t\| > \rho\sigma$ and unwound when long-run equilibrium is restored so that $\|z_t\| \leq \rho\sigma$.

The threshold is also linked to pairs trading profitability as underlined in Equation (1.10). Higher profitablity requires lower thresholds.

1.2.3 Portfolio replication

Portfolio replication requires determination of price leadership in the context of price discovery. In this framework the dominant price is used to replicate the value of the follower. Price discovery measures the contribution of cointegrated assets to reveal information regarding a common factor that measures fundamentals.[7] It can be shown from VECM in Equations (1.2)-(1.4), that the contribution of assets y_t and x_t to price discovery is:

$$PD_y = \frac{\alpha_2}{\alpha_2 - \alpha_1} = \frac{N_y}{N_y + N_x} \tag{1.11}$$

$$PD_x = \frac{-\alpha_1}{\alpha_2 - \alpha_1} = \frac{N_x}{N_y + N_x} \tag{1.12}$$

If new information from both markets is incorporated into the common efficient price, $0 \leq PD_i \leq 1$ for $i = y, x$. Under the extreme case where $\alpha_1 = 0$, the price discovery metrics become $PD_y = 1$ and $PD_x = 0$ then there is a predominance of asset y_t in the price discovery process.[8] If $\alpha_2 = 0$ then we have $PD_x = 1$ and $PD_y = 0$ so that there is a predominance of asset x_t in terms of price discovery. Price discovery is exploited so that the price is used in

[7] See Hasbrouck (1995), Gonzalo and Granger (1995), and Lehmann (2002).
[8] Predominance in this context implies that the common fundamental factor is driven solely from the dominant price.

this model to replicate the value of the follower. Equations (1.11) and (1.12) demonstrate that price discovery relies on the relative number of agents in both markets which defines the relative speed of mean reversion.

1.3 Cointegration and price discovery

In this paper we focus on the European equity market in order to identify potential profitable opportunities underlying pairs trading strategies. We consider companies included in the STOXX Europe 600 index, restricting the analysis to those corporates that are located in the Eurozone such that their prices are quoted in euro currency. The selected sample consists of 292 companies across ten countries of the European common currency area. Daily closing price data are collected over the period dating from January 1st, 2000 to February 6th, 2017. This comprises an average of 4461 trading observations[9]. The data source is Datastream. Our sample period covers the pre-crisis period as well as the post Lehman era, therefore it allows us to analyze pairs trading under the existence of cointegration in different market states.

The presence of cointegration in this context indicates that two non-stationary $I(1)$ variables have a linear combination that is stationary, $I(0)$. In what follows, we identify a matching partner for each stock with the restriction that both stocks should belong to the same industry. According to the Industry Classification Benchmark used by STOXX indices[10], the 292 companies are categorized into ten industries, namely, Financials (63), Industrials (58), Consumer Goods (42), Consumer Services (33), Basic Materials (23), Utilities (20), Health Care (16), Technology (15), Telecommunications (11), and Oil & Gas (11). The model presented in Section 2 demonstrates that the mechanism behind cointegration lies on the existence of an underlying common efficient price. Therefore in addition to the statistical sense, assets are linked into the equilibrium underlined in Equation (1.1) due to the existence of shared fundamentals. To account for common fundamentals we impose two restrictions: (i) the industry restriction; (ii) the firm fundamental factor restriction. The

[9] Note that not all companies have the same starting date.

[10] This classification is based on the companies´ primary revenue source.

industry restriction is expected to convey shared supply and demand conditions as well as common regulation exposures. Given that the sample is restricted to the same currency area, we expect the underlined commonalities to drive prices to parity. In what follows, the industry restriction is taken as the baseline restriction. The firm fundamental restriction is considered jointly with the sector restriction in Section 4.5.

Our empirical analysis is based on the VECM specified in Equation (1.4). Econometric details of the estimation and inference of (1.4) can be found in Johansen (1995) and Juselius (2006).

We start by performing an Augmented Dickey-Fuller test as unit roots are a necessary condition for cointegration. We fail to reject the null hypothesis of a unit root for all price series analyzed. We perform the Johansen cointegration test using a rolling-window approach within a given industry at the 5% significant level. Specifically, we use a three-year window from t to $t+3$ (estimation period) to identify cointegrated pairs and, for each selected pair, estimates of the cointegration vector in the VECM are obtained. The selected pairs and resulting estimates are then applied to trading implementation (see the detailed description in Section 4) for the next six-month window from $t+3$ to $t+3.5$ (trading period). This procedure is repeated through the remaining sample period. This leaves us, for instance, with the second estimation window from $t+0.5$ to $t+3.5$, which is followed by the trading window from $t+3.5$ to $t+4$. Given that not all companies have been listed at the first sample date, January 1st, 2000, the starting date of a possible pair is chosen so that transactions are available on both corporates considered. The resulting paired equities are tied via a long-run arbitrage relationship under the imposed restriction that the error term is stationary. In the sense that paired equities are close substitutes, they tend to move in synchrony.

Once cointegration relationships are estimated, we investigate the lead-lag relationship for each pair to determine which asset dominates the price discovery process. Table 1 reports VECM estimates across industries. Because there are thirty different rolling samples, reported results represent an average value computed from a series of estimates for each percentile. We find from Panel A that the adjustment coefficient α_1 is significantly negative for all industry percentiles[11], suggesting that the price of the follower (y_t) is expected to drop by α_1 units in response to one unit increase in the error correction term. The corresponding estimate of α_2 is positive under all percentiles. Results of α_2 also suggest that it is not significantly different from zero in 80% of the estimations. This implies that there is an asymmetric lead-lag relation in 80% of the paired corporates. For the remaining 20%, both assets contribute to price discovery. However, the general result is that there is a dominant asset (x_t) relative to its matching partner (y_t) in terms of price discovery, and thus the follower (y_t) does all the adjustment to restore long-term equilibrium. Note that this result comes by construction given that the leading asset (x_t) is used to replicate the follower (y_t). Effectively Johansen cointegration estimates are obtained in a context in which the follower (or the dependent variable) is set to be explained by the leader which acts as the independent variable. The existence of cointegration allows using Maximum Likelihood estimators of the cointegrating relation to build our portfolio strategy, instead of OLS regressions as widely used in the statistical arbitrage literature[12]. The (constant) cash amount, γ_0, is required to be positive[13]. The positive sign of γ_0 suggests that long cash positions should be held to replicate the follower with γ_1 units of the leading asset, and thus interest expenses are omitted from the construction of arbitrage profits. Then we look at the estimated values of γ_1 reported in Panel B. The values are varied as different units of the leader asset are required to replicate the follower. Note that the largest range for γ_1 is for Financials and Industrials, the two sectors with highest number of paired companies. This coefficient reflects the sensitivity of one asset to its matching partner and is in essence the hedge ratio in our pairs trading strategy.

[11] Average standard errors by industry can be provided upon request.
[12] see for instance Schaefer and Strebulaev (2008) and Kapadia and Pu (2012).
[13] Estimated values of γ_0 can be provided upon request.

Table 1.1: Estimation of VECM parameters

Sample		5th Percentile	25th Percentile	Median	75th Percentile	95th Percentile
Panel A: Estimated values of α_1 and α_2						
Financials	α_1	-0.013	-0.036	-0.076	-0.150	-0.302
	α_2	0.001	0.005	0.012	0.029	0.076
Industrials	α_1	-0.020	-0.038	-0.066	-0.112	-0.323
	α_2	0.001	0.007	0.015	0.028	0.058
Consumer Goods	α_1	-0.021	-0.052	-0.090	-0.136	-0.201
	α_2	0.002	0.008	0.018	0.036	0.077
Consumer Services	α_1	-0.017	-0.036	-0.063	-0.115	-0.254
	α_2	0.002	0.006	0.013	0.024	0.052
Basic Materials	α_1	-0.040	-0.059	-0.082	-0.112	-0.170
	α_2	0.004	0.012	0.024	0.041	0.064
Utilities	α_1	-0.009	-0.014	-0.028	-0.048	-0.089
	α_2	0.001	0.002	0.005	0.009	0.023
Health Care	α_1	-0.046	-0.060	-0.087	-0.125	-0.200
	α_2	0.004	0.010	0.018	0.032	0.052
Technology	α_1	-0.035	-0.046	-0.067	-0.092	-0.147
	α_2	0.005	0.010	0.020	0.032	0.052
Telecommunications	α_1	-0.011	-0.017	-0.028	-0.038	-0.059
	α_2	0.002	0.004	0.008	0.015	0.023
Oil & Gas	α_1	-0.026	-0.038	-0.053	-0.070	-0.090
	α_2	0.002	0.004	0.011	0.020	0.031
Panel B: Estimated values of γ_1						
Financials		0.14	0.45	1.43	7.06	11.64
Industrials		0.14	0.42	0.94	2.76	10.65
Consumer Goods		0.26	0.58	1.30	2.88	7.99
Consumer Services		0.28	0.49	0.88	1.95	4.96
Basic Materials		0.32	0.46	0.96	1.83	3.33
Utilities		0.23	0.45	0.93	3.34	9.86
Health Care		1.07	1.27	1.79	2.76	4.07
Technology		0.54	0.76	1.15	1.99	5.15
Telecommunications		2.17	2.25	2.62	4.25	6.60
Oil & Gas		0.64	0.80	1.10	2.49	3.99

This table presents the values of α_1 and α_2 obtained using the Johansen cointegration methodology in Panel A. The percentiles for α_1 are computed using the absolute values. Summary statistics of estimated values of γ_1 are reported in Panel B. As the Johansen test is conducted on a rolling-window basis, these reported values are an average value computed from a series of estimates of each percentile. The sample period is January 2000 to February 6th 2017.

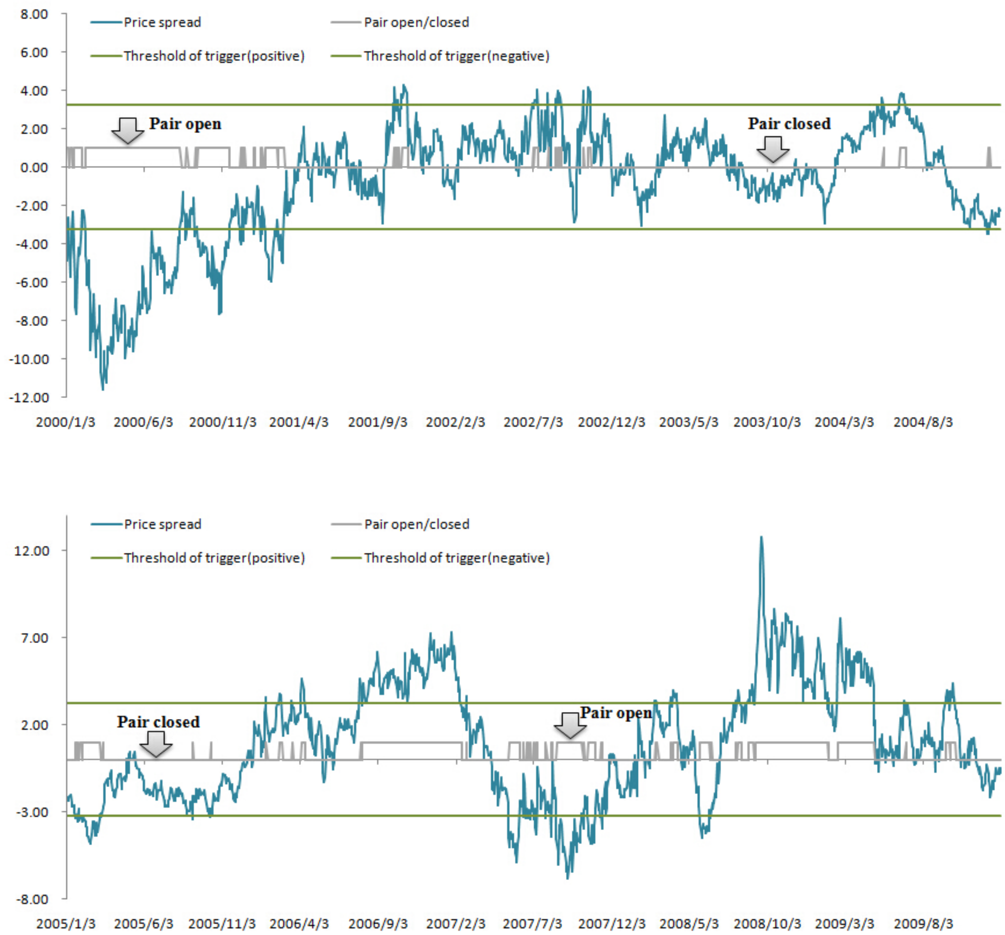
1.4 Profitability of pairs trading

1.4.1 Model-based pairs trading strategies

In this section, we illustrate the proposed pairs trading mechanism presented in Section 2, based on the existence of cointegration. The trading mechanism is described as follows: An arbitrageur opens a long-short position on the day following departure, when the price spread hits the model-derived threshold, denoted as $(1+\alpha_1-\gamma_1\alpha_2)$ units of standard deviation calculated from historical spreads. The initial position is then unwound one day later when price reversion eventually occurs, or is forced to close at the end of a 6-month trading period if no convergence takes place. In other words, we trade according to the rule that delays the opening/close of a position by one day, and the maximum trading horizon is six months. After a pair has completed a round-trip trade, it will be subject to the identical trading rule again. As previously discussed, this paper applies three-year rolling-window approach. This results in a series of cointegration coefficients and speed of mean reversion estimates. These estimates are used to determine pricing errors to design the following six-month out-of-sample trading strategies. The implementation of trading requires construction of replicating portfolios and trading triggers immediately following the three-year estimation period. Since most data are available from January 1st, 2000, their first valid trading day is the first business day in January 2003.

This trading mechanism, which longs the underpriced asset and shorts the overpriced one simultaneously, is implemented according to the sign of the estimated alpha coefficients. Theoretical profits are always positive and defined by return differentials as specified in Equations (1.8) and (1.9). Our pairs selection algorithm is driven by cointegration, which implies that profits generated by the proposed strategy are expected to be stationary. As stated in the model, cointegration guarantees that short-lived price deviations revert towards equilibrium, such that the slow adjustment process can be exploited to make profits. With this trading rule Figure 1 illustrates how to perform the strategy using the cointegrated pair, Air Liquide and BASF, as an example. The fluctuating line in blue represents price spread z_t

Figure 1.1: Price spreads between Air Liquide and BASF and pairs trading establishment



This figure illustrates how to perform pairs trading strategy using the cointegrated pair, Air Liquide and BASF, during the period 2000-2009.

, while the two straight lines in green indicate the border of model-derived threshold (either positive or negative). The line in grey, near to the x-axis, reflects the opening and close of pairs trades on a daily basis. We see that a position is initiated when the price spread moves beyond the border and then closed when price deviation lies between two border lines.

1.4.2 The baseline results

We analyze the performance of pairs trading strategies for each industry, under the "persistence calibrated" standard deviation trigger. Risk and return characteristics are examined at the portfolio level. In addition to forming equal-weighted portfolios, we calculate buy-

Table 1.2: Model-derived trading trigger $1 + \alpha_1 - \gamma_1\alpha_2$

Sample	5th Percentile	25th Percentile	Median	75th Percentile	95th Percentile
Financials	0.65	0.80	0.89	0.93	0.97
Industrials	0.80	0.87	0.91	0.93	0.98
Consumer Goods	0.79	0.84	0.89	0.93	0.97
Consumer Services	0.81	0.86	0.91	0.94	0.96
Basic Materials	0.83	0.85	0.88	0.90	0.92
Utilities	0.90	0.91	0.94	0.95	0.96
Health Care	0.83	0.84	0.86	0.88	0.90
Technology	0.84	0.87	0.90	0.92	0.97
Telecommunications	0.89	0.93	0.94	0.97	0.98
Oil & Gas	0.88	0.90	0.91	0.93	0.94

This table presents the values of model-derived trading trigger $1+\alpha_1-\gamma_1\alpha_2$, which is computed using VECM estimates obtained from the Johansen cointegration methodology. As the trading strategy is conducted on a rolling-window basis, these reported values are an average value computed from a series of threshold numbers of each percentile. The sample period is January 2000 to February 6th 2017.

and-hold portfolio returns, following Gatev et al. (2006). This approach takes into account compounded returns, known as value-weighted portfolio return. The return computation is thus based on daily marked-to-market returns of individual pairs.

Table 2 reports estimated percentiles of trading trigger for each of the industry groups aggregated over the set of 30 rolling samples. Reported results demonstrate that there is error persistence delivering average trading triggers ranging from 0.65 to 0.98. Table 3 reports the excess return distribution by industry group and for the all-pair portfolio, representing all tradable targets, over the whole out-of-sample period since 2003.[14] Panel A shows results from the equal-weighted portfolio, while Panel B presents profit estimates from the value-weighted portfolio. The general observation is that pairs portfolios gain statistically significant positive excess returns. We can see in Panel A that the equal-weighted portfolio generates an annualized average return of at least 4.33%. Moreover only two portfolios

[14] Our strategy's profitability is induced from two positions. In this context the payoff is interpreted as excess return since trading profits or losses are earned from one euro investment in simultaneous long-short positions.

Table 1.3: Summary statistics of excess returns to pairs portfolios

Sample	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: Equal-weighted portfolios								
Financials	0.0746	0.0000	0.1467	0.57	14.44	0.07	-0.07	0.51
Industrials	0.0433	0.0000	0.1318	0.15	10.76	0.07	-0.07	0.33
Consumer Goods	0.0514	0.0000	0.0943	0.69	17.45	0.06	-0.05	0.54
Consumer Services	0.0772	0.0000	0.1235	0.48	14.17	0.09	-0.06	0.63
Basic Materials	0.0662	0.0000	0.1642	0.08	18.54	0.09	-0.14	0.40
Utilities	0.0553	0.0000	0.1585	1.40	29.37	0.14	-0.10	0.35
Health Care	0.0644	0.0000	0.1163	-1.60	43.23	0.05	-0.14	0.55
Technology	0.0587	0.0000	0.1728	0.67	13.63	0.12	-0.07	0.34
Telecommunications	0.0470	0.0000	0.1735	1.08	22.32	0.14	-0.08	0.27
Oil & Gas	0.0579	0.0000	0.1034	0.94	18.12	0.07	-0.05	0.56
All-Pair Portfolio	0.0576	0.0154	0.0681	0.89	14.25	0.04	-0.03	0.85
Panel B: Value-weighted portfolios								
Financials	0.0506	0.0000	0.0700	0.62	16.38	0.03	-0.04	0.72
Industrials	0.0246	0.0000	0.0639	0.26	10.47	0.03	-0.03	0.38
Consumer Goods	0.0272	0.0000	0.0454	0.73	15.61	0.03	-0.02	0.60
Consumer Services	0.0467	0.0000	0.0613	0.56	16.20	0.04	-0.03	0.76
Basic Materials	0.0346	0.0000	0.0808	0.26	18.06	0.05	-0.06	0.43
Utilities	0.0292	0.0000	0.0796	1.88	38.70	0.08	-0.05	0.37
Health Care	0.0455	0.0000	0.0960	-0.12	11.83	0.05	-0.06	0.48
Technology	0.0303	0.0000	0.0849	0.62	13.17	0.06	-0.03	0.36
Telecommunications	0.0313	0.0000	0.0967	1.06	23.31	0.06	-0.06	0.32
Oil & Gas	0.0309	0.0000	0.0504	1.03	18.11	0.03	-0.02	0.61
All-Pair Portfolio	0.0335	0.0051	0.0326	0.90	12.78	0.02	-0.01	1.03

This table presents descriptive statistics of excess returns for each industry group and the all-pair portfolio. We trade according to the rule that opens a position in a pair one day after price spread diverges more than $(1+\alpha_1-\gamma_1\alpha_2)$ units of historical standard deviation. Reported are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return and (annualized) Sharpe ratio. The sample period is January 2000 to February 6th 2017.

earn mean returns lower than 5%. Among the ten industries, pairs from Consumer Services deliver the highest average return equal to 7.72% per annum, followed by 7.46% earned in the Financials industry. Results therefore show clear positive performance which is consistent across different industries. We next look at the risk profile, measured by volatility. We can see that the first four industry portfolios (Financials, Industrials, Consumer Goods, Consumer Services) show lower volatility and maximum return values than the remaining portfolios (Basic Materials, Utilities, Health Care, Technology, Telecommunications and Oil & Gas). This overall evidence indicates diversification benefits created from combining a larger number of pairs in a portfolio. This can be explained by the fact that the first four industries include more companies within the group. For this reason, they exhibit similar statistical properties. Reported results also demonstrate that the return distributions of industry portfolios are positively skewed, only with the exception of Health Care.[15] This implies that reported Sharpe ratio estimates may exhibit a downward bias. The finding of right-skewed distribution is consistent with Gatev et al. (2006) and Jurek and Yang (2007). However it is not supported by the work of Kondor (2009) who argues that arbitrageurs' total return is negatively skewed. Sharpe ratios are reported in the last column to evaluate the risk-adjusted portfolio performance.[16] We find that half of these industry portfolios deliver a Sharpe ratio higher than 0.50. We contend that such impressive performance is clearly associated with the sector that gathers a larger number of companies. Our results therefore demonstrate that the proposed strategy is profitable for every industry categorized under STOXX Europe 600, and therefore the existence of positive profits is not industry dependent. The last row of Panel A (Table 3), examines the overall pairs trading performance taking together all pairs selected within industries, thus allowing investment on all tradable opportunities. The average annual return is 6.0% for this all-pair portfolio. More impressive is the annualized Sharpe ratio of 0.85. This arises due to the large gains arising from diversification across pairs that trade in different industries as can be seen in the reported

[15] The negative skew in the Health Care industry may be due to the high downside risk arising from high investment in R&D.

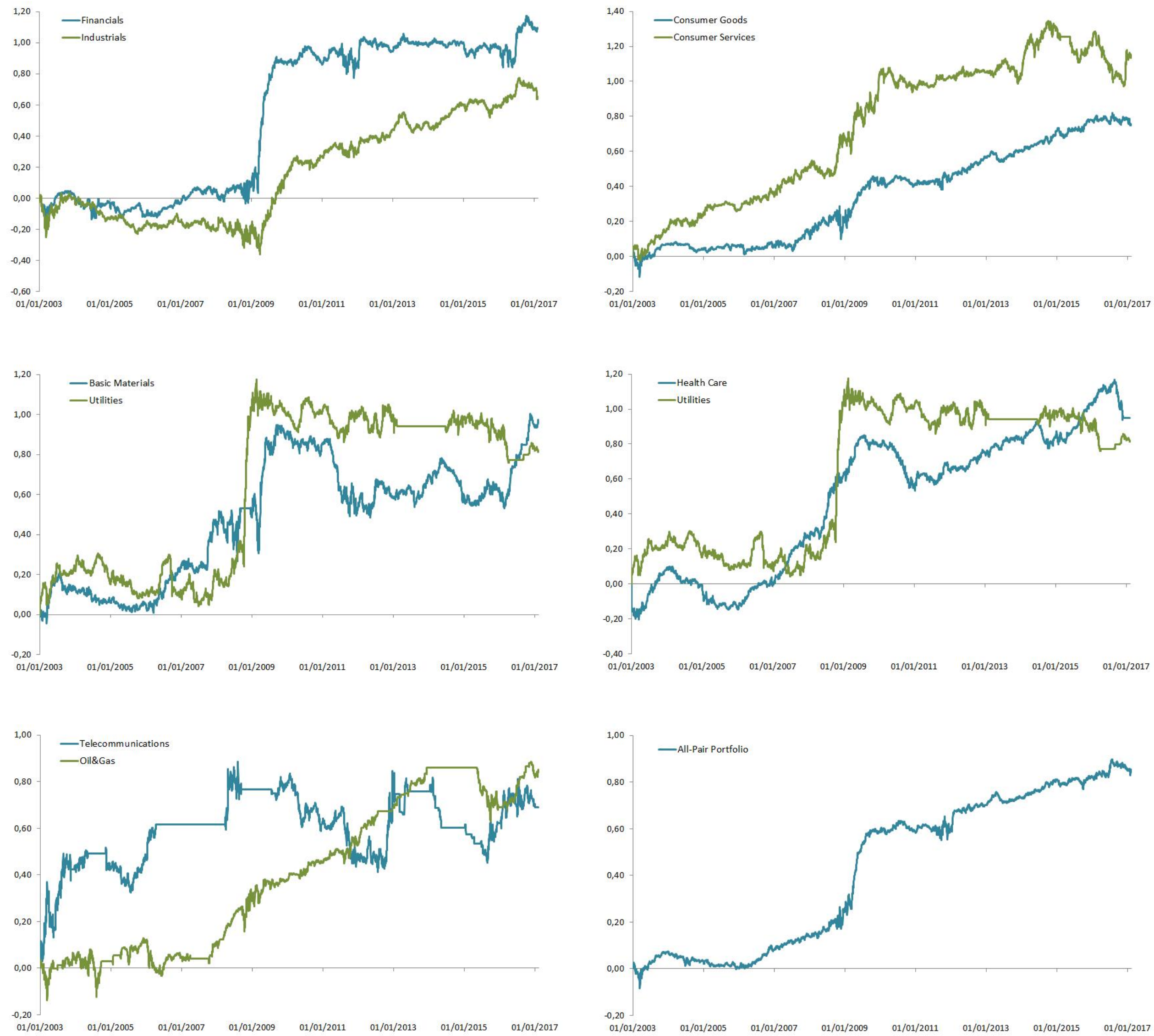
[16] Simplified Sharpe ratios assume zero returns on the risk-free asset exploiting the fact that interest rates have been at historical minimum levels over our sample period. All Sharpe ratios reported in the paper are simplified under this assumption.

volatility of 6.81%. In line with our findings from industry portfolios, cointegration delivers positive skewed portfolio returns.

To unfold the economic significance from arbitrageurs' perspective, we consider the cumulative portfolio returns over our sample period. Figure 2 plots the cumulative profits on the equally weighted industry portfolios and the all-pair portfolio. These reinvested payoffs depict the evolution of an investor´s wealth. It is observed that these portfolios earn cumulative profits of different magnitudes and, more interestingly, their returns exhibit various patterns. For instance, the industry portfolios, Consumer Goods and Consumer Services, accumulate wealth in a steady manner without considerable losses. Other sectors such as Financials or Utilities exhibit clearly pronounced rises over the 2008-2009 period. This evidence indicates that pairs belonging to distinct industries exhibit different response to market conditions leading to unresembling paths of return cumulation. However the common pattern across all portfolios is that there is a significant increase in profitability with the unfolding of the global financial crisis. This therefore suggests that our model-based pairs strategies can be used to hedge away market shocks and simultaneously yield significant profits under abnormal periods. This is consistent with the literature (see for example Alexander et al. (2002) and Do and Faff (2010)). The evolutionary path of wealth underlying the all-pair portfolio suggests that there is persistent profitability without the requirement of stop-loss criterion, over a 14-year period. This portfolio produces a total return of 0.86 at the end of sample.

We report in Panel B of Table 3 the results of value-weighted portfolios. Reported results show positive excess returns. However the resulting profits are lower than those from equal-weighted portfolios. The decrease in profitability can also be found in Chen et al. (2012). Risk-adjusted returns are however improved with respect to the equally weighted cases. The improvement in Sharpe ratio estimates is documented for most of industry portfolios as well as the all-pair portfolio. This arises as a consequence of lower return volatility. Value-weighted portfolios, are by construction less volatile than equal-weighted portfolios as weights in time t are defined according to lagged returns. The all-pair portfolio earns on average an annual excess return of 3.35%, and a Sharpe ratio of 1.03. The latter represents

Figure 1.2: Cumulative returns for industry pairs portfolios and the all-pair portfolio



This figure plots the cumulative excess returns of equal-weighted pairs portfolios over the period January 2000 to February 6th 2017.

Table 1.4: Summary statistics of excess returns to pairs portfolios after transaction costs

Sample	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
				Equal-weighted portfolios				
Financials	0.0325	0.0000	0.0901	0.83	20.76	0.06	-0.05	0.36
Industrials	0.0157	0.0000	0.0834	0.80	22.25	0.07	-0.05	0.19
Consumer Goods	0.0314	0.0000	0.0942	0.69	17.56	0.06	-0.05	0.33
Consumer Services	0.0618	0.0000	0.1228	0.48	14.48	0.09	-0.06	0.50
Basic Materials	0.0470	0.0000	0.1643	0.07	18.49	0.09	-0.14	0.29
Utilities	0.0395	0.0000	0.1532	1.09	22.86	0.08	-0.05	0.27
Health Care	0.0428	0.0000	0.1163	-1.60	43.24	0.05	-0.14	0.37
Technology	0.0472	0.0000	0.1727	0.68	13.68	0.12	-0.07	0.27
Telecommunications	0.0282	0.0000	0.1299	1.35	26.32	0.10	-0.06	0.22
Oil & Gas	0.0487	0.0000	0.1054	0.62	19.31	0.07	-0.06	0.46
All-Pair Portfolio	0.0356	0.0000	0.0674	0.94	14.77	0.04	-0.03	0.53

This table presents descriptive statistics of excess returns net of transaction costs, for each industry group and the all-pair portfolio. Transaction costs are estimated as one-half of the sum of the bid-ask spreads on both assets. Reported are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return and (annualized) Sharpe ratio. The sample period is January 2000 to February 6th 2017.

an improvement of 21% when compared to the equal-weighted all-pair portfolio.

The baseline analysis above concludes that our pairs trading strategy is profitable. The solid performance confirms price convergence after the onset of a pricing anomaly, demonstrating that pricing errors are stationary.

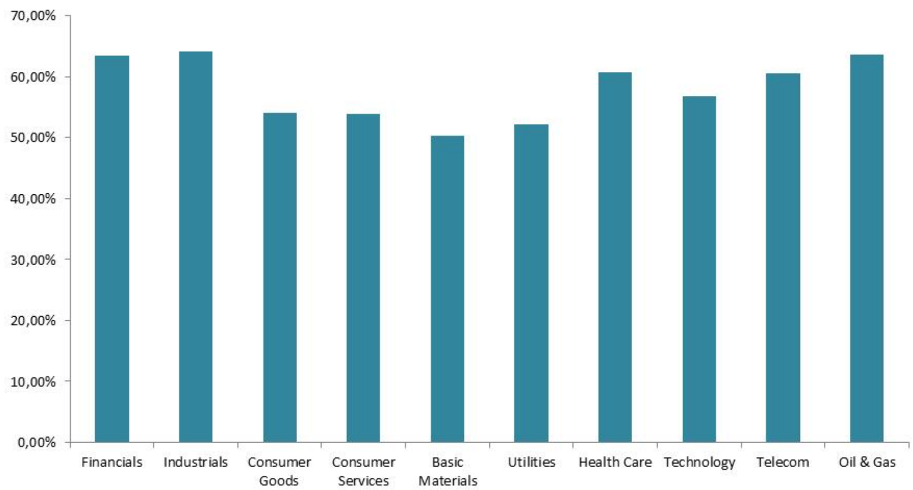
The next step is to introduce transaction costs and assess the impact on strategy returns. Given that closing price data are used to compute abnormal returns, there is an identical probability of being at bid or ask. We impose corrections of these profits to reflect that, in practice, we long at the ask and sell at the bid prices. In other words, we have to subtract trading costs to get an estimate of profits net of transaction costs. To this end, we collect bid and ask prices for each equity pair. Because we long the loser and short the winner asset, transaction costs will reduce profits by one-half of the sum of the bid-ask spreads on both assets every time there is a change in position in the pair. Results reported in Table 4 reveal that, after accounting for transaction costs, Sharpe ratios and mean returns remain positive in all cases.

1.4.3 Price discovery, relative turnover and portfolio replication

Equations (1.11) and (1.12) show that price discovery is defined in terms of the relative number of participants in paired markets. Proxies for the relative number of participants can be found on relative liquidity measures. Another approach in the literature has been to address the relative number of analyst or informed traders following a given firm (see Brennan et al. (1993)). In this paper we follow FFG and use liquidity measures for this purpose. The relationship between price leadership and trading volume has been extensively addressed in the price discovery literature (see Hasbrouck (1995) and references therein). It has also been discussed in the asset pricing literature. Chordia and Swaminathan (2000) and Llorente et al. (2002) show that trading volumes deliver valuable information about future price movements and exerts impact on the speed of adjustment of individual stocks (see also Admati and Pfleiderer (1988)). The underlying presumption is that stocks that trade with higher liquidity respond faster to common information and become information leaders. Stocks that trade with low liquidity are therefore followers as they exhibit lower speed of price adjustment to common information. Variations in relative liquidity in both markets are therefore reflected on the changes in price discovery.

In what follows we check whether the result of price discovery in the theoretical model is confirmed empirically. We use for this purpose turnover as a proxy for market liquidity. Turnover is defined as the ratio of the number of shares traded in a day to the number of shares outstanding at the end of the day. Our interest is to see whether price leadership is associated with higher turnover. Specifically, we check for every pair screened out over each trading period, and calculate the percentage of them meeting the condition that the leader's average turnover is higher than the follower. Figure 3 shows the results. We can see that price

Figure 1.3: The association of price leadership with trading volume of individual stocks



This figure shows, for each industry group, the percentage of leading assets in the price discovery process whose average trading volume is higher than the follower, over the sample period January 2000 to February 6th 2017.

leadership is associated with higher turnover for more than half of the pairs. This ratio is even higher for Financials (63%), Industrials (64%), Health Care (61%), Telecommunications (61%), and Oil & Gas (64%).

We now proceed to demonstrate the importance of price leadership determination for portfolio replication design. For this purpose, we repeat the trading exercise in Section 4.2 using the follower to replicate the leader. We denote these portfolios as "follower" portfolios. For comparison, Table 5 summarizes the excess returns for pairs portfolios which are present in Table 3. From Panel A we find that average returns of industry portfolios are lower compared to the baseline results. The same observation is applied to the performance metrics of Sharpe ratio. Comparing the all-pair portfolio return estimates under the equal-weighted method, the baseline portfolio yields a mean return which is 1.42% higher. It also outperforms the follower portfolio in terms of Sharpe ratio, increasing from 0.74 (follower portfolio) to 0.85 (baseline portfolio). These findings are held for the case of value-weighted portfolios.

Results therefore confirm that the theoretical implications on price leadership of the model exposed in Section 2 are important to maximize pairs trading profitability.

Table 1.5: Excess returns to alternative pairs trading: the switch of leadership

Sample	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: Equal-weighted portfolios								
Financials	0.0533	0.0000	0.1395	1.09	19.12	0.09	-0.08	0.38
Industrials	0.0404	0.0000	0.1530	-0.36	18.89	0.08	-0.11	0.26
Consumer Goods	0.0443	0.0287	0.1182	-0.90	22.70	0.06	-0.11	0.37
Consumer Services	0.0575	0.0000	0.1055	0.46	10.28	0.05	-0.05	0.55
Basic Materials	0.0288	0.0000	0.1618	-0.28	19.74	0.09	-0.14	0.18
Utilities	0.0466	0.0000	0.1137	0.60	19.90	0.08	-0.07	0.41
Health Care	0.0511	0.0000	0.1306	0.07	20.70	0.08	-0.11	0.39
Technology	0.0276	0.0000	0.1509	0.40	12.91	0.07	-0.07	0.18
Telecommunications	0.0331	0.0000	0.1440	0.66	20.82	0.09	-0.07	0.23
Oil & Gas	0.0405	0.0000	0.1579	-0.25	14.72	0.09	-0.08	0.26
All-Pair Portfolio	0.0435	0.0000	0.0585	0.40	12.00	0.03	-0.02	0.74
Panel B: Value-weighted portfolios								
Financials	0.0401	0.0000	0.0635	1.30	17.28	0.04	-0.03	0.63
Industrials	0.0230	0.0000	0.0737	-0.12	16.66	0.04	-0.05	0.31
Consumer Goods	0.0333	0.0117	0.0579	-0.41	18.23	0.03	-0.05	0.58
Consumer Services	0.0332	0.0000	0.0509	0.64	9.44	0.03	-0.02	0.65
Basic Materials	0.0141	0.0000	0.0809	-0.13	18.06	0.05	-0.07	0.17
Utilities	0.0264	0.0000	0.0574	0.83	21.64	0.04	-0.03	0.46
Health Care	0.0244	0.0000	0.0655	0.24	16.95	0.04	-0.05	0.37
Technology	0.0209	0.0000	0.0781	0.62	14.87	0.04	-0.03	0.27
Telecommunications	0.0211	0.0000	0.0719	0.73	19.90	0.04	-0.04	0.29
Oil & Gas	0.0200	0.0000	0.0787	-0.32	14.47	0.04	-0.04	0.25
All-Pair Portfolio	0.0253	0.0013	0.0278	0.79	9.24	0.01	-0.01	0.91

This table presents descriptive statistics of excess returns for each industry group and the all-pair portfolio. We trade according to the rule that opens a position in a pair one day after price spread diverges more than $(1+\alpha_1-\gamma_1\alpha_2)$ units of historical standard deviation. Contrary to portfolios in Table 3, the leadership is switched when establishing pairs portfolios. That is, we use the price of follower to replicate the leader in this alternative trading rule. Reported are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return and (annualized) Sharpe ratio. The sample period is January 2000 to February 6th 2017.

1.4.4 A comparison of performance: Model-based trading algorithm versus GGR (2006)

The purpose of this section is to compare the results of the proposed strategy with the results arising from application of the methodology introduced by Gatev et al. (2006), acknowledged as the benchmark work in the pairs trading literature. GGR identify pairs by minimizing the sum of squared spreads between two normalized price series in a 1-year period. Although both GGR and our trading strategy exploit mean reversion in search of profitability, the seminal work of GGR is based on non-parametric past return correlation. Our strategy instead follows a model-based trading algorithm. To derive results under GGR methodology, we identify pairs within each of the ten industries and rank them by distance. In order to make the two trading algorithms comparable, we set the number of pairs included in an industry portfolio to be the same as the number of cointegrated pairs under our method, for every trading period. In other words, we guarantee that the portfolio size is identical under two different selection methods. The objective is to remove diversification effect caused by sample size.

Following GGR, a trade is initiated if the prices diverge by more than two standard deviations. In Panel A of Table 6, we summarize the return performance delivered by GGR´s equally weighted portfolios. Compared to the baseline results in Table 3, the mean portfolio return associated with our strategy is larger in all cases than that associated with GGR. The magnitude of return gap is between 1% and 3% per annum. On the basis of volatility, both strategies perform closely. In terms of the risk-return tradeoff, we find that our model-based industry portfolios yield Sharpe ratios that exceed GGR portfolios in seven out of ten cases. The magnitude of outperformance from our methodology is also significant. Furthermore, our all-pair portfolio, gains an annualized mean return which is 2.4% higher than that reported for GGR, and generates a Sharpe ratio of 0.85, ten units higher than 0.75 achieved under GGR. Similar conclusions are reached when we compare on the basis of value-weighted portfolios (Panel B of Table 6).

Table 1.6: Summary statistics of excess returns to GGR portfolios

Sample	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: Equal-weighted portfolios								
Financials	0.0685	0.0000	0.1854	0.43	11.14	0.12	-0.12	0.37
Industrials	0.0324	0.0000	0.0870	0.80	45.21	0.06	-0.05	0.37
Consumer Goods	0.0313	0.0000	0.0533	1.59	25.52	0.04	-0.02	0.59
Consumer Services	0.0688	0.0000	0.1430	2.22	40.89	0.22	-0.14	0.48
Basic Materials	0.0167	0.0000	0.0837	-0.82	28.88	0.09	-0.17	0.20
Utilities	0.0326	0.0000	0.0872	1.84	31.66	0.06	-0.04	0.37
Health Care	0.0461	0.0000	0.1017	0.21	54.89	0.09	-0.09	0.45
Technology	0.0403	0.0000	0.1797	1.79	27.40	0.13	-0.09	0.22
Telecommunications	0.0359	0.0000	0.1941	1.23	26.63	0.13	-0.12	0.19
Oil & Gas	0.0262	0.0000	0.0971	1.35	34.28	0.11	-0.07	0.27
All-Pair Portfolio	0.0335	0.0000	0.0445	1.83	27.95	0.03	-0.02	0.75
Panel B: Value-weighted portfolios								
Financials	0.0415	0.0000	0.1021	0.06	29.57	0.10	-0.11	0.33
Industrials	0.0188	0.0000	0.0402	0.81	47.47	0.03	-0.02	0.62
Consumer Goods	0.0151	0.0000	0.0286	1.50	24.06	0.02	-0.01	0.81
Consumer Services	0.0387	0.0000	0.0790	2.25	37.90	0.11	-0.06	0.35
Basic Materials	0.0061	0.0000	0.0828	-0.87	28.83	0.04	-0.08	0.07
Utilities	0.0194	0.0000	0.0434	2.11	33.87	0.03	-0.02	0.45
Health Care	0.0567	0.0000	0.0805	0.57	22.22	0.06	-0.06	0.71
Technology	0.0158	0.0000	0.0918	1.84	29.36	0.07	-0.05	0.17
Telecommunications	0.0264	0.0000	0.1037	1.99	32.23	0.08	-0.05	0.25
Oil & Gas	0.0128	0.0000	0.0513	1.24	34.24	0.06	-0.04	0.20
All-Pair Portfolio	0.0206	0.0000	0.0249	1.69	21.80	0.01	-0.01	0.83
Panel C: Equal-weighted all-pair portfolios at two different thresholds								
1.5 standard deviations	0.0313	0.0000	0.0463	1.71	25.50	0.03	-0.02	0.68
3 standard deviations	0.0173	0.0000	0.0386	1.54	38.96	0.03	-0.02	0.45

This table presents statistics of excess returns for industry groups and the all-pair portfolio under GGR. We trade based on the rule that opens a position in a pair one day after price spread diverges over 2 standard deviations in Panel A and B, while the same rule is applied but subject to 1.5 and 3 standard deviations in Panel C. Reported are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return and (annualized) Sharpe ratio. The sample period is Jan. 2000 to Feb. 6th 2017.

Panel C of Table 6 shows results under GGR for alternative trading thresholds. Specifically we look at the profitability of the 1.5 and 3 standard deviations thresholds. It is clear that the alternative thresholds underperform the commonly used 2 standard deviations and therefore our model-based algorithm.

Reported results therefore demonstrate that the combination of cointegration and a persistence-linked trading trigger proposed in this framework superceeds the benchmark methodology proposed in the literature.

1.4.5 Commonality within industry groups: Fundamental factors analysis

The theoretical framework presented in Section 2 shows that two firms that share common fundamentals are linked via a long-run cointegration relationship. In our baseline empirical application, shared fundamentals arise due to common shocks at the industry and geographical level. In this section we look at pairs trading performance imposing firm-specific fundamentals (or factor) restrictions. Common firm fundamentals are expected to result in stable equilibrium relationships. Furthermore we test to which extent firm-level commonalities improve the performance of pairs trading strategies.

We follow the asset pricing literature (see Fama and French (1993), Fama and French (1996), and Asness et al. (2013)) to consider a value factor measuring the long-run (or book) value relative to its current market value (BV/MV). We additionally control for the size effect using market capitalization and trading volume as proxy measures. In particular we sort firms within a given industry, based on each of the factors (market cap, book-to-market ratio, and trading volume), and classify them into tercile portfolios. Then within each tercile, we identify a matching partner for every firm. Note that under this restriction, the number of selected pairs is much lower than under the baseline study in Section 4.2. Given that firms are ranked into three groups, we are restricted to do this analysis on four industries for which the number of firms is more than 30: Financials, Industrials, Consumer Goods, and Consumer Services. Tables 7 and 8 present percentiles for coefficient estimations and

Table 1.7: VECM estimates and the model-derived trading trigger for pairs controlling book-to-market ratio

Sample		5th Percentile	25th Percentile	Median	75th Percentile	95th Percentile
		Panel A: Estimated values of α_1 and α_2				
Financials	α_1	-0.013	-0.041	-0.079	-0.146	-0.306
	α_2	0.001	0.005	0.014	0.030	0.076
Industrials	α_1	-0.023	-0.039	-0.067	-0.120	-0.312
	α_2	0.002	0.008	0.016	0.029	0.064
Consumer Goods	α_1	-0.023	-0.047	-0.090	-0.142	-0.200
	α_2	0.003	0.009	0.017	0.034	0.067
Consumer Services	α_1	-0.037	-0.052	-0.077	-0.122	-0.260
	α_2	0.006	0.009	0.016	0.026	0.052
		Panel B: Estimated values of γ_1				
Financials		0.18	0.45	1.35	4.58	7.92
Industrials		0.19	0.45	0.94	2.68	10.08
Consumer Goods		0.40	0.74	1.54	2.96	7.03
Consumer Services		0.55	0.70	0.98	1.66	3.19
		Panel C: Model-derived trading trigger $1 + \alpha_1 - \gamma_1\alpha_2$				
Financials		0.63	0.80	0.88	0.93	0.96
Industrials		0.79	0.86	0.91	0.94	0.96
Consumer Goods		0.80	0.83	0.88	0.94	0.98
Consumer Services		0.80	0.86	0.91	0.93	0.96

This table presents the values of α_1 and α_2 obtained using the Johansen cointegration methodology in Panel A. The percentiles for α_1 are computed using the absolute values. Summary statistics of estimated values of γ_1 are reported in Panel B. Panel C presents the values of model-derived trading trigger $1 + \alpha_1 - \gamma_1\alpha_2$, which is computed using the resulting VECM estimates. As the Johansen estimation and the following trading activities is conducted on a rolling-window basis, these reported values are an average value computed from a series of numbers of each percentile. The sample period is January 2000 to February 6th 2017.

pairs trading triggers. While coefficient estimations are highly consistent with the baseline estimation, the results suggest lower dispersion in the cointegration coefficient and higher speed of equilibrium reversion due to the gain in commonality arising from firm-specific factor restrictions (note that simultaneous multi-factor restriction is not possible due to resulting low observations in each tercile portfolio).

Table 9 displays the strategy performance for each of the factor restrictions. Results show that while the restriction on size does not provide a clear improvement with respect to the baseline case, the book-to-market and volume restrictions clearly outperform the baseline. Specifically, Panel B shows that in all industry groups the mean portfolio return is substantially larger under the book-to-market restriction. Such increase in magnitude counteracts with the resulting increase of volatility level. Accordingly, Sharpe ratios are superior to the baseline results. Finally Panel C reports the effect of trading volume on performance. Except from Consumer Services industry case, all analyzed portfolios deliver stronger performance in terms of mean return. They also yield higher Sharpe ratios relative to the baseline cases. The results of each tercile are available upon request.

The overall result suggests that the imposed additional restriction based on fundamental factors, for the search of tradable pairs, is likely to boost mean reversion equilibrium and pairs trading performance. In our analysis, the superior results are delivered under the book-to-market ratio and trading volume restrictions. This evidence can be attributed to the increased commonality shared between paired assets. Our evidence therefore demonstrates that the imposed common factor restrictions strengthen cointegration and contributes to the performance of pairs trading strategies. Factor (as well as industry) commonality drives investors to exploit arbitrage between paired assets.

Table 1.8: VECM estimates and the model-derived trading trigger for pairs controlling trading volume

Sample		5th Percentile	25th Percentile	Median	75th Percentile	95th Percentile
		Panel A: Estimated values of α_1 and α_2				
Financials	α_1	-0.012	-0.036	-0.078	-0.152	-0.294
	α_2	0.001	0.004	0.012	0.029	0.077
Industrials	α_1	-0.021	-0.037	-0.063	-0.121	-0.252
	α_2	0.002	0.007	0.014	0.026	0.054
Consumer Goods	α_1	-0.030	-0.061	-0.091	-0.134	-0.189
	α_2	0.005	0.009	0.020	0.038	0.068
Consumer Services	α_1	-0.027	-0.039	-0.066	-0.090	-0.136
	α_2	0.006	0.008	0.015	0.022	0.036
		Panel B: Estimated values of γ_1				
Financials		0.14	0.48	1.48	6.84	9.84
Industrials		0.20	0.43	0.89	2.92	10.67
Consumer Goods		0.37	0.82	1.69	3.27	7.87
Consumer Services		0.35	0.48	0.82	1.64	3.06
		Panel C: Model-derived trading trigger $1+\alpha_1-\gamma_1\alpha_2$				
Financials		0.65	0.80	0.89	0.93	0.97
Industrials		0.81	0.87	0.91	0.93	0.96
Consumer Goods		0.79	0.83	0.88	0.92	0.95
Consumer Services		0.81	0.87	0.91	0.94	0.96

This table presents the values of α_1 and α_2 obtained using the Johansen cointegration methodology in Panel A. The percentiles for α_1 are computed using the absolute values. Summary statistics of estimated values of γ_1 are reported in Panel B. Panel C presents the values of model-derived trading trigger $1+\alpha_1-\gamma_1\alpha_2$, which is computed using the resulting VECM estimates. As the Johansen estimation and the following trading activities is conducted on a rolling-window basis, these reported values are an average value computed from a series of numbers of each percentile. The sample period is January 2000 to February 6th 2017.

Table 1.9: Summary statistics of excess returns to pairs portfolios controlling common factors

Sample	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: Market capitalization								
Financials	0.04	0.00	0.16	0.46	14.39	0.09	-0.06	0.26
Industrials	0.03	0.00	0.17	-0.16	11.66	0.07	-0.11	0.15
Consumer Goods	0.03	0.00	0.18	0.24	18.54	0.12	-0.13	0.18
Consumer Services	0.14	0.00	0.24	0.83	17.38	0.16	-0.13	0.56
Panel B: Book-to-market ratio								
Financials	0.11	0.00	0.17	0.70	13.07	0.09	-0.08	0.65
Industrials	0.08	0.00	0.19	0.68	16.11	0.12	-0.09	0.40
Consumer Goods	0.09	0.01	0.13	0.27	8.76	0.07	-0.07	0.69
Consumer Services	0.12	0.00	0.19	0.33	12.54	0.11	-0.09	0.63
Panel C: Trading volume								
Financials	0.10	0.00	0.17	0.68	14.58	0.09	-0.09	0.59
Industrials	0.06	0.00	0.17	0.30	11.10	0.09	-0.08	0.33
Consumer Goods	0.09	0.00	0.17	0.26	14.00	0.10	-0.10	0.55
Consumer Services	0.07	0.00	0.19	0.19	7.64	0.09	-0.06	0.40

This table presents descriptive statistics of excess returns for four industry groups, controlling three different fundamental factors. We trade according to the rule that opens a position in a pair one day after price spread diverges more than $(1+\alpha_1-\gamma_1\alpha_2)$ units of historical standard deviation. Reported are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return and (annualized) Sharpe ratio. The sample period is January 2000 to February 6th 2017.

1.5 Conclusion

In this paper we adapt the demand and supply framework introduced by FFG to show how equities that share common fundamentals can be combined to exploit pairs trading opportunities. Market clearing conditions are derived under a demand schedule including an arbitrage component and persistent cointegration errors. Equilibrium dynamics are represented via a VECM framework where long-run equilibrium convergence allows profits from pairs trading. Our model exploits price discovery so that the leading asset can be used to replicate the follower and shows how pairs trading profitability is linked to the speed of equilibrium reversion. Based on this presumption, our model derives a persistence dependent trading threshold used to trigger pairs trading strategies. In an out-of-sample exercise applied to STOXX Europe 600 equity price daily data, we show that: (a) price leadership is an important determinant of pairs trading profitability; (b) model-based pairs trading strategies yield positive Sharpe ratios that are higher than those obtained from competing pairs trading methodologies. Portfolio outperformance is enhanced under imposed firm fundamental factor as well as industry restrictions.

Chapter 2

A Matket Approach for Convergence Trades

2.1 Introduction

Traditional portfolio building needs to pick stocks from the cloud of possible alternatives in the market. Cointegration, has been used to identify closely related assets because it is believed that trading those assets will produce higher profits ceteris paribus. The cointegration-based portfolio literature establishes that portfolios of cointegrated assets are formed subject to some optimal investment weights, in order to maximize the expected portfolio value over a period of time, the investment horizon. This paper answers the following open questions in the literature: (a) What are the dynamics of such portfolios? (b) What positions need to be taken to maximize portfolio value? (c) How should the portfolio be managed? This paper addresses these questions by solving the dynamic maximization problem faced by an arbitrageur. We show that the maximum expected spread portfolio value over the time span (t, T) is equivalent to the price of call option on the spread of the two stocks.

Specifically we present a simple analytical model to describe the stationary convergence of the two non-stationary stocks. The underlying presumption is that if at least one of the mean-reversion parameters exists, and they have opposite signs, then the two stocks will be cointegrated, as their deviation will be stationary and of reduced rank. This is consistent with the discrete vector error correction model. In this sense, the model implies a simple framework to test for spread convergence, and therefore provides a straightforward mechanism to construct pairs. As the individual stocks deviate from the implied equilibrium level an arbitrageur will bet on the joint convergence by going short the overvalued asset and long the undervalued one. The long-short positions will eliminate the price discrepancy of the two assets from their joint equilibrium level. In the absence of a cointegration relation, the two asset prices in our model follow independent geometric Brownian motions that allow for different expected returns. In short, an arbitrageur chooses between two risky and possibly cointegrated stocks and a risk-free asset and allocates his wealth optimally. This guarantees a self-financing portfolio with zero initial value.

The model makes several empirical predictions. First it expects that cointegration-based portfolios yield higher returns compared to randomly selected portfolios. Second the model suggests that cointegration leads to bounded risk in the selected portfolios. While volatility exhibits time decay in cointegrated portfolios, non-cointegrated pairs will feature increasing volatility with the square root of time and are therefore expected to underperform. Third, the model suggests that when the speed of convergence of two paired assets is zero, the assets are not cointegrated. Using this parameter restriction we propose a simple pairs selection mechanism which demonstrates to outperform alternative pairs selection methodologies.

We associate the optimal asset allocation problem with option valuation. We find that the expected maximum value of the portfolio from t to T is equivalent to the value of a spread option with strike price equal to zero (see Margrabe (1978)). In consequence the optimal investment size is the corresponding delta hedging strategy. This paper therefore provides a market approach for valuing and managing portfolio strategies. Moreover this formulation allows for the extension of optimal strategy derivation to a multi-variable spread framework. Daily dynamic management and complete hedging is then possible using the Black–Scholes Greeks.

This article is related to studies that examine two risky but cointegrated assets and their impact on option pricing, e.g., Lo and Wang (1995), and Duan and Pliska (2004).[1] This literature focuses on the effect of cointegration on pricing arising from comovements in the asset prices. It starts from discrete cointegrated processes with GARCH(1,1) to construct the continuous counterparts. This involves pricing options on two assets and calculating sensitivities in a Black-Scholes fashion. We contribute to this literature by linking option pricing with the expected optimization of a portfolio on two cointegrated assets including an alternative asset, the risk-free bond. By linking portfolio optimization to option theory we extend the understanding of long-term comovement on two assets and the binding role of stationarity in portfolio optimization.

Our framework is also associated with the pairs trading literature of statistical arbitrage where the equilibrium notion of cointegration is exploited for trading purposes, as Gatev et al. (2006) and Figuerola-Ferretti et al. (2016). The former work selects pairs according to a canonical measure while the latter determines pairs according to Johansen (1991) in a price discovery model explicitly adapted to pairs trading. In this literature trading is typically triggered when price spread reaches a given level (see also Elliott et al. (2005) and Do et al. (2006)).[2] In a continuous setting, Jurek and Yang (2007) and Liu and Timmermann (2013) consider continuous processes to describe how an investor maximizes his wealth subject to preferences. Jurek and Yang (2007) employ an Ornstein-Uhlenbeck (O-U)[3] framework to account for horizon and divergence risks. Liu and Timmermann (2013) describe cointegrated pairs with a mean-reverting error and provide closed-form solutions under recurring arbitrage opportunities for stopped cointegrated processes.[4] The work of Lei and Xu (2015) is built on Liu and Timmermann (2013) to incorporate transaction costs.

[1] See also Stulz (1982) for the consideration of two related but not cointegrated assets.

[2] Another strand of the literature focuses on relative value trading based on technical analysis and on the company fundamentals, e.g. Brock et al. (1992), Chhaochharia and Grinstein (2007), Cumming et al. (2011). Mean-variance and uninformed investors, e.g. Marquering and Verbeek (2004), Campbell and Thompson (2008), also relate to this. They target the prediction of expected returns and use that knowledge to make profits. The literature is long and divided, but they agree that the out-of-sample performance of these prediction models is poor. In the longer term there is no stable relationship between the explanatory variables and the level of returns.

[3] Uhlenbeck and Ornstein (1930)

[4] Under stopped cointegration processes asset prices are cointegrated before the difference reaches zero. These processes can also be used to model strategies of convergence traders who close out their positions when prices converge.

Our framework presents cointegrated assets that are traded continuously, combined with a risk-free bond to set up a stationary portfolio. This implementation is in the spirit of Ankirchner et al. (2012) who point out that cross-hedging can be optimally achieved when the spread is stationary. We therefore construct a hedging portfolio[5] that minimizes risk, without sacrificing return following the spirit of Black and Scholes (1973). Our optimal strategy is employed in a hedging framework to pursue positive excess returns. In this context, optimal portfolio holdings are derived as closed-form solutions to the general form of value functions. These solutions reflect risk-return tradeoffs as underlined in the benchmark literature of mean-variance portfolio (see for example Marquering and Verbeek (2004)).

Our empirical application uses Dow Jones Industrial Average (DJIA thereafter) daily stock price data for the 1997-2015 period to test the proposed optimal trading strategy. This involves the use of portfolios weights that correspond to the deltas of the mapping spread option. We document remarkably consistent profits and impressive Sharpe ratios over the sample period. Our results also suggest that the proposed pairs selection algorithm supersedes the selection method based on Johansen (1991) methodology. In agreement with the literature, see Figuerola-Ferretti et al. (2016), we find that portfolio performance is stronger during the recent financial crisis. Moreover, portfolio profitability increases with the level of pairs' spread volatility and the return distribution is less dispersed when pairs with a high speed of convergence are considered. This can be explained by the error correction term $\lambda_i z_t$ $(i = 1, 2)$ interpreted as the excess return in the model. The intuition is that a pair candidate should include two well-matched assets, not only moving towards each other under stronger reverting force, measured by the speed, but also exhibiting an ample room of deviation in a volatile context, measured by the pricing error.

The remainder of the paper proceeds as follows. In Section 2 we introduce the investor's portfolio choice problem. Section 3 shows the connection between the derived optimal pairs portfolio and the spread option. In section 4 we conduct an empirical application of the optimal strategy to the constituents of DJIA. Section 5 concludes.

[5]Statistical arbitrage profits are subject to limiting arbitrage critique, but our portfolio is hedged, see Kondor (2009).

2.2 The continuous-time error correction

2.2.1 The setup

We assume there are three securities in the market, a risk-free bond and two risky assets, which can be traded in a friction-less continuous-time setting. The price of the risk-free bond is denoted by B_t, which provides a constant rate of return r. Its price dynamics are defined by:

$$dB_t = rB_t dt \tag{2.1}$$

The two risky assets are also tradable in the market. Their prices, y_t and x_t, follow the dynamic process:

$$\frac{dy_t}{y_t} = \mu_y dt + \lambda_1 z_t dt + \sigma_y dW_{y,t} \tag{2.2}$$

$$\frac{dx_t}{x_t} = \mu_x dt - \lambda_2 z_t dt + \sigma_x dW_{x,t} \tag{2.3}$$

$$z_t = ln y_t - ln x_t \tag{2.4}$$

where μ_y, μ_x, λ_1, λ_2, σ_y, and σ_x are constant parameters. $W_{y,t}$ and $W_{x,t}$ are standard Brownian motions with zero drift rate and unit variance rate. The above stochastic processes are specified under the physical measure P. We define z_t as price spread between the logarithms of the two asset prices, which displays the temporary mispricing. We refer to λ_1 and λ_2 as the speed of convergence of y_t and x_t, respectively. The higher the mean-reverting speed, the greater is the force that drives asset prices back to equilibrium. $\lambda_1 z_t$ and $-\lambda_2 z_t$ represent the error correction terms under the continuous VECM framework. This setup therefore includes two cointegrated variables with a unit cointegrating vector.

The above framework models price dynamics of two price series with mean-reverting characteristics. First, μ_y and μ_x are used to specify the expected rate of return per unit time in Equations (2.2) and (2.3). These are not defined to be equal, so they exhibit different market premiums arising from different portfolio betas. In this framework cointegration does not imply perfect comovement of paired assets, and thus no restriction on the beta is imposed. Our model incorporates the error correction term to describe a mean-reverting process that delivers abnormal returns via adjustment dynamics through convergence trade.

Consistent with Lei and Xu (2015), the restriction on positive mean-reverting speed implies stationarity of the error term. A model-implied test for cointegration therefore requires formally testing for this restriction. When $\lambda_1 = \lambda_2 = 0$ there is no equilibrium reversion and therefore no evidence of cointegration.

Additionally our model shows that speed of reversion parameters also affect pairs trading profitability. When $\lambda_1 = \lambda_2 = 0$ asset prices follow a geometric Brownian motion and thus price convergence cannot be exploited to gain arbitrage profits. By contrast, if λ_1 and λ_2 are non-zero, our model includes error correction terms that capture relative mispricing and produce mean reversion. Extra risk premium processes are incorporated for y_t and x_t in the drifts. As a result, the instantaneous expected rate of return can be explicitly defined as $\mu_y + \lambda_1 z_t$ and $\mu_x - \lambda_2 z_t$, respectively. The model therefore implies that positive return premium is created by stationary mispricings as Brennan and Wang (2010).

Based on the analysis above, we derive Proposition 1.

Proposition 1. *Equations* (2.2)-(2.4) *specify a dynamic VECM in continuous time with a stationary price spread* $z_t = lny_t - lnx_t$*, which can be represented under the measure* P *by the following process, under the assumption that the correlation between the two asset-return processes is constant and given by* $E[dW_{y,t}dW_{x,t}] = \rho_{xy}dt$, $\rho_{xy} \in (-1, 1)$.

$$dz_t = \alpha(\mu_z - z_t)dt + \sigma_z dW_{z,t} \tag{2.5}$$

with $\alpha = -(\lambda_1 + \lambda_2)$ *is the mean-reverting speed of* z_t *and satisfy* $\alpha > 0$*;* $\sigma_z = \sqrt{\sigma_y^2 - 2\rho_{xy}\sigma_y\sigma_x + \sigma_x^2}$*, and* $W_{z,t}$ *is a Brownian motion defined by* $W_{z,t} = (\sigma_y W_{y,t} - \sigma_x W_{x,t})/\sigma_z$*;* $\mu_z = \frac{1}{\alpha}(\mu_y - \frac{1}{2}\sigma_y^2 - \mu_x + \frac{1}{2}\sigma_x^2)$ *is the long-term equilibrium level.*

The dynamics of y_t and x_t determine the movement of price spread.[6] The process z_t is therefore a Gaussian process with mean-reverting force. The error correction terms, $\lambda_1 z_t$ and $-\lambda_2 z_t$, play an important role in determining the drift of price spread. It is the error correction term that drives mean reversion in y_t and x_t in order to ensure the stationarity of z_t.

When $\alpha > 0$ the process is strictly stationary[7] and the variance has a time decay. That is as for big t, the process converges to $\sigma_z^2/2\alpha$, unlike the Geometric Brownian motion that has an unbounded variance (it grows infinitely). The O-U process has a normal density function:

$$p\left(z_t = z; t; z_{t_0} = z_0, t_0\right) = \frac{1}{\sqrt{2\pi s^2\left(t\right)}} e^{-\frac{(z-m(t))^2}{2s^2(t)}} \tag{2.6}$$

where $m\left(t\right) = \mu_z + \left(z_0 - \mu_z\right) e^{-\alpha(t-t_0)}$ is the mean of the process, and the variance is $s^2\left(t\right) = \frac{\sigma_z^2}{2\alpha}\left[1 - e^{-2\alpha(t-t_0)}\right]$.

2.2.2 The optimal portfolio

Suppose there is an investor who maximizes the value of his portfolio, composed by three securities in the economy. Let φ_1, φ_2 and φ_3 be, respectively, the number of shares held on risky assets y_t and x_t, and the number of risk-free bond in the portfolio. The value of that portfolio, Π_t, is then represented by:

$$\Pi_t = \varphi_1 y_t + \varphi_2 x_t + \varphi_3 B_t \tag{2.7}$$

We suppose that the investor maximizes the expected portfolio value. Then the value function that solves this maximization problem, $V\left(t, \Pi, y, x\right)$, is defined as:

$$V\left(t, \Pi, y, x\right) = \underset{\varphi_1,\varphi_2}{max} E_t\left[\Pi_T\right] \; where \, t \in [0, T] \tag{2.8}$$

where Π_T is the wealth at time T, corresponding to the optimal trading strategy with the initial wealth $\Pi_t = \Pi$, and asset prices $y_t = y$ and $x_t = x$ at time t.

[6] Note that this specification excludes negative prices so that the lognormality assumption is satisfied.
[7] If $\alpha > 0$ then also, $t \to \infty$ or $T - t \to \infty$.

The optimal investment policies are obtained and summarized as Proposition 2. The proof of this proposition is left for the Appendix B.

Proposition 2. *When two asset-return processes are correlated and their correlation is given by* $E[dW_y dW_x] = \rho_{xy}dt$, $\rho_{xy} \in (-1, 1)$[8], *the optimal investment policies,* $(\varphi_1^*, \varphi_2^*)$, *on the individual risky assets following a continuous VECM* (2.2)-(2.4), *under the physical measure* P, *and for every* t *are given by :*

$$\varphi_1^* = -\frac{V_\Pi}{V_{\Pi\Pi}} \cdot \frac{\theta_y - \rho_{xy}\theta_x}{\sigma_y y(1-\rho_{xy}^2)} - \frac{V_{\Pi y}}{V_{\Pi\Pi}} \tag{2.9}$$

$$\varphi_2^* = -\frac{V_\Pi}{V_{\Pi\Pi}} \cdot \frac{\theta_x - \rho_{xy}\theta_y}{\sigma_x x(1-\rho_{xy}^2)} - \frac{V_{\Pi x}}{V_{\Pi\Pi}} \tag{2.10}$$

where $\theta_y = \frac{(\mu_y - r) + \lambda_1 z_t}{\sigma_y}$ *and* $\theta_x = \frac{(\mu_x - r) - \lambda_2 z_t}{\sigma_x}$ *are the excess returns for* y_t *and* x_t*, respectively.*

The optimal portfolio holdings, given by Equations (2.9)-(2.10), have a simple structure and can be decomposed into two components, as specified in the literature applying utility-based objective functions (see for instance Jurek and Yang (2007), Basak and Chabakauri (2010) and Kondor and Vayanos (2014)). The first component comes from the myopic demand, whereas the second one represents the intertemporal hedging demand. It can be seen that the optimal investment size is jointly determined by both assets' risk premiums θ_y and θ_x, which are linked through the correlation parameter ρ_{xy}. An increase in correlation leads to greater investment size suggesting a positive link between them. If we ignore the existence of correlated BMs the solution leads to underinvestment on paired assets (in absolute terms). In this sense the importance of return correlation is highlighted. Moreover, impacts of asset volatility or interest rate are conveyed through the risk premium and exerted on the investor's behavior. The higher the asset volatility or bond interest rate treated as funding cost, the lower the investment allocation that the investor will assign to a given asset.

[8]The correlation parameter ρ_{xy} is either positive or negative, given that imperfectly integrated markets are considered in this study. Perfect correlation are thus excluded, in which special cases suggest that two paired assets provide identical risk premium in absolute term.

We solved a more general portfolio choice problem allowing for different asset allocations for each asset and thus improve on the studies with equal weight setups, $[-1, 1]$, see Liu and Longstaff (2004), Gatev et al. (2006), and Jurek and Yang (2007). Our weights are required to be of opposite signs if stationarity is to be guaranteed. Optimal portfolio allocations incorporate the funding cost and are of different sizes, in line with the mean-variance portfolio literature (see Marquering and Verbeek (2004) and Basak and Chabakauri (2010)).

Given the proposed solution to the dynamic asset allocation problem, Corollary 1 following Proposition 2 is derived to determine expected abnormal returns.

Corollary 1. *Using the optimal investment policies φ_1^* and φ_2^* derived in Proposition 2, we obtain the following process f_t representing the expected abnormal returns under the measure P:*

$$
\begin{aligned}
f(y_t, x_t, t) &= \varphi_1^* y_t \left(\mu_y - r + \lambda_1 z_t\right) + \varphi_2^* x_t \left(\mu_x - r - \lambda_2 z_t\right) \\
&= \varphi_1^* y_t \sigma_y \theta_y + \varphi_2^* x_t \sigma_x \theta_x \qquad (2.11)
\end{aligned}
$$

Consequently, abnormal returns from pairs trading are expected to be higher when paired assets are linked through cointegration.

Corollary 1 demonstrates that pairs trading deliver higher excess returns under the existence of cointegration relationship. Note that, in addition to the expected return $\mu - r$ implied by asset pricing models, an extra excess return would be earned by exploiting price convergence driven by cointegration, until relative mispricing is eliminated. This is possible when price deviation exists between cointegrated assets, such that λ_1 and λ_2 are nonzero. When $\lambda_1 z_t > 0$, y_t will be relatively underpriced and an upward movement is expected to restore equilibrium. In this case, $\lambda_2 z_t < 0$ is expected as x_t is overpriced, implying that a downward movement is expected in response to spread mean reversion. These mean-reverting activities can be exploited via pairs trading. Therefore $\lambda_1 z_t$ and $-\lambda_2 z_t$ in Equation (2.11), represent extra risk premiums originated from relative mispricing between two paired assets. This is consistent with Brennan and Wang (2010), showing that return premium induced by the mispricing can be expected to be positive if the pricing error z_t follows a stationary process. Conversely, if price adjustment occurs immediately, there is no mispricing in either asset. In this sense $\lambda_1 = \lambda_2 = 0$ and there is no cointegration. Therefore, as is presented in the proposed framework, additional profits are gained if relative mispricing exists and prices eventually revert back to the long-term level.

2.3 Optimal portfolio selection and exercise strategy of spread options

In what follows we solve the dynamic asset allocation problem under a well-defined functional form of V. This involves deriving the candidate value function and obtaining the optimal investment strategies. Prior portfolio optimization literature derives the arbitrageur´s optimal dynamic investment strategy for a set of preference specifications. A commonly used form is the constant-relative-risk-aversion (CRRA) utility function (see for instance Merton (1971), Kim and Omberg (1996), Brennan and Xia (2002), and Liu (2007) etc.). This preference allows for well-defined tractable solutions but does also have important limitations. Suen (2009) argue that this function is not bounded, implying that the difficulties may be faced when applying the standard tools of dynamic programming. In this paper we aim to identify an appropriate value function not associated with utility preference. To do so, we first prove that our portfolio is a martingale (see Appendix C). We then link the optimal portfolio with spread options, as both the proposed portfolio and the underlying asset of spread options are designed to be stationary. The optimization problem is solved by identifying a spread option written on underlying cointegrated assets, which serves as the replicating instrument of the proposed portfolio. This allows the use of market data and compute the optimal investment allocations.

2.3.1 The interpretation of portfolio strategy using spread options

A spread option in this framework is written on paired assets which are linked in the long run via cointegration. In this section we link the wealth process Π_t under the optimal policies with European-type spread options. In what follows, we show that the optimal investment holdings can be associated with the spread option deltas of the involved individual assets. This departs from the previous literature which considers general portfolio choice problems with cointegrated assets (see for example Liu and Timmermann (2013)) by introducing forward-looking information embedded in option prices. We offer simple and tractable solutions using information from the OTC option market.

The portfolio Π_t is self-financing if

$$d\Pi_t = \varphi_1 dy_t + \varphi_2 dx_t + r(\Pi_t - \varphi_1 y_t - \varphi_2 x_t)dt$$

Let $C(y_t, x_t, t)$ be the option value. We use Ito to produce the option´s dynamics as:

$$dC(y_t, x_t, t) = \left[\frac{\partial C}{\partial t} + \frac{1}{2}\sigma_y^2 y_t^2 \frac{\partial^2 C}{\partial y^2} + \rho \sigma_y y_t \sigma_x x_t \frac{\partial^2 C}{\partial y \partial x} + \frac{1}{2}\sigma_x^2 x_t^2 \frac{\partial^2 C}{\partial x^2}\right] dt + \frac{\partial C}{\partial y} dy_t + \frac{\partial C}{\partial x} dx_t \tag{2.12}$$

We now build a stationary portfolio, Π_t using the hedging condition which requires that we go short on the spread portfolio and long the option written on the spread, $C(y_t, x_t, t)$. Differentiating with respect to the two underlying assets we obtain $\varphi_{1t} = \frac{\partial}{\partial y} C(y_t, x_t, t)$ and $\varphi_{2t} = \frac{\partial}{\partial x} C(y_t, x_t, t)$. Then we have a risk-free portfolio:

$$d(C(y_t, x_t, t) - \Pi_t) \tag{2.13}$$

$$d(C(y_t, x_t, t) - \Pi_t) = \left[\frac{\partial C}{\partial t} + \frac{1}{2}\sigma_y^2 y_t^2 \frac{\partial^2 C}{\partial y^2} + \rho_{xy} \sigma_y y_t \sigma_x x_t \frac{\partial^2 C}{\partial y \partial x} + \frac{1}{2}\sigma_x^2 x_t^2 \frac{\partial^2 C}{\partial x^2} - r(\Pi_t - \varphi_{1t} y_t - \varphi_{2t} x_t)\right] dt \tag{2.14}$$

Then under perfect replication, the hedging condition will earn the risk-free rate of return:

$$d(C(y_t, x_t, t) - \Pi_t) = r(C(y_t, x_t, t) - \Pi_t)dt \tag{2.15}$$

With the dynamics of the option and the self-financing portfolio, we then derive the Black-Scholes PDE

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma_y^2 y_t^2 \frac{\partial^2 C}{\partial y^2} + \rho\sigma_y y_t \sigma_x x_t \frac{\partial^2 C}{\partial y \partial x} + \frac{1}{2}\sigma_x^2 x_t^2 \frac{\partial^2 C}{\partial x^2} + \frac{\partial C}{\partial y} r y_t + \frac{\partial C}{\partial x} r x_t - rC = 0 \tag{2.16}$$

where $C(y_T, x_T, T) = \psi(y_T, x_T)$. $\psi(y_T, x_T)$ represents the payoff function for a general European option.

If $\Pi = C(y, x, 0)$, then $d(C(y_t, x_t, t) - \Pi_t) = 0$. And thus $\Pi_T = C(y_T, x_T, T) = \psi(y_T, x_T)$.

Setting up a convergence trade requiring a long position in one asset and short in another asset, will provide an equivalent payoff to that offered by spread options.

Suppose we have an option written on $y_t - x_t$. This option can be considered as a spread option that gives the holder the right to exchange one asset for another. In this framework we take spread option with zero strike.

The payoff function for this option at maturity T is $\psi(y_T, x_T) = max(y_T - x_T, 0)$, so the spread option can be defined as (see Appendix D):

$$C(y_t, x_t, t) = y_t\Phi(d_1) - x_t\Phi(d_2) \tag{2.17}$$

where

$$d_1 = \frac{ln(\frac{y_t}{x_t}) + \frac{1}{2}\sigma_z^2(T-t)}{\sigma_z\sqrt{T-t}}, \;\; d_2 = d_1 - \sigma_z\sqrt{T-t}$$
$$\sigma_z = \sqrt{\sigma_y^2 - 2\rho\sigma_y\sigma_x + \sigma_x^2}$$

This is the Margrabe (1978)´s result. In order to hedge the option position, the writer needs to pursue delta hedging. This requires holding Δ_t shares of the underlying asset at time t. It is simple to demonstrate that the hedging strategy for the two-asset spread option is:

$$\begin{aligned} \Delta_{y,t} &= \frac{\partial C}{\partial y} = \Phi(d_1) + \left(\phi(d_1) - \frac{x_t}{y_t}\phi(d_2)\right)\frac{1}{\sigma_z\sqrt{T-t}} \\ &= \Phi(d_1) \end{aligned} \tag{2.18}$$

$$\begin{aligned} \Delta_{x,t} &= \frac{\partial C}{\partial x} = -\Phi(d_2) - \frac{y_t}{x_t}\left(\phi(d_1) - \frac{x_t}{y_t}\phi(d_2)\right)\frac{1}{\sigma_z\sqrt{T-t}} \\ &= -\Phi(d_2) \end{aligned} \tag{2.19}$$

In this way we transform the optimization problem of the investment strategy under P, to the risk-neutral measure $\mathbb{Q}$ by doing valuation of spread options and its resulting delta hedging strategies. This requires spread option valuation and the use of market data on this derivative. The analysis above is summarized in Proposition 3.

Proposition 3. *Consider an investor who invests in a portfolio with initial value Π_t, containing two cointegrated assets y_t and x_t and a risk-free bond with constant return r. The investor tries to maximize the portfolio value at terminal time T. According to the principle of perfect replication, the portfolio value is equivalent to purchasing a spread option on $y_t - x_t$ with zero strike price. Consequently, the optimal investment policies are equivalent to the hedging strategies of the spread option, namely,*

$$\varphi_1^* = \Delta_y, \quad \varphi_2^* = \Delta_x \tag{2.20}$$

2.3.2 The optimal value function

In the Appendix B we show that the value function V satisfies the Hamilton-Jacobi-Bellman (HJB) equations of stochastic control theory. The next step is to obtain an appropriate functional form of V. In the previous section, we show that the portfolio value is equivalent to the payoff function of a spread option written on two cointegrated assets. The Black-Scholes price is the maximum discounted payoff under no arbitrage, the option formula accordingly serves as an appropriate candidate function which optimizes the portfolio value.

Proposition 4. *For $0 \leq t \leq T$, let $V(t,\Pi,y,x) = max\, E_t\,[\Pi_T \geq 0]$ we have the optimal value function written as:*

$$\begin{aligned} V(t,\Pi,y,x) &= y_t\Phi(d_1) - x_t\Phi(d_2) \\ &= \Delta_{y,t}\cdot y_t + \Delta_{x,t}\cdot x_t \end{aligned} \tag{2.21}$$

Consequently, it can be affirmed that for this function $V(t,\Pi,y,x)$, we get

$$V_\Pi = \Delta_{y,t} + \Delta_{x,t} = \Phi(d_1) - \Phi(d_2)$$

$$V_{\Pi\Pi} = \Gamma_y - \Gamma_x = \frac{\phi(d_1)}{y_t\sigma_z\sqrt{T-t}} - \frac{\phi(d_2)}{x_t\sigma_z\sqrt{T-t}}$$

$$V_{\Pi y} = \frac{\phi(d_1)-\phi(d_2)}{y_t\sigma_z\sqrt{T-t}}, \quad V_{\Pi x} = \frac{-\phi(d_1)+\phi(d_2)}{x_t\sigma_z\sqrt{T-t}}$$

The implication of Proposition 3 and 4 is that we associate the value of the investor´s portfolio, made of two cointegrated assets and a risk-free bond, with the spread option. This allows us to determine the optimal investment policies in terms of option deltas. We therefore contribute to the portfolio selection literature by introducing forward-looking information from option markets to solve the portfolio optimization problem. Previous studies in the literature use in-sample observations to determine optimal weights. However a more recent line of work (see Bailey and Lopez de Prado (2014), Bailey et al. (2014) and Harvey and Liu (2014)) demonstrates that calibrating trading practice via backtest delivers overfitting estimates, and thereby underperformance. We shed light to this problem by using market data to calibrate portfolio solutions. We therefore provide a market-based approach to the portfolio selection problem that can be extended to sophisticated portfolio composition beyond the spread trading. As opposed to the extant literature our proposed optimal investment strategies arise from a generic value function solution that does not require assumptions regarding investor´s preferences. Under our framework, the Black-Scholes spread option specification is considered as an appropriate value function.

2.4 Empirical study

2.4.1 Data and methodology for performance evaluation

In what follows we test our model using data on U.S. equity market. We collect from Datastream daily data on prices for the companies included in the DJIA during the period January 1997 to December 2015. This index aims to track 30 leading blue-chip companies based in the US, hence summarizing the performance of the industrial sector.[9]

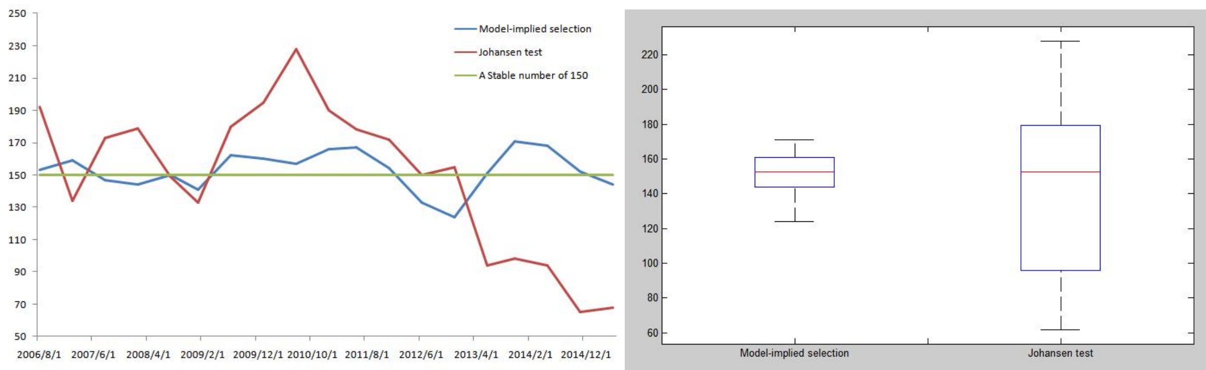
Our options data are collected from IvyDB (OptionMetrics), which provides data on all index and stock options listed in the U.S. market. The data on options ends on August 31st, 2015. We use the volatility surface file, which delivers a smoothed volatility surface for a variety of standard maturities and option deltas reflecting a set of strike prices. We choose the at-the-money implied volatilities for calls and puts written on all individual stocks included in the DJIA for a 6-month maturity. Then the at-the-money Black-Scholes volatility is calculated as the average volatility for a call and a put with absolute delta level equal to 0.5.[10]

The first step towards implementing the proposed trading strategy is to estimate the parameters under the continuous cointegration model (2.2)-(2.3), $\Theta = (\hat{\mu_y}, \hat{\mu_x}, \hat{\lambda_1}, \hat{\lambda_2}, \hat{\sigma_y}, \hat{\sigma_x}, \hat{\rho_{xy}})$. These are estimated applying maximum likelihood methods (ML thereafter) following the spirit of Lei and Xu (2015). The derivation procedure is present in detail in Appendix E. Our analysis depends on a rolling-sample approach. We choose an estimation window of $t = 10$ years length and then the next 6-month window as the trading period (this ranges from $T = t$ to $T = t + 0.5$). A 10-year data frame is therefore used to predict the parameters needed for the implementation of our trading strategy. This process is moved forward by adding price data for the following 6 months and dropping the earliest 126 observations. The process is repeated until the end of the dataset is reached. Point estimates are used for portfolio construction for every sequence of 6-month windows. The first estimation window is from January 1997 to December 2006, which is followed by the first trading period from

[9]The company, Visa Inc., has been excluded from the sample because of a shorter period of price data which are only available since March 19th, 2008.

[10]Data on stock options are available until August 31st, 2015, such that the empirical study is terminated on that date.

Figure 2.1: Number of cointegrated pairs: Model-implied selection criteria vs. Johansen (1991) method



This figure plots, in the left panel, the number of cointegrated pairs for each trading period, identified according to the model-implied criteria and the Johansen (1991) test over the period January 1st, 2007-August 31st, 2015. The right panel is the boxplot of these number showing their distributions.

January 2007 to June 2007. We therefore have nineteen trading periods for the out-of-sample exercise, over the period January 1st, 2007-August 31st, 2015.

We evaluate the performance of portfolio strategies under three criteria: (1) average excess return; (2) return volatility (standard deviation); (3) portfolio Sharpe ratio. The latter is used to assess the risk-adjusted performance.

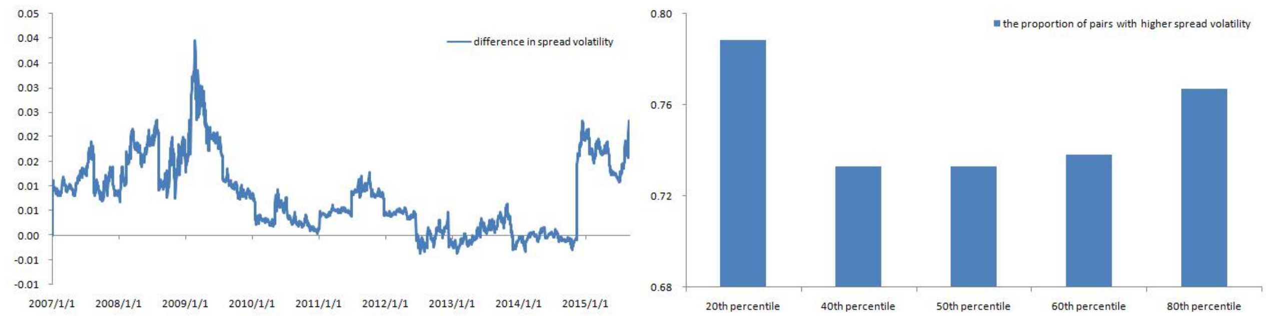
2.4.2 Pairs formation and comparison

Next step would be picking pairs that will constitute the trading portfolios. There are 406 possible combinations among 29 targeted firms. The pair identification employs the model-implied criteria, based on the restriction of convergent speed $\alpha > 0$, that renders stationary z_t. Every pair in the portfolio is mean-reverting.[11] To be robust, we compare the performance of the model-implied method with the Johansen (1991) cointegration method in both pairs selection and trading profitability.

Figure 1 plots the number of cointegrated pairs (left panel) identified under the Johansen and the model-implied methods for each 6-month trading period between January 2007 and August 2015. We see that while the number is time-varying, the number of pairs

[11] It implies that λ_1 and λ_2, exhibit opposite signs (see Equations (2.2)-(2.3)).

Figure 2.2: Spread volatility comparison at the level of individual pairs



This figure plots, in the left panel, the time series of the difference in the average of spread volatility between the model-implied pairs and Johansen (1991)'s pairs. The right panel shows, at different percentiles, the proportion of the model-implied pairs with higher spread volatility, relative to Johansen (1991)'s pairs, over the period January 1st, 2007-August 31st, 2015.

selected under the continuous model is not as volatile as the number selected by Johansen methodology.[12] The boxplot (right panel) supports this observation. Both methods screen out, on average, around 150 pairs of assets over the entire trading period. However the distribution of number of pairs expands under the Johansen method, as reflected by a lower first quartile and a higher third quartile. This evidence is also observed in terms of the minimum and maximum values. This is an extra risk in the portfolio selected with the traditional statistical approach that affects performance and possibly explains why Kondor (2009) claims that those portfolios couldn't survive many equilibriums. In the perspective of practitioners, a stable number of tradable pairs is favorable given no wide adjustment of portfolio components period by period.

Considering the fact that different pairs are traded under the two methods, it is of great interest to analyze the difference between them by focusing on spread volatility. We examine the volatility gap at the individual level. The left panel of Figure 2 shows that the mean volatility of all tradable pairs is higher under the model-implied method in most of the time. The value of mean difference peaks in crisis times, and then drops to a low level and even becomes zero. Overall, the magnitude of volatility gap is limited to a small range.

[12] We apply Johansen (1991) cointegration test to the natural logarithm of price series in search of pairs at the 10% significant level over a 10-year period. The 10-year window moves forward for 6-month so as to conduct a rolling-sample test, until the end of the dataset.

The right panel shows the proportion of pairs exhibiting relatively higher spread volatility under our selection method. The finding is consistent at five different percentiles, revealing a proportion of over 70%.

2.4.3 Pairs portfolio strategies using option deltas as investment policies

In this section we look into the trading process. Trading will exploit the joint convergence by going long the underpriced asset and shorting the overpriced one in the pair, according to the deltas of the corresponding option on the spread. These options are traded in OTC markets where we get the individual implied volatilities. Correlation is estimated using maximum likelihood. The implied volatility of the spread is then calculated using Equation (2.17). All these are applied to pairs under both model-implied and Johansen methods. Then two equally weighted portfolios are constructed.

Figure 3 shows the periodic performance out of sample. We observe that both portfolios deliver synchronous performance. In fact the correlation coefficients of average return, return volatility and Sharpe ratio for both portfolios are 0.89, 0.93, and 0.90, respectively. More importantly, while both portfolios deliver positive abnormal profits and Sharpe ratios, our portfolio performs better in most of periods. In agreement with previous literature (see Figuerola-Ferretti et al. (2016)) portfolio outperformance is maximized over the recent crisis, June 2008-June 2009. An increase of portfolio volatility is accompanied as expected, reaching values of 10% per annum before falling back to pre-crisis mean volatility level. This is implied by the high level of volatility embedded on option prices. Our portfolio also experiences fewer losses when the market conditions are less favorable for the use of pairs strategy.[13]We therefore offer a novel approach for pairs selection that is closely linked but outperforms the Johansen selection method.

To understand the economic significance, Figure 4 plots the cumulative profits by sum-

[13]The negative returns are small, and documented in five out of nineteen trading periods. At the beginning of global financial crisis, the first half of 2008, occurs the first one, while the rest are distributed in after-crisis period.

Figure 2.3: The evolution of performance: Model-implied portfolio vs. Johansen's portfolio

This figure describes the performance of two pairs portfolios in terms of three criterias, for each trading period, over the period January 1st, 2007-August 31st, 2015. It is shown that both portfolios deliver very close performance, whose correlation coefficients are 0.89, 0.93 and 0.90 based on the mean return, volatility and Sharpe ratio.

Figure 2.4: Cumulative excess returns of pairs portfolios

This figure plots cumulative excess return of our model-implied and Johansen's portfolios over the period January 1st, 2007-August 31st, 2015.

ming up the gains on an equal-weighted portfolio of all open positions. Both portfolios accumulate returns in a similar evolution path. Two portfolios earn cumulative returns of similar magnitudes over the first two years, but then outstanding performance is observed from our portfolio during the latest crisis, particularly in 2009. After that this performance disparity is maintained and cumulative profits of both portfolios increase consistently until the end of sample. Our portfolio accumulates profits of 0.41 against 0.32 earned by the benchmark Johansen (1991) portfolio. In addition to the outperformance, the evidence so far supports the choice of spread option deltas as optimal weights in the cointegrated assets.

Table 1 reports descriptive statistics for both out-of-sample portfolios. In particular, our portfolio earns an average excess return of 4.54% per year from January 2007 until August 2015. This is about 1% more per annum for the period, relative to the benchmark Johansen (1991) portfolio. Big variation in the number of constituents included in Johansen's portfolio impacts on the profit & loss statement (PnL), thus increasing its return volatility resulting in a Sharpe ratio 0.23 lower than our portfolio. The reported average return and Sharpe ratio suggest the outperformance of our model-based selection criteria. Moreover, we consider a passive strategy giving equal weights to a pair of assets, as another relevant benchmark for our portfolio subject to daily rebalancing. This is commonly applied to the studies of

Table 2.1: Out-of-sample performance of pairs portfolios

Pairs portfolios	Mean	Median	Stdev	Skew	Kurtosis	Max	Min	Sharpe	Cumulative returns
Model-implied portfolio	0.0454	0.0000	0.0522	0.41	9.42	0.03	-0.03	0.87	0.41
Johansen's portfolio	0.0353	0.0000	0.0552	0.72	11.41	0.03	-0.02	0.64	0.32
Delta-neutral portfolio	0.0287	0.0000	0.0360	0.35	8.98	0.02	-0.02	0.80	0.26

This table reports the performance of our portfolio formed with selected pairs under the model-implied method, and a portfolio formed with pairs under the Johansen (1991) test. A delta-neutral portfolio is also constructed using pairs chosen by the model-implied method. Reported are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return, the (annualized) Sharpe ratio, and cumulative returns. The sample period is January 1st, 2007-August 31st, 2015.

convergence trades and is thought to be delta neutral, for instance Liu and Longstaff (2004), Gatev et al. (2006), Jurek and Yang (2007). This seems intuitively reasonable but may not be optimal. Note that this exercise is conducted on pairs identified by our model-implied selection method. The results of this delta-neutral strategy are reported in the last row of Table 1. By comparison, we find that both pairs portfolios produce a higher average return, particularly the mean return of our model-implied portfolio exceeding by about 1.7% per annum. Not surprisingly the delta-neutral strategy delivers a lower return volatility given no requirement of rebalancing; however our proposed model-implied portfolio yields the highest Sharpe ratio. In sum, conventional delta-neutral strategy seems suboptimal relative to our dynamic portfolio holdings modeled by spread option´s deltas.

2.4.4 Information flow and portfolio performance

Our portfolio strategy exploits forward-looking information that are implied by current option prices. From the perspective of an investor, it is critical to understand the relation between the degree of information flow and the performance of an option-based portfolio. The informational role of options is supported by a large body of literature given nonpublic information brought to the option market by informed traders. As has been also found, the adjustment of stock prices is not achieved immediately to reflect information embedded in the option prices, indicating the existence of information asymmetry. One can expect the level of information asymmetry is higher in times of crisis compared to tranquil times. This would make option prices more informative, indicating a pronounced advantage of option-

Table 2.2: Out-of-sample performance of pairs portfolios in periods of crisis and no crisis

Pairs portfolios	Mean	Median	Stdev	Skew	Kurtosis	Max	Min	Sharpe	Cumulative returns
Panel A: January 1st, 2007-December 31st, 2009									
Model-implied portfolio	0.0820	0.0000	0.0697	0.47	7.56	0.03	-0.03	1.18	0.26
Johansen's portfolio	0.0535	0.0000	0.0709	0.87	9.59	0.03	-0.02	0.75	0.17
Delta-neutral portfolio	0.0521	0.0000	0.0481	0.37	7.16	0.02	-0.02	1.08	0.16
Panel B: January 1st, 2010-August 31st, 2015									
Model-implied portfolio	0.0259	0.0000	0.0399	-0.12	4.59	0.01	-0.01	0.65	0.15
Johansen's portfolio	0.0256	0.0000	0.0447	0.17	8.28	0.02	-0.02	0.57	0.15
Delta-neutral portfolio	0.0162	0.0000	0.0275	-0.03	4.63	0.01	-0.01	0.59	0.09

This table shows the out-of-sample performance for the three different pairs portfolios as present in Table 1. The crisis period is from January 2007 to December 2009, while the non-crisis period is between January 2010 and August 2015.

based portfolio relying on current market data. Therefore, it is reasonable to conceive our option-based strategy to perform remarkably well in the latest crisis (2007-2009).

Our findings support this expectation. Panel A of Table 2 shows that our portfolio yields, on average, an excess return of 8.20% per annum during the global financial crisis, more than three times higher than the return earned over the post-crisis period 2010-August 2015. This finding suggests that pairs portfolio is a low-beta portfolio given its long-short position on two close-related assets. The boost of return volatility is accompanied as expected, but the magnitude is not as large as the increase of profitability achieved under distressed conditions. We therefore find a superior Sharpe ratio for the turbulent period, and the difference is substantial which is 1.18 relative to 0.65. Then we look into the performance of delta-neutral portfolio, which is formed with identical pairs as our model-implied portfolio. It produces, on average, an annualized return of 5.2% during crisis in contrast to 1.6% in tranquil times. The same observation is documented in terms of Sharpe ratio, which is 1.08 against 0.59. Moreover, the comparison between our portfolio and the delta-neutral portfolio is interesting. We find that the magnitude of return disparity is larger in the crisis period, whose value is up to 3% per annum against 1% in the non-crisis era. The substantial difference should be attributed to the portfolio compositions based on option prices including richer information in crisis times.

The superior performance in crisis is also achieved under the Johansen (1991) method, although the performance gap is not so large. Some interesting evidence are found after comparing Johansen´s portfolio with ours. Our portfolio is found to produce stronger performance in the bear market while both portfolios' performance is close in the post-2009 sample period. The resulting outperformance mainly results from the profit-generation ability, represented by the mean return. This non-trivial finding uncover that our pairs strategy may be more able to screen out pairs embedded with strong mean-reverting characteristics so as to discover and capture trading opportunities.

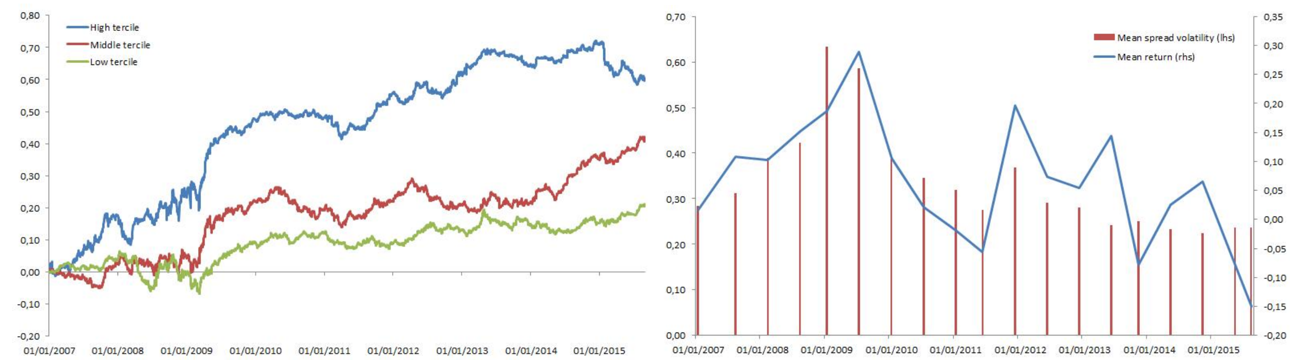
2.4.5 The sensitivity of portfolio performance

In this section we explore the sensitivity of portfolio performance with respect to two key parameters. We concentrate on spread volatility, as a measure of idiosyncratic risks, and the speed of equilibrium convergence, as a condition of eliminating deviations.

As spread volatility is a time-series, we take the average volatility of each selected pair for every trading period as the proxy of volatility level. These pairs are then sorted on the mean spread volatility and divided into three groups corresponding to the low, middle, and high terciles of volatility. Pairs in each tercile are used to form an equally weighted portfolio. The left panel of Figure 5 shows that the high-volatility group, whose final return reaches at 0.60, apparently dominates the other two portfolios providing respective total return of 0.42 and 0.21. This is in line with the principle, more risk, more return. Specifically, the summary statistics for the three portfolios reveals that, the resulting high return volatility for the top tercile (8.26%), relative to the middle tercile (6.97%) and bottom tercile (5.33%), is sufficiently compensated by the resulting mean return, thus leading to a better Sharpe ratio of 0.82 (compared to 0.67 and 0.44, respectively). In addition, the top tercile portfolio also delivers greater maximum and minimum returns. These evidence thus indicates a greater profitability of pairs with a high level of spread volatility.

Further, we explore how the portfolio return varies as the evolution of spread volatility.

Figure 2.5: The sensitivity of portfolio performance to spread volatility

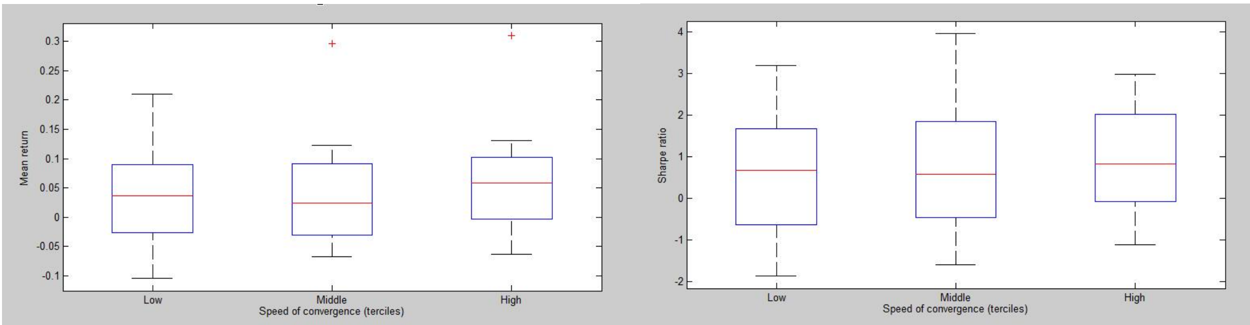


The left panel of the figure describes the cumulative return for three portfolios with pairs corresponding to the high, middle, and low terciles of spread volatility. The right panel displays the mean return of top tercile portfolio against the evolution of mean spread volatility. The sample period is January 1st, 2007-August 31st, 2015.

We take the top tercile portfolio for example. The right panel of Figure 5 displays a strong comovement between the average spread volatility and the portfolio' mean return. In times of high volatility, such as the recent global financial crises, the mean return reach the maximum level while in low-volatility conditions the portfolio exhibits less impressive performance. One possible explanation for this evidence is the existence of higher pricing errors during abnormal market conditions.

Then we look into the speed of convergence. Figure 6 displays boxplots of the mean returns and Sharpe ratios conditioning on the low, middle, and high terciles of convergent speed. It is observed that not only the median, but also the first and third quartiles are higher for the top tercile of pairs, in terms of both evaluation metrics. The distribution of mean return and Sharpe ratio shrinks as pairs with a higher speed are considered, indicating a less dispersed return distribution avoiding jumps/drops in profitability. In contrast, pairs at the bottom tercile are more likely to induce extreme returns in the slow reversion process, and thereby possibly lead to a big tail risk. It seems that the speed of convergence is associated with the portfolio´s performance stability.

Figure 2.6: Performance distribution of pairs portfolios at terciles of convergent speed



This figure describes mean return (left panel) and Sharpe ratio (right panel) distributions for the high, middle, and low terciles of convergent speed at the portfolio level, over the period January 1st, 2007-August 31st, 2015.

2.5 Concluding remarks

In this paper we propose a simple analytical model to describe the stationary convergent process in the continuous setting. This model implies a simple method to check for cointegration based on the speed of equilibrium reversion. This novel application is useful for pairs selection. The model also suggests that cointegration-based portfolio generates higher returns and bounded risk relative to a portfolio formed with randomly selected assets. Our paper contributes to the pairs portfolio and spread option literature by exploiting a continuous model of cointegration to derive optimal portfolio solutions. In doing this, we formulate optimal weights that correspond to the delta strategies of spread options. We depart from the current literature by deriving closed-form solutions without the assumption of a utility function for the portfolio optimization problem. The proposed strategy can easily be implemented retrieving forward-looking option market information. In this sense, we extend the understanding of long-term comovement on two assets and the binding role of stationarity in portfolio optimization.

Our model is tested empirically out of sample using daily DJIA data between 1997 and 2015. Maximum likelihood estimations are conducted on continuous VECM parameters for strategy implementation. Our empirical findings are summarized as follows: (a) The proposed strategy offers consistent profitability over the analyzed period in terms of Sharpe ratios and cumulative profits. This supports the efficacy of introducing spread option's deltas into the portfolio choice problem. (b) Our pairs selection algorithm outperforms the benchmark Johansen methodology commonly used in the pairs trading literature.

Chapter 3

Supercointegrated

3.1 Introduction

Statistical arbitrage, as a quantitative method of speculation, has been increasingly prevalent along with the evolution of computational finance. One of the most popular statistical arbitrage strategies is pairs trading, which is widely used by hedge funds and investment banks since the mid-1980s. Pairs trading strategies exploit temporary mispricing between paired assets by taking long-short positions. If there is price discrepancy, which is short-lived according to past price information, a trading opportunity arises and can be profited from price adjustment process. In other words, pairs trading is a convergence trade strategy identifying a pair of assets whose prices historically move together. Then, arbitrage profits are generated by unwinding the position in the case that prices converge to the long-run equilibrium over time. Conversely, arbitrageurs suffer losses if by any chance price spreads widen further.

This article shows the performance of pairs trading portfolios when sorted by the level of cointegration of their constituents. We focus on supercointegrated portfolios, that are formed by pairs which do not reject standard cointegration tests at 1% level of confidence. From an out-of-sample exercise, we show that this elite group outperforms our benchmark, the market portfolio, for the period under study: an initial investment of \$1.00 in the market index in November 2000 yields \$1.60 in July 2016. When invested in our supercointegrated portfolio, the initial investment results in \$2.00 at the period end. In terms of realized returns, the supercointegrated portfolio yields an annual return of 6.18%, in contrast to the 3.33% for the market index. This superior performance of the supercointegrated portfolio is related to the volatility level that is much lower than the market portfolio. Notably, our pair strategy leads to an impressive Sharpe ratio of 0.60, against 0.17 for the benchmark during the same time period.

Interested on the sources of risk of the supercointegrated portfolio, they seem to be linked to aggregate market volatility. A regression of the returns of the supercointegrated portfolio against the 3-factor Fama and French (1992) model shows that the market is a statistically significant source of covariance with the returns of the pairs strategy. Momentum and book-to-market ratios do not exhibit covariance with the strategy. The maximum-likelihood estimates of an autoregressive AR(1) model on the realized variance series show that the square root of the unconditional variance is two times lower in the case of pairs portfolio. Surprisingly enough, the autoregressive parameter is similar for the pairs portfolio and the market portfolio. As shown throughout the article, this issue can be employed to obtain a superior performance in terms of Sharpe ratio to index-based strategies.

Whether the superior performance is exclusive of the supercointegrated portfolio is an additional contribution of this article. Remarkably, we also find that the degree of performance of the pairs strategy is positively related to the level of cointegration among pairs. A sorted collection of pairs in terms of their degree of cointegration reveals interesting results. For the portfolio consisting of higher cointegrated pairs in the first quintile, the average excess return of 6.2% is earned per annum which is 2.3% higher than the portfolio with pairs in the second quintile. The similar finding is documented in terms of the Sharpe ratio, which is 0.6 versus 0.5. These results strongly suggest that the performance of pairs portfolio stresses when the cointegration relationship ties up.

A possible explanation to our results relies on the interesting risk-return relationship of the pairs strategy. A projection of contemporaneous returns onto the lagged realized variance exhibits a statistically significant and positive beta coefficient. The source of value seems to be linked to increments of aggregate market volatility. When market volatility rises, a higher level of cointegration among pairs is found, triggering a virtuous circle for our strategy: the number of cointegrated pairs increases --more pairs--, and their quality improves --assets are tied-up under a more intense relationship--. This situation results in the outstanding performance of the supercointegrated portfolio.

In parallel to our findings, this article conducts several exercises to assess the robustness of its main results. The most critical point is the threshold of the pairs strategy. In the history of pairs-trading, a constant threshold, which is determined as the unconditional standard deviation, is commonly applied to trigger trades. To capture the time-varying nature of price spreads, we employ another criteria built on the constant threshold method. This new trigger, named dynamic threshold, is determined as the conditional standard deviation calculated using a 1-year moving window. As shown later, the superior performance of the supercointegrated portfolio to index-based strategies is robust to different threshold criteria. Additional checks on different aspects of the sample as data frequency or transaction costs indicate that the strength of the supercointegrated portfolio is not sample dependent.

Thus, this article studies the performance of supercointegrated portfolios, examining their sources of time-varying risk. The rest of the paper proceeds as follows. Section 2 stresses our contribution to the existing literature. Section 3 details the data and portfolio construction, and Section 4 analyzes the supercointegrated portfolio in an out-of-sample exercise. Section 5 provides international evidence. Section 6 conducts some robustness checks, and conclusions are left to Section 7.

3.2 Contribution to existing literature

The main contribution of this article stresses the importance of the cointegration level on the performance of pairs trading strategies. We adopt a novel perspective for the analysis moving from individual to aggregate sides of the question: instead of the traditional approach, where pairs trading strategies are studied as individual investments, we examine the behaviour of a portfolio, the so-called supercointegrated portfolio, composed by several highly cointegrated pairs. The major result is that the level of cointegration matters.

This paper is related to the area of relative asset pricing which is concerned potential price differentials of two substituted assets that should be priced identically. When price spread appears, relative-value arbitrage involving a long-short position, i.e. pairs trading, exploits the violation of the law of one price and makes profits from expected restoration of price equilibrium. In this way, the principal innovations on the literature could be gathered in two main streams: the improvement of trigger indicators, and the design of alternative methods for selecting the pairs.

With regard to the trigger indicators, the work of Gatev et al. (2006) examining a large sample of liquid U.S. equities has received a remarkable attention in the pairs trading domain. Their trading algorithm is simply implemented based on the historical standard deviation of the spread: trades are opened when price divergence exceeds two standard deviations, and liquidated upon the spread converges under the threshold value. The strategy delivers economically and statistically significant excess returns of around 11% per annum over the period 1962-2002. Following Gatev et al. (2006), Papadakis and Wysocki (2007) and Engelberg et al. (2009) implement similar empirical analysis relying on different samples of U.S. equity market to evaluate and explain the profitability of pairs-trading.[1] The wide usage of this strategy can be attributed to its fairly clear merit: the method based on constant trading trigger is easy to use and, given the nonparametric nature, it is not subject to model misspecifications. Nevertheless, the constant threshold method is selected somehow arbitrarily, ignoring the possibility of dramatic price swing as time passes, and thus whether the fixed threshold is sensible to trigger trading signals is suspicious.

Moving beyond the standard deviation, Elliott et al. (2005) explicitly models the spread using a mean reverting Gaussian Markov chain, observed in Gaussian noise. In continuous time, the application of well-known Ornstein-Uhlenbeck (OU) process allows spread estima-

[1] The popularity of Gatev et al. (2006)'s trading rule is reflected on the empirical tests on many international markets. For instance, Andrade et al. (2005), Broussard and Vaihekoski (2012), and Bowen and Hutchinson (2014) replicate the strategy in Taiwan, Finland and U.K. equity markets, respectively.

tion, which serves as the base of this trading rule. When the observed spread exceeds the estimated value, an investor opens a position in the spread upon the elimination of deviation. As such, this stochastic approach bases the determination of trade decisions on the model's prediction. This approach has two primary strengths. First, the evolution of price spread is described with an OU process, capturing the mean reverting property which is the key to pairs trading. Second, this stochastic model allows parameter forecasting. Parameter estimates is straightforward to obtain given the known distribution of OU process in closed form. Built on the spirit of Elliott et al. (2005), Do et al. (2006) propose a general approach of stochastic residual spread that describes mispricing at the return level. It means that the spread is defined as the return difference of a pair of assets. They suggest taking a position whenever the accumulated spread is larger than certain threshold values. However, no guidance is given on how to further specify the threshold.

Concerning the selection pairs criteria, a common critique to Gatev et al. (2006) is so-called cointegration relationship of selected pairs without resorting to formal testing. As a consequence, the relationship implied from price comovement may be spurious since high correlation does not necessarily indicate mean-reversion properties. This potential drawback is unveiled by Do and Faff (2010). Extending the original sample used by Gatev et al. (2006), the authors illustrate the inefficiency of the method with the finding that 32 percent of all identified pairs do not converge.

In order to better exploit mean reversion properties, Vidyamurthy (2004) also employs a cointegration method for pairs-trading based on the cointegration theory proposed by Engle and Granger (1987). But the spirit of his trading rule is still in similar fashion to the constant threshold method in Gatev et al. (2006), suggesting a open/close position according to the magnitude of price divergence relative to a threshold value. Under the assumption that price spread is a Gaussian white noise, Vidyamurthy (2004) derives the threshold value maximizing profit function for each particular pair, in contrast of fixing the threshold level to be two standard deviations as Gatev et al. (2006). Despite this progress, the kernel of both methods is the same and relies on a constant trigger level. Similarly, cointegration approach is also applied to develop pairs trading strategies for achieving the minimum required profit per trade given a selected trading threshold, i.e. Lin et al. (2006) and Puspaningrum et al. (2010). Undoubtedly the constant threshold method occupies an important position in the development of pairs trading.

The most recent contributions to pairs-trading research fall into the domain of stochastic control and continuous-time cointegration; see Jurek and Yang (2007) and Liu and Timmermann (2013), respectively. Both models require daily rebalancing. In this sense, their dynamic feature leads to heavy transaction costs so that a comparison with less dynamic threshold methods would shed light on their effectiveness in face of market frictions. Similarly, Cartea et al. (2015) also provides an interesting discussion of pairs trading within the context of algorithmic trading using continuous-time cointegration.

3.3 Data, pairs selection and portfolio construction

This analysis focuses on the U.S. equity market, and particularly investigates traded equities of the S&P100 index. This index comprises 100 leading, large cap companies across multiple industries. This provides a blue-chip representation of sector leaders in the U.S. market, regarded as a proxy of the overall US stock market. The constituents of S&P100 index are large-cap stocks that can be traded in stock exchanges with complete access to market participants. The dataset comprises daily closing prices for S&P100 constituents from January 1st, 1998 to June 24th, 2016, leading to 4822 observations for most of these stocks.[2] Once excluded the series with missing values or data shortness, our sample of stocks results in 3486 possible combinations of equity pairs. Data has been collected from Datastream.

To identify the matching partners, we proceed in a two-step procedure. Roughly speaking, this two-step selection procedure combines Johansen (1991) method and Engle and Granger (1987) method. Johansen (1991) method is applied to confirm a pair of assets shar-

[2]Exceptions are found on the following sixteen companies whose price data are available for a shorter period: Accenture, Alphabet, Facebook, General Motors, Goldman Sachs, Kinder Morgan, Mastercard, Metlife, Mondelez International, Monsanto, Pfizer, Philip Morris International, Priceline Group, Twenty-First Century Fox, United Parcel, and Visa. These companies have been excluded from the sample.

ing long-run equilibrium, while Engle and Granger (1987) method is used to determine the cointegration error z_t given that the cointegration regression estimated using OLS is practically easy for portfolio construction. In this way, we firstly test the existence of cointegration among each possible pair by conducting the Johansen (1991) cointegration test at the 1% confidence level.

In a second step, we refine our selection of pairs by running the OLS regression,

$$y_t = \alpha + \beta x_t + z_t \tag{3.1}$$

where y_t and x_t are Asset 1 and Asset 2 constituents of a possible pair, and z_t is a normally distributed error term. The idea is that the cash amount of $\hat{\alpha}$ and $\hat{\beta}$ units of x_t are invested to replicate the prices of y_t. We then claim that y_t is equivalent to the replicating portfolio $(\hat{\alpha} + \hat{\beta}x_t)$ under the existence of cointegration relationship. Considering that the essence of pairs trading is an arbitrage strategy with simultaneous long-short positions on paired assets, a positive value of $\hat{\beta}$ is required. Otherwise, positions in the same direction would be taken, long-long or short-short, which is the case violating the mechanism of pairs trading. In addition, $\hat{\alpha}$ is also required to be positive so as to rule out the possibility of borrowing money, for the purpose of constructing self-financing portfolios. In other words, we exclude those pairs with negative OLS estimates $\hat{\alpha}$ and $\hat{\beta}$.

The trading mechanism is described as follows. First, we set a threshold to trigger the strategy. When price spread exceeds the threshold, we open a long-short position one day after the appearance of the trading signal, particularly longing the underpriced asset and shorting the overpriced one. Similarly, the initial positions will be liquidated one day later when prices converge to a level that is within the border of threshold values. In other words, the opening/close of pairs positions is delayed by one day. After a round-trip trade has been completed, a pair will be subject to the same trading rule again.

The prior literature commonly uses a constant threshold as the trading trigger; see, among others Gatev et al. (2006) and Do and Faff (2010). This constant threshold uses to be the unconditional standard deviation from historical spreads, and it keeps constant across the entire trading period. Instead of the constant threshold, this article proposes a dynamic criteria which takes advantage of the most recent information in the time series of the price spreads. This dynamic threshold is determined as the standard deviation calculated using a 1-year rolling-window --the 252 previous observations-- of the price spread. Given a particular trading day, we compare the estimated 1-year rolling-window volatility with the realized price deviation of the pair in that day. If the realized deviation is higher than one standard deviation of the volatility estimate, the pairs strategy is opened. The dynamic threshold introduces the dynamics into the traditional pairs strategy, as the strategy is now conditional to the most recent price information.[3]

Finally, we construct an equal-weighted portfolio constituted by our selection of supercointegrated pairs. This procedure benefits from the risk of diversification, and it has been widely employed in the asset pricing literature; see, for instance, Goetzmann and Kumar (2008), DeMiguel et al. (2009), and Tu and Zhou (2011). The excess return of the portfolio is measured as the return on committed capital. In other words, the mark-to-market portfolio payoffs are scaled by the number of selected pairs during a trading interval.[4] The payoff is interpreted as excess return since the trading profits or losses are earned from one dollar investment in simultaneous long-short positions. In parallel to the supercointegrated portfolio, we construct as a control an additional naive strategy holding the S&P100 index under a simple buy-and-hold trading rule.

3.4 Supercointegrated portfolio

This section examines the performance of the supercointegrated portfolio. We evaluate the strategy in an out-of-sample scenario, a most demanding framework. Some complementary analysis are left to the robustness section.

[3]The pairs strategy is also tested for the constant threshold in Section 6. The main conclusion is that results are robust to the choice of the threshold.

[4]This measure is obviously conservative, since it considers the opportunity cost of capital if a pair does not trade for some days of the trading interval. This scenario takes place commonly, as it is rare all selected pairs meet the trading criteria in a given day. As such, the obtained return in our study would not be the upper bound of trading profits and, in this sense, its superior performance is more convincing if this is the case.

3.4.1 Anatomy of supercointegrated pairs returns

To endorse the competence of the proposed strategies, we examine the performance out of sample. The procedure is as follows: we select a three-year window (formation period) to identify cointegrated pairs via Johansen (1991) method and, for these chosen pairs, estimate the cointegration relationship by OLS. Then, the resulting estimates are applied to trades during a six-month forward-ahead period. This procedure is repeated through the remaining sample period.[5] It can be argued that this procedure can be improved by increasing the frequency of re-estimation of the portfolio pairs, and many other alternatives for the out-of-sample exercise could be implemented. In this way, we have adopted a conservative perspective on presenting here the out-of-sample strategy, and other different variations to our strategy design have been left to the robustness check section.

To measure the performance of the portfolio we have considered the Sharpe ratio and the performance measure proposed by Goetzmann et al. (2007) (PMG henceforth). Given the portfolio returns, r_t, and the corresponding returns of the risk-free asset, $r_{f,t}$, PMG is calculated as

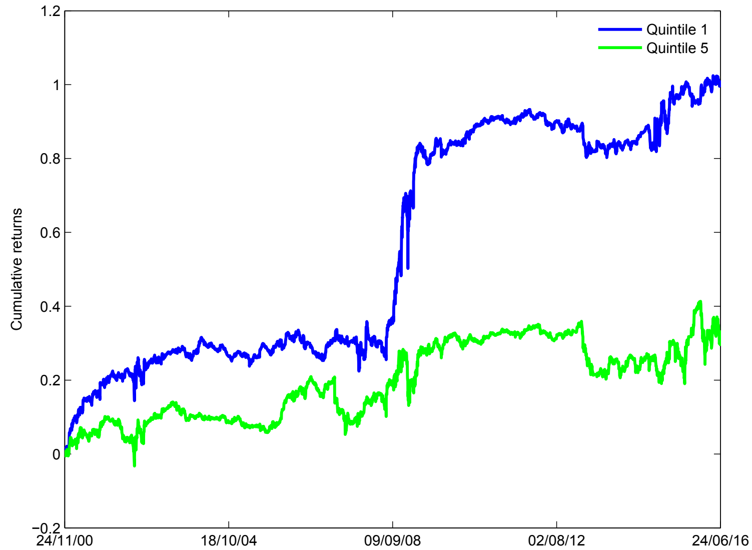
$$PMG = \frac{1}{(1-\gamma)\,\Delta t} \ln \left(\frac{1}{T} \sum_{i=1}^{T} \left(\frac{1+r_t}{1+r_{f,t}} \right)^{1-\gamma} \right) , \tag{3.2}$$

where T is the length of the return series and γ is the risk-aversion coefficient. The PMG measure has been widely used to study the performance of alternative investments, such as hedge funds, generating robust ratings; see, among others, Brown et al. (2008), Bali et al. (2013) and Jackwerth and Slavutskaya (2016).

Once settled the strategy, we study the effect of the cointegration degree on the performance. To this end, we sort the pairs, screened out at the 1% level, in ascending order according to their p-value of Johansen (1991) test. Then, we divide the cross-section into quintiles; Quintile 1 (Q1) is the set of pairs with lowest p-value, and Quintile 5 (Q5) is the

[5] A detailed explanation of our procedure is provided here. In an initial step, we estimate a first set of cointegration parameters $(\hat{\alpha}_{(1)}, \hat{\beta}_{(1)})$ using data from January 1998 to December 2000. These parameters are employed to implement the strategy during period January 2001 to June 2001, which is left for trading. Then, we move six-month ahead our window, and re-estimate a new set of cointegration parameters $(\hat{\alpha}_{(2)}, \hat{\beta}_{(2)})$ from July 1998 to June 2001, and the period July 2001 to December 2001 is now left for trading. This procedure is repeated in thirty-two non-overlapping six-month trading periods. Note that the last trading period is a bit less than six months since the sample ends June 24th, 2016.

Figure 3.1: Cumulative excess returns of quintile portfolios: Q1 and Q5



This figure displays the cumulative excess returns for two equally-weighted portfolios constructed by the same number of pairs. These pairs are sorted in ascending order according to the p-value of Johansen (1991) test, and then divided into quintiles. Q1 (Q5) corresponds to the portfolio of most (lowest) cointegrated pairs. Results are obtained using the dynamic threshold method over the out-of-sample period. Data spans from November 24th, 2000 to June 24th, 2016.

one with the highest p-value.[6] Figure 1 displays the cumulative excess returns of the Q1 and Q5 portfolios. The most interesting result is that a higher level of cointegration exerts positive impacts on the strategy performance. The Q1 portfolio, whose final return reaches its highest level at 1.0, apparently dominates Q5, which is formed with less cointegrated pairs. Q5 represents the most relaxing relation of cointegration, and it performs weakly all the time. The difference in cumulative return reflects straightforwardly the strength of Q1, exhibiting a return gap up to around 0.70 between Q1 and Q5. Interestingly enough, this gap enlarges during the financial distress period in September 2008.

Table 1 also provides a different perspective on the results. This table reports the main statistics for the five sorted portfolios (Panel A), and the aggregate portfolio that contains all pairs (Panel B). Again, the main finding is that the portfolio at a higher level of cointegration (Q1) is likely to achieve more attractive excess returns, with a similar risk than the remainder portfolios. For instance, the Q1 portfolio yields an average return of 6.2% per annum, which

[6] Pairs in each quintile are used to form an equally weighted portfolio such that performance comparison is still proceeded at the portfolio level.

Table 3.1: Performance of five equal-weighted pairs portfolios

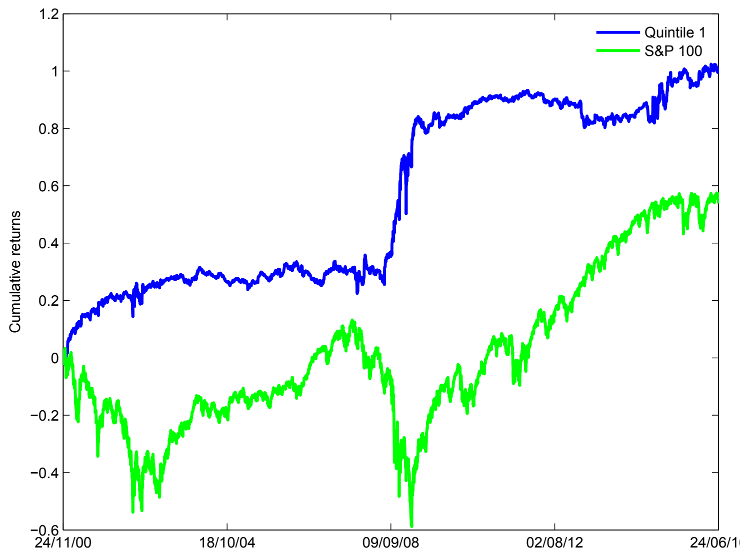
Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe	PMG ($\gamma = 2$)	PMG ($\gamma = 3$)
					Panel A: Quintile portfolios					
Q1	0.0618	0.0000	0.1033	1.67	57.32	0.10	-0.10	0.60	0.0368	0.0315
Q2	0.0386	0.0000	0.0757	0.32	15.38	0.04	-0.04	0.51	0.0184	0.0155
Q3	0.0278	0.0000	0.0976	-1.47	65.84	0.08	-0.11	0.29	0.0038	-0.0011
Q4	0.0329	0.0000	0.1109	1.91	101.99	0.13	-0.13	0.30	0.0062	0.0002
Q5	0.0181	0.0000	0.0878	0.02	19.67	0.05	-0.05	0.21	-0.0040	-0.0079
					Panel B: All-pairs portfolio					
All-pairs	0.0348	0.0000	0.0731	1.12	63.52	0.07	-0.08	0.48	0.0151	0.0125

This table reports the main statistics for five independent portfolios (Panel A) ranked according to the p-value of Johansen (1991) test, and the portfolio containing all pairs (Panel B). Portfolios in Panel A are constructed using the same number of pairs, and then sorted by quintiles. Reported statistics are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return, the (annualized) Sharpe ratio and the performance measure of PMG. Results are obtained using the dynamic threshold method over the out-of-sample period. Data spans from November 24th, 2000 to June 24th, 2016.

represents 2.3% more than Q2, the second best portfolio. This significant return of Q1, joint with a comparable volatility to the remaining portfolios, pushes up the Sharpe ratio of the top quintile. As shown in Panel A, the Sharpe ratio improves in a positive, monotonic pattern, increasing from 0.2 to 0.6. Such evident enhancement is definitely worthy of close attention, and provides investors an important practical implication: pairs of assets should be selected at the 1% significant level, in order to focus on the top quintile.

From Table 1, we also find that the Q1 also improves the performance of the portfolio that aggregates all pairs under a risk-adjusted basis. Q1 exhibits a melioration of the Sharpe ratio of 0.12 points when compared to the baseline result of the all-pairs portfolio. In contrast, the performance of the rest portfolios drops dramatically in terms of Sharpe ratio. In line with these previous results, the PMG measure reveals that most cointegrated portfolio (Q1) still exhibit a better performance than the rest. Lastly, we find a positively skewed distribution of excess returns which is common to four quintiles and the all-pairs portfolio. Combined

Figure 3.2: Cumulative excess returns of the portfolio Quintile 1 and the market benchmark



This figure displays the cumulative excess returns for the highest cointegrated portfolio Q1 (blue line) against the cumulative returns of the S&P 100 index (green line). Results for Q1 portfolio are obtained using the dynamic threshold method over the out-of-sample period. Data spans from November 24th, 2000 to June 24th, 2016.

with the large values of kurtosis, a high level of tail risk is suggested as plenty of hedge fund strategies.

The cumulative excess returns for the Q1 portfolio and the S&P100, our market benchmark, are depicted in Figure 2. Some interesting conclusions arise from the inspection of this figure. First, the Q1 portfolio yields a superior return than the naive strategy. The departure of cumulative returns between two portfolios is specially remarkable during the periods from December 2000 to January 2002, and June 2008 to December 2009, where the Q1 portfolio performed exceptionally well. In contrast to these dramatic increments of profitability in Q1, the market benchmark suffered a dramatic decline due to the dot-com crash in 2000, and the global financial crisis in 2008. Second, from observation of Figure 2, the volatility of Q1 seems to be significantly lower than the market. In statistics not provided here, but available upon request, the annual volatility of the S&P100 is 19.34% for the period under analysis, against a 10.33% of the Q1 portfolio. This results in a Sharpe ratio of 0.17 for the market benchmark, distant from the 0.60 coefficient of the Q1 portfolio.

Table 3.2: OLS regressions of returns on Fama and French (1992) factors

Model	Intercept	MktRF	SMB	HML
Q1	0.0021**	0.0006*	0.0000	0.0005
	(0.0009)	(0.0002)	(0.0005)	(0.0006)
S&P100	0.0047	0.0055***	-0.0011	0.0009
	(0.0037)	(0.0009)	(0.00014)	(0.0014)

OLS regressions of supercointegrated portfolio (Q1) and market benchmark (S&P100). The three factors are the excess market return, SMB, and HML. Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

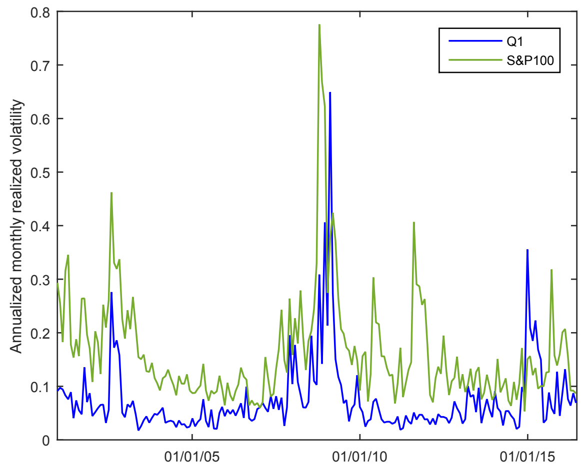
These two previous observations allude to an interesting (and positive) risk-return balance in pairs trading strategies, where increments in the aggregate market volatility are related to a higher profitability of the cointegrated portfolio.[7]

Interested on possible explanations to the performance of pairs portfolios, we project the returns of the supercointegrated portfolio and the market benchmark onto the three factor model of Fama and French (1992). Table 2 reports the robust OLS estimates. The estimates show that, while market beta is positive and statistically significant for the pairs portfolio, the documented figure is extremely small. This indicates that the performance of pairs strategies is not closely linked to the dynamics of the market, i.e. these strategies are close to market-neutral, which has been shown in the prior literature, e.g. Gatev et al. (2006), Liu and Timmermann (2013), and Rad et al. (2016).[8] In the view of asset allocation, the feature of market neutrality brings about diversification benefits to a portfolio holder whose positions are strongly correlated with the market. Moreover, "high-minus-low" and size factors are not statistically significant at standard confidence levels, and their betas are close to zero.

[7]This result is similar to that found by Barroso and Santa-Clara (2015) and Daniel and Moskowitz (2016) within the context of momentum portfolios.

[8]Market neutrality means small/insignificant exposures to the market. This outcome is achieved in pairs strategies due to the simultaneous long-short positions on two close-substituted assets. Given this, Liu and Timmermann (2013) ex ante assume that a pair of assets have identical market betas.

Figure 3.3: The monthly realized volatility of portfolios



This figure shows the realized volatilities of the Q1 portfolio (most cointegrated pairs, blue line) and the S&P 100 index. The monthly realized volatility is obtained from daily returns.

3.4.2 The time-varying risk of pairs trading portfolio

It is always interesting to look at the other side of a coin. We thus analyze the time-varying risk of pair trading strategies. Inspired in Barroso and Santa-Clara (2015), we compute the realized variance estimate from the past 21 observations of daily return,

$$RV_{i,t} = \sum_{j=0}^{20} r^2_{i,t-j} \tag{3.3}$$

where i stands for the S&P 100 and Q1 portfolio, and $r_{i,t-j}$ represents the daily returns of each strategy.

Figure 3 displays the time series of the realized volatility for the market index and the Q1 portfolio (most cointegrated pairs). As expected, the realized volatility of the Q1 portfolio is lower than for the market index. At the limit, the beta of the pair trading portfolio should converge to zero since it involves long and short weights, therefore the systematic risk of the pair trading portfolio should be negligible. In our case, the Q1 portfolio only incorporates the most cointegrated pairs, thus the market risk is reduced but not completely erased. Therefore, we can also appreciate a similar up-down pattern between both series, suggesting a link between the market and Q1 realized volatilities. The correlation between

Table 3.3: AR(1) estimates of realized variance

Model	Q1	S&P100
c	0.00065	0.00095
	(0.00054)	(0.00071)
ϕ	0.2718***	0.6913***
	(0.0591)	(0.0382)
σ^2	8.0×10^{-6}***	1.58×10^{-5}***
	(1.2×10^{-7})	(3.9×10^{-7})
LogLk	-858.67	-793.14
Obs.	193	193

OLS regressions of supercointegrated portfolio (Q1) and market benchmark (S&P100). The three factors are the excess market return, SMB, and HML. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

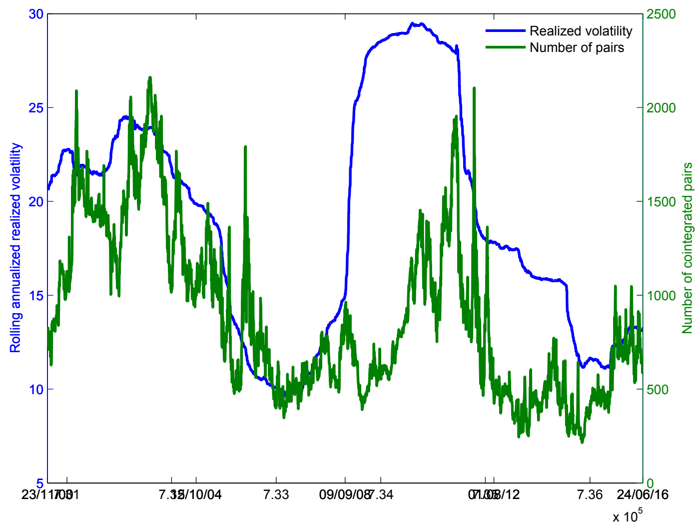
both realized variance series is equal to 0.56. This finding suggests that our strategy is not a pure market-neutral portfolio. Another important aspect of the realized variance series is their persistence. To evaluate this parameter, we estimate an autoregressive AR(1) model,

$$RV_{i,t} = c + \phi RV_{i,t-1} + \epsilon_t \tag{3.4}$$

where $RV_{i,t}$ is the realized variance observation at time t, c is a constant, and ϕ the autoregressive coefficient. The error term ϵ_t is normally distributed with zero mean and variance σ^2. The results from the estimated AR(1) model are provided in Table 3. First, we verify that the realized variance of the Q1 portfolio exhibits autocorrelation; the estimate of ϕ is strongly significant and close to 0.27. Notably, the differences in scale are also relevant: pairs trading strategies exhibit a much lower variance on average compared to the market benchmark, as it is inferred from the σ^2 estimate, which is 2 times lower for realized variance of the Q1 portfolio than for the S&P 100 one.

Despite the similarities with the market realized variance (see Figure 3 and Table 3) a second factor may affect the portfolio risk: the number of highly cointegrated pairs. The pairs

Figure 3.4: The rolling 3-year realized volatility of S&P 100 vs. number of cointegrated pairs



This figure displays the evolution of the number of cointegrated pairs, against the rolling realized volatility of S&P 100. The calculation is done by implementing a rolling 3-year window over the entire sample period.

trading strategy is implemented in a moving window, so the number of cointegrated pairs may vary: as time moves on, the latest price is included in the formation period searching for cointegration relationship within a pair of assets. The new arrival of information, reflecting the aggregate market sentiment, affects price movements and further the existing relationship among assets. The variation in the number of pairs is expected to be substantial as the degree of closeness is distinct under diverse market conditions. The number of cointegrated pairs is connected to the evolution of the realized variance of S&P 100. Figure 4 shows a high correlation between these two series, with a coefficient up to 0.53. We then may conclude that the pool of cointegrated pairs gets larger as the market volatility goes up, such relation even evident for the case of 1% level cointegration. This indeed supports our focus on pairs accepted at 1% level in this article, given a reliable conjecture that high sensitivity to the market may result in large profitability potentials.

Given the persistent pattern of the realized volatility it becomes interesting to analyze the risk-return relationship for the supercointegrated portfolio in order to establish a relationship between past realized volatility and contemporaneous returns. For this purpose we estimate by OLS the following model,

Table 3.4: Risk-return relationship

	η_0	η_1
estimate	0.000153*	0.10421***
s.e.	(7.9×10^{-5})	(0.02572)
N	192	
R^2	0.0795	

This table displays the OLS estimation of the monthly returns of the Q1 portfolio against the one month lagged realized volatility.

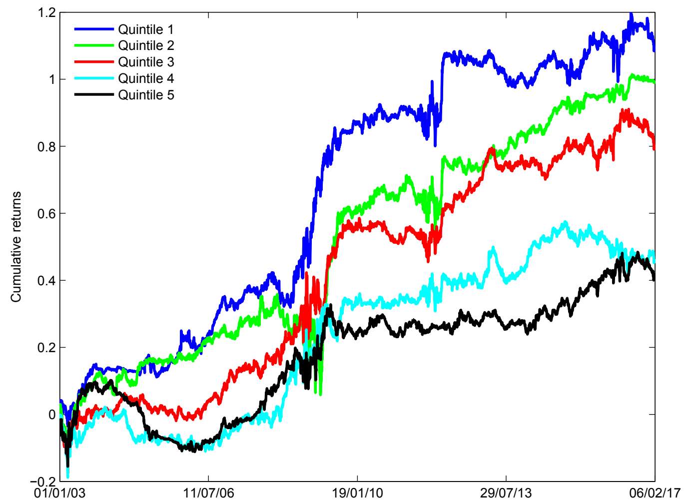
$$R_t = \eta_0 + \eta_1 RV_{t-1} + \xi_t \ , \quad \xi_t \sim N(0,\, 1) \tag{3.5}$$

where R_t stands for the t month realized return of Q1 portfolio, RV_{t-1} is the corresponding one month lagged realized volatility (depicted in Figure 3), ξ_t is the error term and $(\eta_0\ ,\ \eta_1)$ are the coefficients of the linear fit. Table 4 exhibits the results of the OLS estimation. This results show a positive relationship between last month realized variance and one month ahead future return. This finding is interesting because as volatility increases, stock markets use to fall, but the pair trading portfolio increases its expected return. Therefore, the hedging properties of the pair trading portfolio seem relevant.

3.5 International evidence

We now examine whether our results are also extensible to other markets. The international evidence of pairs trading has already been tested in an international context; see, for instance, Dunis and Lequeux (2000), Broussard and Vaihekoski (2012), Bowen and Hutchinson (2014) and Dunis and Ho (2016). We then analyze the behaviour of the supercointegrated portfolio using the companies included in the STOXX Europe 600 index. Our analysis is restricted to those firms located in the Eurozone whose stock prices are quoted in euro. The selected sample comprises 292 companies across ten countries of the European common currency area. Data frequency is daily, and it spans from January 1st, 2000 to February 6th, 2017,

Figure 3.5: Cumulative excess returns of quintile portfolios for European stocks



This figure displays the cumulative excess returns for five equal-weighted portfolios constructed by the same number of pairs. These pairs are sorted in ascending order according to the p-value of Johansen (1991) test, and then divided into quintiles. Q1 (Q5) corresponds to the portfolio of most (lowest) cointegrated pairs. Results are obtained using the dynamic threshold method over the out-of-sample period. Data spans from January 1st, 2000 to February 6th, 2017.

resulting in 4,461 trading observations on average. The data source is Datastream.

The cumulative returns of pairs portfolios sorted by their level of cointegration are depicted in Figure 5. Once again, these results exhibit a monotonic behaviour of cumulative returns according to their level of cointegration. As shown, the cumulative returns of the Q1 portfolio are higher than others. For instance, the Q1 portfolio yields almost three times more return than the Q5 portfolio. This outstanding performance of Q1 improves during periods of high volatility, as the concerns about the Greek sovereign debt during 2010, and the incertitude in Italian and Spanish debt during August 2011.

Table 5 provides a complementary perspective on the results of Figure 5. This table shows the annual excess returns for the five quintiles. We observe that the average return of the strategy decreases as the degree of cointegration does. The volatility of the strategy is quite similar among portfolios, resulting in an improvement of the Sharpe ratio of the Q1 portfolio. Interestingly, the kurtosis of the supercointegrated portfolio is lower than the rest.

Table 3.5: Annual excess returns for European stock portfolios

Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Q1	0.0811	0.0093	0.1198	0.78	12.36	0.07	-0.05	0.68
Q2	0.0673	0.0125	0.1231	1.10	20.53	0.08	-0.05	0.55
Q3	0.0539	0.0039	0.1101	0.44	19.39	0.07	-0.07	0.49
Q4	0.0308	0.0021	0.1007	0.74	13.45	0.07	-0.04	0.31
Q5	0.0286	0.0013	0.0894	0.47	13.47	0.05	-0.04	0.32

This table reports the main statistics for five independent portfolios of European stocks ranked according to the p-value of Johansen (1991) test. Portfolios in are constructed using the same number of pairs, and then sorted by quintiles. Reported statistics are the mean and median excess return (annualized), the (annualized) standard deviation, skew, kurtosis, the maximum and minimum daily excess return, the (annualized) Sharpe ratio. Results are obtained using the dynamic threshold method over the out-of-sample period. Data spans from January 1st, 2000 to February 6th, 2017.

Thus, these results show evidence of the performance of the supercointegrated portfolio in an international context. The existence of a close relationship between the cointegration degree and the performance of pairs portfolios are detected in other markets.

3.6 Robustness check

This section performs some additional analysis to check the robustness of our results to different specifications of the trading model. The main conclusion of this section is that results are robust to different specifications of the trigger, the design of out-of-sample exercise, and some microstructure issues as sampling frequencies and transaction costs, among others.

3.6.1 Alternative triggers

The trigger plays a critical role in pairs strategies. In this paper, the trigger consists of a dynamic threshold computed as the standard deviation using a 1-year rolling-window of price spreads. Taking advantage of the predictability of variances, we wonder whether a refinement of the trigger leads to an improvement of the portfolio performance. To this end, we forecast the standard deviation of the spread series using a GARCH model Bollerslev (1986), using this predicted standard deviation as the dynamic threshold value.[9]

[9] The most widely used GARCH(1,1) model is applied to our analysis. Let the standard deviation of the spread series to be

Table 3.6: The baseline dynamic method vs. GARCH (1,1)

Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: All pairs								
Dynamic method	0.0348	0.0000	0.0731	1.12	63.52	0.07	-0.08	0.48
Dynamic method_GARCH	0.0335	0.0000	0.0661	0.98	65.68	0.06	-0.07	0.50
Panel B: Pairs portfolios under GARCH (1, 1)								
Q1	0.0546	0.0000	0.0877	1.65	71.63	0.09	-0.10	0.62
Q5	0.0146	0.0000	0.0808	-0.02	20.49	0.06	-0.05	0.18
Top 20	0.0420	0.0000	0.0840	1.51	72.48	0.09	-0.09	0.50

This table reports in Panel A the performance of the original dynamic strategy using 1-year rolling standard deviation as the threshold, relative to the dynamic trading rule based on estimated volatility from GARCH (1, 1) model. The selected pairs are divided into quintiles according to their p-value of the Johansen test. Panel B shows results of three different pairs portfolios with sorted pairs under the GARCH (1, 1)-based dynamic method. The "Top 20" portfolio includes the 20 pairs with the lowest p-value in each non-overlapping trading period. These results are obtained between November 24th, 2000 and June 24th, 2016.

The question whether this refinement enhances the performance of the strategy is answered in Table 6. Results in Panel A shows that both methods deliver similar performance in terms of Sharpe ratio. Our strategy generates a higher mean return, while the predictability ability of GARCH (1, 1) results in a lower return volatility. Results in Panel B are again in favor of the portfolio including higher cointegrated pairs. Then we move further to compare the performance between two methods across three independent portfolios. The evidence is the same as what we find above, both methods expressing comparable profitability. Figure 6 reveals the superiority of our dynamic method in generating absolute returns, while the lower return volatility under the effect of GARCH (1,1) boosts the Sharpe ratio slightly.

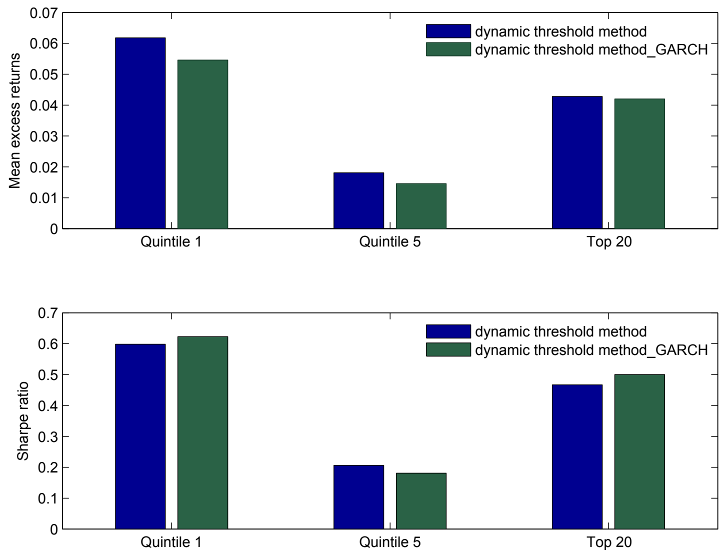
In sum, the sophisticated GARCH (1,1) model seems not able to significantly improve the trading performance, which may enhance confidence about the profitability of our simple

σ_t, where σ_t is a nonnegative process. The GARCH (1,1) model is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{3.6}$$

where $\alpha_0 > 0$, $\alpha_1 > 0$, and $\beta_1 > 0$; α_0 is the constant, α_1 is the first order of the ARCH term, and β_1 is the first order of the GARCH term. The procedure for the out-of-sample analysis is as follows: we estimate a GARCH (1,1) model for each spread serieas, using a 3-year estimation window of data that includes 756 observations. Therefore, the conditional standard deviations, σ_t, for each nonoverlapping trading period, are calculated recursively using the forecasted variance equation. For each day, there is a particular predicted standard deviation serving as the threshold.

Figure 3.6: Mean excess returns and Sharpe ratios: Baseline vs. GARCH



This figure displays the performance of the baseline dynamic method and the dynamic trading rule based on estimated volatility from GARCH (1, 1) model, in terms of annualized mean excess return and Sharpe ratio. The selected pairs are divided into quintiles according to their p-value. The performance results are shown for three different pairs portfolios, Quintile 1, Quintile 5, and "Top 20". The "Top 20" portfolio includes the 20 pairs with the lowest p-value in each non-overlapping trading period. These results are from November 24th, 2000 to June 24th, 2016.

trading method.

3.6.2 The design of out-of-sample exercise

A potential critique to our results might rely on the arbitrary choice of some features of the out-of-sample exercise. For instance, the baseline trading model used throughout the article employs a 3-year window to estimate the cointegration vector, applying the resulting point estimates to trade during the subsequent 6-month window. In other words, the parameters of the model are constant during the 6-month trading period, but the trigger is still conditional to past volatility.

To improve this procedure, we enhance the dynamics of the strategy by estimating the OLS coefficients using a 3-year window which moves forward every day. Then, the model parameters are updated on a daily basis. We call this strategy as advanced dynamic method. Table 7 summarizes the excess returns (Panel A) for the advanced dynamic strategy. For the ease of explanation, we also introduce in Panel A the baseline dynamic strategy which has been shown to be superior. We see that this upgraded method achieves a substantial improvement in the excess return by 25 percent, increasing from 3.48% to 4.36% per annum. Also importantly, the boost of return is not at the expense of a dramatic rise of volatility, only upward by 13 percent. The Sharpe ratio of the advanced dynamic strategy is 0.53 and accordingly gets enhanced. Panel B in Table 7 explores the performance of different pairs portfolios under the advanced dynamic method. We sort pairs in ascending order according to their p-value from Johansen (1991) test, and partition them in quintiles like in the previous section to form independent portfolios of equal size. Panel B shows that, the Q1 portfolio, with higher cointegrated pairs, yields the best performance among the three portfolios, offering an average excess return of 6.04% and the Sharpe ratio up to 0.66. The dominance of Q1 over Q5, including the least cointegrated pairs, is apparent given about two times increment in mean return and Sharpe ratio. This evidence is consistent with our previous results about level of cointegration and strategy performance. We also follow the literature to form a portfolio of the top 20 pairs. It is observed that the "Top 20" portfolio provides a similar performance as the Q1, which in turn supports the positive role of p-value in strategy performance.

Figure 7 further compares the mean excess return and Sharpe ratio at different levels of cointegration for the advanced (blue bars) and baseline (green bars) dynamic methods, respectively. We clearly observe the improvement incurred from the advanced strategy in both criteria, at least maintaining a high performance. Some insights are as follows. The Sharpe ratio of Q1 achieves a moderate increase by 10.0%. The improvement is more evident for the Q5 and the "Top 20" portfolios. In specific, the mean return and Sharpe ratio of the "Top 20" portfolio grows from 4.28% to 5.18%, and from 0.47 to 0.60.[10] This suggests a strong effect of the advanced method on the top ranking pairs. Figure 8 depicts the cumulative returns earned from investment on portfolios under two dynamic strategies. The

[10]The "Top 20" portfolio under the baseline dynamic method achieves the annualized mean return of 4.28%, the volatility of 9.17%, and the Sharpe ratio of 0.47.

Table 3.7: The baseline dynamic method vs. the advanced method

Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: All pairs								
Dynamic method_Adv	0.0436	0.0000	0.0827	2.88	69.98	0.10	-0.05	0.53
Dynamic method	0.0348	0.0000	0.0731	1.12	63.52	0.07	-0.08	0.48
Panel B: Pairs portfolios under the advanced dynamic method								
Q1	0.0604	0.0000	0.0921	1.72	44.96	0.10	-0.07	0.66
Q5	0.0318	0.0000	0.0971	0.03	96.52	0.12	-0.11	0.33
Top 20	0.0518	0.0000	0.0868	2.97	75.56	0.10	-0.05	0.60

This table reports in Panel A the performance of the advanced and baseline dynamic threshold methods applied to pairs strategies. The advanced dynamic pairs strategy refers to a threshold method of rolling standard deviations relative to dynamic price spreads based on rolling OLS estimates. The selected pairs are divided into quintiles according to their p-value of the Johansen test. Panel B shows results of three different pairs portfolios with sorted pairs under the advanced dynamic method. The "Top 20" portfolio includes the 20 pairs with the lowest p-value in each non-overlapping trading period. These results are obtained between November 24th, 2000 and June 24th, 2016.

cumulative returns are 0.70 for the advanced dynamic method with all pairs. While Q1 generates the highest cumulative return of about 0.97, the relevant portfolio "Top 20" also offers a decent outcome of 0.84.

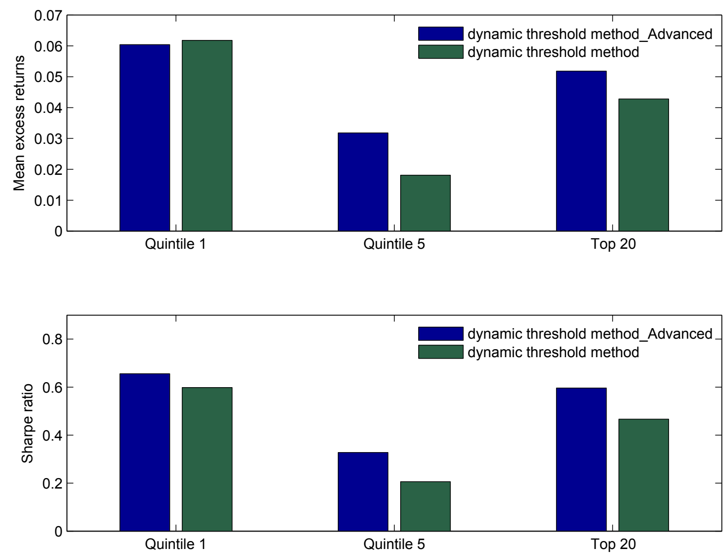
We conclude from this analysis that the advanced dynamic strategy leads to a significant improvement in the out-of-sample performance as a consequence of the introduction of rolling estimates for all relevant parameters in the trading process.

3.6.3 Out-of-sample analysis using open prices

A common concern about the performance of pairs portfolio strategies is that its implementation is not realistic: closing prices are used for liquidating trading positions, and some risks related to the strategy are, for instance, close positions with next day prices, used to be ignored.

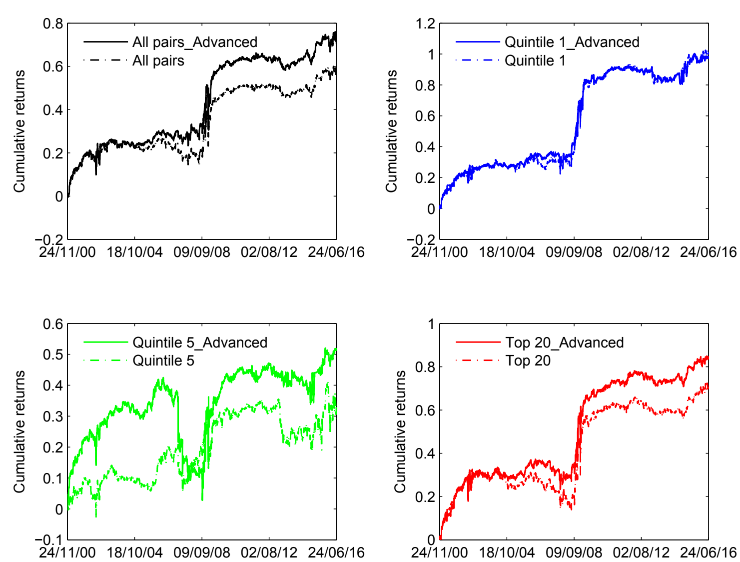
To examine the impact of this question in our strategy, we explore the out-of-sample performance using daily open prices for trades. This new dataset is applied to the dynamic

Figure 3.7: Mean excess returns and Sharpe ratios: Baseline vs. Advanced method



This figure displays the performance of the advanced and baseline dynamic threshold methods, in terms of annualized mean excess return and Sharpe ratio. The advanced dynamic pairs strategy refers to a threshold method of rolling standard deviations relative to dynamic price spreads based on rolling OLS estimates. The selected pairs are divided into quintiles according to their p-value. The performance results are shown for three different pairs portfolios, Quintile 1, Quintile 5, and "Top 20". The "Top 20" portfolio includes the 20 pairs with the lowest p-value in each non-overlapping trading period. These results are from November 24th, 2000 to June 24th, 2016.

Figure 3.8: Cumulative returns: Baseline vs. Advanced method



This figure describes the cumulative returns of four different pairs portfolios. The graph in top left corner refers to the portfolio including all selected pairs, while the rest of graphs represents portfolios including a set of pairs which are sorted. These results are obtained from the advanced and baseline dynamic threshold methods between November 24th, 2000 and June 24th, 2016.

Table 3.8: The performance of pairs portfolios using open prices

Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
				Panel A: All pairs				
Dynamic method	0.0312	0.0000	0.0706	1.98	44.29	0.06	-0.06	0.44
Constant method	0.0268	0.0000	0.0695	1.97	45.14	0.06	-0.06	0.38
				Panel B: Pairs portfolios under the dynamic method				
Q1	0.0619	0.0000	0.1024	2.32	35.04	0.10	-0.07	0.60
Q5	0.0223	0.0000	0.0867	0.51	14.65	0.05	-0.04	0.26
Top 20	0.0371	0.0000	0.0876	2.07	45.28	0.08	-0.07	0.42

This table reports in Panel A the performance of dynamic and constant threshold methods applied to pairs strategies using open prices. The selected pairs are divided into quintiles according to their p-value of the Johansen test. Panel B shows results of three different pairs portfolios with sorted pairs. The "Top 20" portfolio includes the 20 pairs with the lowest p-value in each non-overlapping trading period. These results are obtained between November 24th, 2000 and June 24th, 2016.

and constant methods.[11] Table 8 summarizes the results obtained under open prices. As shown in Panel A, both pairs strategies experience a slight drop in the average excess return compared to closing prices. Although this decline drags down the performance of Sharpe ratio for the dynamic method, its outperformance maintains as we have shown. Panel B of Table 8 exhibits the performance of sorted portfolios by cointegration degree. The first quintile of cointegrated pairs, Quintile 1, provides almost the same outcomes in the mean return and Sharpe ratio, as documented under the usage of closing prices. Again, the superiority of Quintile 1 is obvious to the rest portfolios. In this context, the positive association holds between the degree of performance and the closeness of paired equities measured by cointegration.

In summary, the main conclusions about pairs strategy with closing prices also extend to the usage of open prices.

3.6.4 Performance analysis at different frequencies

Table 9 summarizes the excess return distribution of two particular trading rules in Panel

[11]The prior literature commonly uses a constant threshold as the trading trigger; see, among others Gatev et al. (2006) and Do and Faff (2010). This constant threshold uses to be the unconditional standard deviation from historical spreads, and it keeps constant across the entire trading period.

Table 3.9: The performance of pairs portfolios at different frequencies

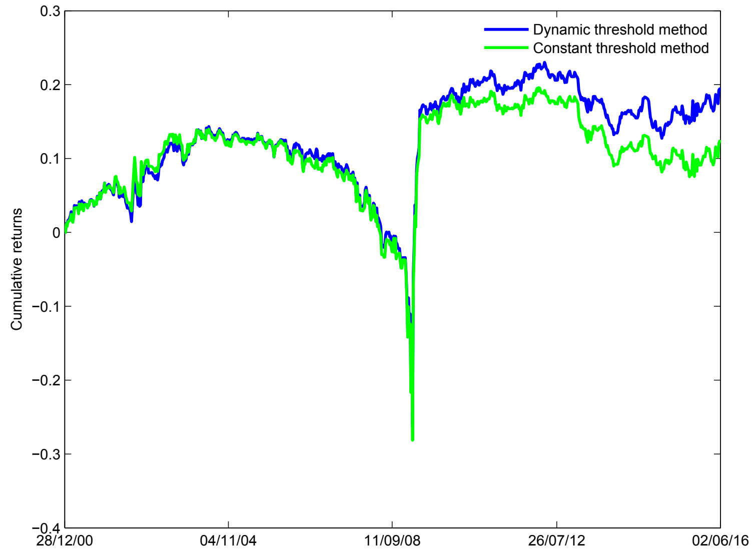
Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
			Panel A: Daily data, 6-month dynamic threshold					
Dynamic method	0.0332	0.0000	0.0748	1.12	63.84	0.07	-0.08	0.44
Constant method	0.0279	0.0000	0.0734	1.08	72.85	0.07	-0.08	0.38
			Panel B: Weekly data, 1-year dynamic threshold					
Dynamic method	0.0125	0.0000	0.0899	4.03	125.17	0.12	-0.14	0.14
Constant method	0.0079	0.0000	0.0915	3.99	123.53	0.12	-0.14	0.09

This table reports the performance of dynamic and constant threshold methods. Pairs trading with daily prices using 6-month dynamic threshold are present in Panel A, while Panel B shows results of both pairs strategies obtained under the use of weekly price data. These results are obtained from November 24th, 2000 to June 24th, 2016.

A and B, respectively. We see from Panel A that pairs trading triggered by the 6-month dynamic threshold performs well persistently, although getting a little less impressive, compared to the baseline results in Panel B of Table 1. Shortening the interval in determining a dynamic threshold doesn't bring about substantial improvement out of sample, may unveiling the usage of 1-year dynamic threshold is appropriate.

Then we look at the performance induced from weekly prices, the low-frequency data, in Panel B. We observe a substantial decline of mean return for both pairs strategies, dropping from 3.48% to 1.25% and from 2.79% to 0.79%. The downward pressure in performance is also fueled by an increase of volatility. The accompanying deterioration is unavoidable in the Sharpe ratio. This unfavorable result is to some extent within the expectation. The use of weekly prices definitely lowers the trading frequency such that much of trading opportunities cannot be detected. This would be a severe strike for pairs trading strategies dependent on a dynamic trigger to search for trading possibilities, in contrast to a buy-and-hold strategy. Low-frequency trading also leads to the rise of volatility since portfolio positions cannot be adjusted in a short time. This is supported by Figure 9 showing that cumulative returns have experienced a deep fall and bottomed in March 2009. This drop coincides with the market collapse during the financial crisis since 2008. A possible explanation is that a pair of assets

Figure 3.9: Cumulative excess returns of pairs portfolios using weekly prices



This figure displays the cumulative excess returns for the dynamic (blue line) and the constant (green line) threshold methods using weekly price data over the entire out-of-sample period.

moving along with the market requires immediate position adjustment but not allowed due to the frequency of data.

The above findings suggest that our dynamic pairs strategy works well as before, under the dynamic threshold conditional on a shorter interval. It is also worth emphasizing that the low-frequency data is not fit to pairs trading strategies.

3.6.5 Transaction costs

Previous analysis ignores transaction costs. This cost effect is the common concern on trading strategies. Pairs trading is expected to be triggered more frequently under the dynamic method than the constant one. Not only may a higher frequency on trading lead to larger returns but also transaction costs. To explore this issue, we evaluate the cost impact on trading profits following the procedure adopted in Gatev et al. (2006).

Basically, Gatev et al. (2006) compares trading profits of two trading rules, that is, open/close a position on the day of divergence/convergence versus open/close a position one day after divergence/convergence. The decrease in the excess return indicates an estimate of transaction costs incurred from trading activities. This estimation stems from the following logic. Assume the extreme case where, at opening a position, the winner trades at ask price and the loser trades at bid price. If next day prices have an identical probability of being at bid or ask, then postponing trades by one day will lower the excess returns by one-half of the sum of the bid-ask spreads of both assets. If at liquidating a position, the price of the winner is bid price and the loser is ask price, the returns will be reduced again by half of the sum of the bid-ask spreads of both assets due to one day delay of trading. In this way we estimate the transaction costs.

Table 10 exhibits the out-of-sample returns in two scenarios in Panels A and B, no waiting versus one-day delay.[12] Comparing these results, we see that the mean excess return falls from 5.09% to 3.48% for the dynamic method, whereas from 4.36% to 2.79% for the constant method. The difference in the return implies an annualized transaction cost of 1.61% for the dynamic method and 1.57% for the constant method. Then we assess if the strategy returns survive in the presence of transaction costs. If the prices, when delaying trades by one day, used to calculate the mean return of 3.48% (Panel B) have the identical probability of being at bid or ask, we have to adjust these returns to reflect that in practice we long at the ask and sell at the bid prices. To do so, we subtract trading costs of 1.61% from 3.48%, for the dynamic method, to get an estimate of returns net of transaction costs. Clearly the returns reported in Panel B cover the estimate of transaction costs, which is still higher than the constant method's return after trading costs.

In addition, we use bid and ask prices to compute bid-ask spreads as an alternative method measuring transaction costs. Data are collected from WRDS. Because we long the loser and short the winner asset, transaction costs will reduce returns by one-half of the sum of the bid-ask spreads on both assets every time there is a change in position in the pair. Panel C in Table 10 reveals that, after accounting for transaction costs, the average return goes down from 3.48% to 1.93% and from 2.79% to 1.53%, for two respective methods. The return

[12]Note that the rest of the paper reports results of pairs strategies that delay trading activities by one day.

Table 3.10: The performance of pairs portfolios considering transaction costs

Portfolio	Mean	Median	Stdev	Skew	Kurtosis	Max.	Min.	Sharpe
Panel A: Trade on day of trigger								
Dynamic method	0.0509	0.0000	0.0753	1.64	65.39	0.07	-0.08	0.68
Constant method	0.0436	0.0000	0.0748	1.63	75.70	0.07	-0.08	0.58
Panel B: Trade one day after trigger								
Dynamic method	0.0348	0.0000	0.0731	1.12	63.52	0.07	-0.08	0.48
Constant method	0.0279	0.0000	0.0734	1.08	72.85	0.07	-0.08	0.38
Panel C: Trade one day after trigger, net of bid-ask spreads								
Dynamic method	0.0193	0.0000	0.0724	1.13	67.58	0.07	-0.08	0.27
Constant method	0.0153	0.0000	0.0722	1.12	79.55	0.07	-0.08	0.21

This table reports the performance of dynamic and constant threshold methods. We trade according to the rule that positions are opened at the end of the day that the threshold criteria is triggered (Panel A). The results in Panel B correspond to a trading rule that delays the opening of positions by one day. Panel C reports results obtained by subtracting returns in Panel B by one-half of the sum of the bid-ask spreads computed with real data, on a pair of assets. These results are from November 24th, 2000 to June 24th, 2016.

gap is thus the transaction cost embedded in the trading activities, shown to be 1.55% and 1.26% per annum under the dynamic and constant methods. These recorded costs are close to the ones estimated above. This alternative analysis provides evidence supporting again the survival of pairs strategies after adjusting for trading costs.

Therefore, transaction costs would not be a major factor explaining the differences in strategy performance documented.

3.7 Conclusion

This paper examines how temporary deviations between two assets sharing cointegration relation can be exploited using pairs trading strategy. We concentrate on the supercointegrated portfolio that is established by pairs at 1% confidence level of Johansen (1991) test. The out-of-sample analysis shows the superiority of this high-quality group of pairs relative to the benchmark market index, in terms of Sharpe ratio from a portfolio perspective. We also find a positive relationship between the performance of pairs portfolio and the level of cointegration among pairs. It means that the performance of pairs portfolio improves in a monotonic pattern as the cointegration relationship gets closer. These evidence are also documented in an international context, by analyzing listed companies in the European stock market. With respect to the risk profile, we find a close connection of the time-varying risk of pairs strategy to aggregate market volatility, which are shown to be persistent. In addition, a positive risk-return relationship of the strategy is also found.

Appendices

Appendix A

A Theoretical Supply and Demand Model for Pairs Trading Dynamics

Assume that a trader has identified two financial instruments whose prices y_t and x_t are cointegrated. The underlying long term equilibrium between both markets can be specified as:

$$y_t = \gamma_0 + \gamma_1 x_t + z_t \tag{A.1}$$

This implies that the value of asset y_t can be replicated by a portfolio using asset x_t. Portfolio replication will be established on the basis of price leadership. z_t represents the stationary arbitrage opportunities in two cointegrated markets arising from market imperfections. A trader exploits temporary mispricings from equilibrium by pursuing pairs trading strategies that short sell the outperforming asset and buy the underperformer. We consider now the aggregate market demand function for all agents who perform pairs trading strategies taking simultaneous positions in y_t and x_t in period t. This is represented by:

$$\begin{aligned} &H\left((\gamma_1 x_t + \gamma_0) - y_t\right) \quad , H \succ 0 \\ &= H\left(z_t\right) \qquad\qquad\qquad , H \succ 0 \end{aligned} \tag{A.2}$$

where H is the elasticity of demand for pairs trading strategies. It increases when transaction costs are negligible, and other market imperfections decrease as vehicles for cross-market trading improve. In the limit, markets become perfectly integrated and the elasticity H tends to infinity. When transaction costs are significant and there are market restrictions that impede inter-market trading, the elasticity of demand for pairs strategies is finite.

We assume that there are N_y agents in the market for asset y_t and N_x agents in the market for asset x_t. These investors will take positions in asset y_t and asset x_t as well as pursue pairs trading in the two markets:

Let $Q_{i,t}$ be the number of shares owned by the i^{th} participant in period t and $B_{i,t}$ the bid price at which that agent is willing to hold quantity $Q_{i,t}$. Then the demand function of the i^{th} agent in the market for stock y_t in period t is

$$Q_{i,t} - A\left(y_t - B_{i,t}\right) \tag{A.3}$$

with $i = 1, ..., N_y$ where $A \succ 0$, is the demand elasticity, assumed to be the same for all market agents.

The demand function for agent j in the market for stock x_t is

$$Q_{j,t} - A\left(x_t - B_{j,t}\right), A \succ 0, j = 1, ..., N_x \tag{A.4}$$

The market for stock y_t will clear at the value of y_t that solves,

$$\sum_{i=1}^{N_y} Q_{i,t} = \sum_{i=1}^{N_y} \left(Q_{i,t} - A\left(y_t - B_{i,t}\right)\right) + H\left(\left(\gamma_1 x_t + \gamma_0\right) - y_t\right) \tag{A.5}$$

with $H \succ 0$.

The market for stock x_t will clear at the value of x_t such that:

$$\sum_{j=1}^{N_x} Q_{j,t} = \sum_{j=1}^{N_x} \left(Q_{j,t} - A\left(x_t - B_{j,t}\right)\right) + H\left(\left(\gamma_1 x_t + \gamma_0\right) - y_t\right) \tag{A.6}$$

Solving Equations (A.5) and (A.6) for y_t and x_t as a function of the mean bid price set by market agents in y_t $\left(B_t^y = N_y^{-1}\sum_{i=1}^{N_y} B_{i,t}\right)$ and the mean bid price $\left(B_t^x = N_x^{-1}\sum_{j=1}^{N_x} B_{j,t}\right)$ for market agents in x_t, we obtain:

$$\begin{aligned} y_t &= \frac{(AN_x+H\gamma_1)N_yB_t^y+HN_x\gamma_1B_t^x+HN_x\gamma_0}{(H+AN_y)N_x+HN_y\gamma_1} \\ x_t &= \frac{HN_yB_t^y+(AN_y+H)N_xB_t^x-HN_y\gamma_0}{(H+AN_y)N_x+HN_y\gamma_1} \end{aligned} \tag{A.7}$$

In what follows we derive the dynamic price relationships. This requires characterizing the model in Equation (A.7) with a description of the evolution of bid prices. It is assumed that immediately after the market clearing period $t-1$ the i^{th} agent in y_t was willing to hold a position of $Q_{i,t}$ at a price y_{t-1}. Following FFG, this implies that y_{t-1} was his bid price after that clearing. We assume that this bid price changes to $B_{i,t}$ according to the equation

$$\begin{aligned} B_{i,t} &= y_{t-1}+e_t+w_{i,t} \\ B_{j,t} &= x_{t-1}+e_t+w_{j,t} \end{aligned} \tag{A.8}$$

$$\begin{aligned} cov\left(e_t, w_{i,t}\right) &= 0, \forall i \\ cov\left(w_{i,t}, w_{j,t}\right) &= 0, \forall i \neq j \end{aligned}$$

with $i = 1, ..., N_y$ and $j = 1, ..., N_x$. Where the vector $\begin{pmatrix} e_t, & w_{i,t}, & w_{j,t} \end{pmatrix}$ is vector white noise with finite variance.

The price change $B_{i,t} - y_{t-1}$ reflects the arrival of new information between period $t-1$ and period t which changes the price at which the i^{th} participant is willing to hold a position of $Q_{i,t}$ in the market y_t. This price change has a component common to all market agents (e_t) and a component idiosyncratic to the i^{th} agent $(w_{i,t})$.

The Equations in (A.8) imply that the mean bid price in each market in period t will be

$$\begin{aligned} B_t^y &= y_{t-1}+e_t+w_t^y \\ B_t^x &= x_{t-1}+e_t+w_t^x \end{aligned} \tag{A.9}$$

where $w_t^y = \frac{\sum_{i=1}^{N_y} w_{i,t}^y}{N_y}$ and $w_t^x = \frac{\sum_{j=1}^{N_x} w_{j,t}^x}{N_x}$. Substituting expressions (A.9) into (A.7) yields the following vector model:

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \frac{H\gamma_0}{d}\begin{pmatrix} N_x \\ -N_y \end{pmatrix} + M\begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_t^y \\ u_t^x \end{pmatrix} \tag{A.10}$$

where

$$\begin{pmatrix} u_t^y \\ u_t^x \end{pmatrix} = M\begin{pmatrix} e_t + w_t^y \\ e_t + w_t^x \end{pmatrix} \tag{A.11}$$

$$M = \frac{1}{d}\begin{bmatrix} N_y\left(\gamma_1 H + AN_x\right) & \gamma_1 HN_x \\ HN_y & \left(H + AN_y\right)N_x \end{bmatrix} \tag{A.12}$$

And

$$d = \left(H + AN_y\right)N_x + \gamma_1 HN_y \tag{A.13}$$

We next convert Equation (A.10) into a Vector Error Correction Model (VECM) by subtracting $(y_{t-1}, x_{t-1})'$ from both sides, with

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = \frac{H\gamma_0}{d}\begin{pmatrix} N_x \\ -N_y \end{pmatrix} + (M - I)\begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_t^y \\ u_t^x \end{pmatrix} \tag{A.14}$$

$$M - I = \frac{1}{d}\begin{bmatrix} -HN_x & \gamma_1 HN_x \\ HN_y & -HN_y\gamma_1 \end{bmatrix} \tag{A.15}$$

Rearranging terms,

$$\begin{pmatrix} \Delta y_t \\ \Delta x_t \end{pmatrix} = \frac{H}{d}\begin{pmatrix} -N_x \\ N_y \end{pmatrix}\begin{pmatrix} 1 & -\gamma_1 & -\gamma_0 \end{pmatrix}\begin{pmatrix} y_{t-1} \\ x_{t-1} \\ 1 \end{pmatrix} + \begin{pmatrix} u_t^y \\ u_t^x \end{pmatrix} \tag{A.16}$$

Appendix B

Optimal Portfolio Holdings in Continuous Time

Suppose the portfolio consists of two risky assets and a risk-free asset. We denote φ_1 and φ_2, as the number of shares held respectively by the investor in the risky assets y_t and x_t at time t, while the number of the risk-free asset is given by φ_3. Let Π_t be the portfolio value, represented by:

$$\Pi_t = \varphi_1 y_t + \varphi_2 x_t + \varphi_3 B_t \tag{B.1}$$

Portfolio dynamics are therefore given by:

$$\begin{aligned} d\Pi_t &= \varphi_1 dy_t + \varphi_2 dx_t + \varphi_3 dB_t \\ &= \left[r\Pi_t + \varphi_1 y_t(\mu_y - r + \lambda_1 z_t) + \varphi_2 x_t(\mu_x - r - \lambda_2 z_t) \right] dt + \varphi_1 \sigma_y y_t dW_y + \varphi_2 \sigma_x x_t dW_x \end{aligned} \tag{B.2}$$

We suppose that the investor maximizes the expected portfolio value. Then the value function that solves this maximization problem, $V(t, \Pi, y, x)$, is defined as:

$$V(t, \Pi, y, x) = \underset{\varphi_1, \varphi_2}{max} E_t \left[\Pi_T \right]$$

where $t \in [0, T]$. By Ito's Lemma, we get:

$$dV = V_{\Pi}d\Pi + V_y dy + V_x dx + \frac{1}{2}V_{\Pi\Pi}d\Pi^2 + \frac{1}{2}V_{yy}dy^2 + \frac{1}{2}V_{xx}dx^2 + V_{\Pi y}d\Pi dy + V_{\Pi x}d\Pi dx + V_{yx}dydx - V_t dt$$

Suppose correlation exists between two Brownian motions, which is given by

$$E[dW_y dW_x] = \rho_{xy}dt \; where \, \rho_{xy} \in (-1,1) \tag{B.3}$$

The Hamilton–Jacobi–Bellman (HJB) equation for the problem, $E_t[dV(\cdot)] = A$, therefore becomes:

$$\begin{aligned} A \;=\; & \underset{\varphi_1,\varphi_2}{max} V_{\Pi}[r\Pi_t + \varphi_1 y(\mu_y - r + \lambda_1 z_t) + \varphi_2 x(\mu_x - r - \lambda_2 z_t)] \\ & + V_y[\mu_y y + \lambda_1 y z_t] + V_x[\mu_x x - \lambda_2 x z_t] \\ & + \frac{1}{2}V_{\Pi\Pi}(\varphi_1^2\sigma_y^2 y^2 + \varphi_2^2\sigma_x^2 x^2 + 2\varphi_1\varphi_2\sigma_y y\sigma_x x\rho_{xy}) \\ & + \frac{1}{2}V_{yy}\sigma_y^2 y^2 + \frac{1}{2}V_{xx}\sigma_x^2 x^2 + V_{\Pi y}(\varphi_1\sigma_y^2 y^2 + \varphi_2\sigma_y y\sigma_x x\rho_{xy}) \\ & + V_{\Pi x}(\varphi_2\sigma_x^2 x^2 + \varphi_1\sigma_y y\sigma_x x\rho_{xy}) + V_{yx}\sigma_y y\sigma_x x\rho_{xy} - rBV_{\tau} \end{aligned}$$

The first-order conditions for φ_1 and φ_2 are:

$$\frac{\partial A}{\partial \varphi_1} = V_{\Pi}[y(\mu_y - r + \lambda_1 z_t)] + V_{\Pi\Pi}(\varphi_1\sigma_y^2 y^2 + \varphi_2\sigma_y y\sigma_x x\rho_{xy}) + V_{\Pi y}\sigma_y^2 y^2 + V_{\Pi x}\sigma_y y\sigma_x x\rho_{xy} = 0$$

$$\frac{\partial A}{\partial \varphi_2} = V_{\Pi}[x(\mu_x - r - \lambda_2 z_t)] + V_{\Pi\Pi}(\varphi_2\sigma_x^2 x^2 + \varphi_1\sigma_y y\sigma_x x\rho_{xy}) + V_{\Pi y}\sigma_y y\sigma_x x\rho_{xy} + V_{\Pi x}\sigma_x^2 x^2 = 0$$

Solving the two equations above, we get the optimal portfolio holdings φ_1^* and φ_2^*, under the condition that correlation is imposed:

$$\begin{aligned}
\varphi_1^* &= -\frac{V_\Pi}{V_{\Pi\Pi}} \cdot \frac{1}{\sigma_y y(1-\rho_{xy}^2)} \cdot [\frac{(\mu_y - r) + \lambda_1 z_t}{\sigma_y} - \frac{(\mu_x - r) - \lambda_2 z_t}{\sigma_x}\rho_{xy}] - \frac{V_{\Pi y}}{V_{\Pi\Pi}} \\
&= -\frac{V_\Pi}{V_{\Pi\Pi}} \cdot \frac{\theta_y - \theta_x \rho_{xy}}{\sigma_y y(1-\rho_{xy}^2)} - \frac{V_{\Pi y}}{V_{\Pi\Pi}} \qquad \text{(B.4)}
\end{aligned}$$

$$\begin{aligned}
\varphi_2^* &= -\frac{V_\Pi}{V_{\Pi\Pi}} \cdot \frac{1}{\sigma_x x(1-\rho_{xy}^2)} \cdot [\frac{(\mu_x - r) - \lambda_2 z_t}{\sigma_x} - \frac{(\mu_y - r) + \lambda_1 z_t}{\sigma_y}\rho_{xy}] - \frac{V_{\Pi x}}{V_{\Pi\Pi}} \\
&= -\frac{V_\Pi}{V_{\Pi\Pi}} \cdot \frac{\theta_x - \theta_y \rho_{xy}}{\sigma_x x(1-\rho_{xy}^2)} - \frac{V_{\Pi x}}{V_{\Pi\Pi}} \qquad \text{(B.5)}
\end{aligned}$$

where $\theta_y = \frac{(\mu_y - r) + \lambda_1 z_t}{\sigma_y}$ and $\theta_x = \frac{(\mu_x - r) - \lambda_2 z_t}{\sigma_x}$ are the excess returns for y_t and x_t, respectively.

Appendix C

The Martingale Argument

We prove that the portfolio Π_t made of two cointegrated equities and a risk-free bond, is a martingale.
As defined above, the dynamic of portfolio value can be written as:

$$d\Pi_t = \varphi_{1t} dy_t + \varphi_{2t} dx_t + \varphi_{3t} dB$$

so that

$$\begin{aligned} d(e^{-rt}\Pi_t) &= \varphi_{1t} d(e^{-rt} y_t) + \varphi_{2t} d(e^{-rt} x_t) + \varphi_{3t} d(e^{-rt} B) \\ &= \varphi_{1t} e^{-rt} \sigma_y y_t (\theta^p_{y,t} dt + dW_{y,t}) + \varphi_{2t} e^{-rt} \sigma_x x_t (\theta^p_{x,t} dt + dW_{x,t}) \end{aligned}$$

According to the probability measure $\mathbb{Q}$ defined under the Girsanov theorem,

$$\widetilde{W}_t = W_t + \int_0^t \theta^p_u du \tag{C.1}$$

therefore,

$$d(e^{-rt}\Pi_t) = \varphi_{1t} e^{-rt} \sigma_y y_t d\widetilde{W}_{y,t} + \varphi_{2t} e^{-rt} \sigma_x x_t d\widetilde{W}_{x,t} \tag{C.2}$$

We call $\mathbb{Q}$ the risk-neutral measure because it renders the discounted stock price $e^{-rt}\Pi_t$ into a martingale. Therefore,

$$e^{-rt}\Pi_t = \varphi_{1t}(y_0 + \int_0^t e^{-rt} \sigma_y y_u d\widetilde{W}_{y,u}) + \varphi_{2t}(x_0 + \int_0^t e^{-rt} \sigma_x x_u d\widetilde{W}_{x,u})$$

where under $\mathbb{Q}$ the process $\int_0^t e^{-rt} \sigma_y y_u d\widetilde{W}_{y,u}$ and $\int_0^t e^{-rt} \sigma_x x_u d\widetilde{W}_{x,u}$ are a martingale.

We substitute $dW_{y,t} = -\theta^p_{y,t} dt + d\widetilde{W}_{y,t}$ back to dy_t and $dW_{x,t} = -\theta^p_{x,t} dt + d\widetilde{W}_{x,t}$ back to dx_t, and thus obtain:

$$d\Pi_t = r\Pi_t dt + \varphi_{1t}\sigma_y y_t d\tilde{W}_{y,t} + \varphi_{2t}\sigma_x x_t d\tilde{W}_{x,t} \tag{C.3}$$

Under this specification, the undiscounted portfolio value Π_t has mean rate of return equal to the interest rate under the measure $\mathbb{Q}$.

It is clearly underlined that the investor has two asset categories: (1) a risk-free bond with rate of return r, and (2) two risky assets with mean rate of return r under $\mathbb{Q}$. The mean rate of return for the formed portfolio will be the risk-free rate r under $\mathbb{Q}$, and thus the discounted value of the portfolio, $e^{-rt}\Pi_t$, will be a martingale. This is true regardless how investor allocates money to these investment options.

Appendix D

An Option to Exchange One Asset for Another

The payoff function for this option at maturity T is $\psi(y_T, x_T) = max(y_T - x_T, 0)$, so the spread option can be defined as

$$\begin{aligned} C(y,x,t) &= e^{-r(T-t)} E^{\mathbb{Q}} \left[Max(y_T - x_T, 0) | y_t = y, \, x_t = x \right] \\ &= E^{\mathbb{Q}} \left[x_t (\frac{y_t}{x_t} - 1)^+ | y_t = y, \, x_t = x \right] \\ &= E^{\mathbb{Q}} \left[x_t S(\zeta, t) \mid y_t = y, \, x_t = x \right] \end{aligned} \tag{D.1}$$

where $\zeta = \frac{y_t}{x_t}$.

Then it follows

$$\frac{\partial C}{\partial y} = x_t \frac{\partial S}{\partial \zeta} \frac{1}{x_t} = \frac{\partial S}{\partial \zeta}, \;\; \frac{\partial^2 C}{\partial y^2} = \frac{\partial^2 S}{\partial \zeta^2} \frac{1}{x_t}$$

$$\frac{\partial C}{\partial x} = S + x_t \frac{\partial S}{\partial \zeta} \left(-\frac{y_t}{x_t^2} \right) = (S - \zeta \frac{\partial S}{\partial \zeta})$$

$$\frac{\partial^2 C}{\partial x^2} = \left(\frac{\partial S}{\partial \zeta} \left(-\frac{y_t}{x_t^2} \right) + \left(\frac{y_t}{x_t^2} \right) \frac{\partial S}{\partial \zeta} - \zeta \frac{\partial^2 S}{\partial \zeta^2} \left(-\frac{y_t}{x_t^2} \right) \right) = \frac{1}{x_t} \zeta^2 \frac{\partial^2 S}{\partial \zeta^2}$$

$$\frac{\partial^2 C}{\partial y \partial x} = -\frac{y_t}{x_t^2} \frac{\partial^2 S}{\partial \zeta^2}, \;\; \frac{\partial C}{\partial t} = x_t \frac{\partial S}{\partial t}$$

The partial differential equation (2.16) becomes

$$\frac{\partial S}{\partial t} + \frac{1}{2}\sigma_z^2 \zeta^2 \frac{\partial^2 S}{\partial \zeta^2} = 0 \tag{D.2}$$

This equation is just the Black-Scholes equation for a single asset with $r = 0$ and with a volatility of σ_z. Note that

$$S(\zeta, T) = (\zeta - 1)^+ \tag{D.3}$$

which follows that

$$C(y_t, x_t, t) = x_t \left[\zeta \Phi(d_1) - \Phi(d_2)\right] = y_t \Phi(d_1) - x_t \Phi(d_2) \tag{D.4}$$

where

$$d_1 = \frac{ln(\frac{y_t}{x_t}) + \frac{1}{2}\sigma_z^2 (T-t)}{\sigma_z \sqrt{T-t}}, \;\; d_2 = d_1 - \sigma_z \sqrt{T-t}$$

$$\sigma_z = \sqrt{\sigma_y^2 - 2\rho\sigma_y\sigma_x + \sigma_x^2}$$

This is the Margrabe (1978)´s result.

Appendix E

Maximum Likelihood Estimation of Continuous VECM

As we have defined, the prices of y_t and x_t follow the dynamic process:

$$\frac{dy_t}{y_t} = \mu_y dt + \lambda_1 z_t dt + \sigma_y dW_{y,t} \tag{E.1}$$

$$\frac{dx_t}{x_t} = \mu_x dt - \lambda_2 z_t dt + \sigma_x dW_{x,t} \tag{E.2}$$

Let $Y_t = lny_t$ and $X_t = lnx_t$. We define the spread as the difference between the log of the two asset prices, $z_t = \ln y_t - \ln x_t = Y_t - X_t$. Ito's lemma implies that

$$dY_t = \left(\mu_y + \lambda_1 z_t - \frac{1}{2}\sigma_y^2\right) dt + \sigma_y dW_{y,t} \tag{E.3}$$

$$dX_t = \left(\mu_x - \lambda_2 z_t - \frac{1}{2}\sigma_x^2\right) dt + \sigma_x dW_{y,t} \tag{E.4}$$

If observations are spaced Δt apart, the Euler approximation of (E.3) and (E.4) are given by

$$Y(t_i) - Y(t_{i-1}) = \left(\mu_y + \lambda_1(Y(t_{i-1}) - X(t_{i-1})) - \frac{1}{2}\sigma_y^2\right) \Delta t + \sigma_y \sqrt{\Delta t} Z_y \tag{E.5}$$

$$X(t_i) - X(t_{i-1}) = \left(\mu_x - \lambda_2(Y(t_{i-1}) - X(t_{i-1})) - \frac{1}{2}\sigma_x^2\right)\Delta t + \sigma_x\sqrt{\Delta t}Z_x \tag{E.6}$$

where Z_y and Z_x are two standard normal random variables with correlation ρ_{xy}.

Thus, the transition density to $(Y(t_i),\, X(t_i)) = (\ln y(t_i),\, \ln x(t_i))$ given previous observations $(Y(t_{i-1}),\, X(t_{i-1}))$ is written as

$$\begin{aligned} & f\left(Y(t_i),\ X(t_i) \mid Y(t_{i-1}),\ X(t_{i-1});\, \mu_y, \mu_x, \lambda_1, \lambda_2, \sigma_y, \sigma_x, \rho_{xy}\right) \\ = \;& f\left(Y(t_i),\ X(t_i) \mid Y(t_{i-1}),\ X(t_{i-1});\, \Theta\right) \\ = \;& \mid \det(J_i) \mid \cdot \phi(Z_y,\, Z_x; \Sigma) \end{aligned} \tag{E.7}$$

where

$$Z_y = \frac{Y(t_i) - Y(t_{i-1}) - \left(\mu_y + \lambda_1(Y(t_{i-1}) - X(t_{i-1})) - \frac{1}{2}\sigma_y^2\right)\Delta t}{\sigma_y\sqrt{\Delta t}} \sim N(0,1) \tag{E.8}$$

$$Z_x = \frac{X(t_i) - X(t_{i-1}) - \left(\mu_x - \lambda_2(Y(t_{i-1}) - X(t_{i-1})) - \frac{1}{2}\sigma_x^2\right)\Delta t}{\sigma_x\sqrt{\Delta t}} \sim N(0,1) \tag{E.9}$$

$$\Sigma = \begin{pmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{pmatrix} \tag{E.10}$$

$\phi(Z_y,\, Z_x; \Sigma)$ is the probability density function of bivariate normal distribution with zero mean and covariance matrix Σ, and $\det(J)$ is the Jacobian determinant.

Next we calculate the Jacobian,

$$\det(J) = \det\begin{bmatrix} \frac{\partial Z_y}{\partial Y(t_i)} & \frac{\partial Z_y}{\partial X(t_i)} \\ \frac{\partial Z_x}{\partial Y(t_i)} & \frac{\partial Z_x}{\partial X(t_i)} \end{bmatrix} = \frac{1}{\sigma_y\sigma_x\Delta t} \tag{E.11}$$

The joint density of Z_y and Z_x is given by

$$\phi(Z_y,\, Z_x; \Sigma) = \frac{1}{2\pi}(\det\Sigma)^{-1/2}\exp\left[-\frac{1}{2}\left(Z_y,\, Z_x\right)\Sigma^{-1}\begin{pmatrix} Z_y \\ Z_x \end{pmatrix}\right]$$

where

$\det \Sigma = 1 - \rho_{xy}^2, \qquad \Sigma^{-1} = \frac{1}{1-\rho_{xy}^2} \begin{pmatrix} 1 & -\rho_{xy} \\ -\rho_{xy} & 1 \end{pmatrix}.$

Therefore we obtain

$$\phi(Z_y,\, Z_x; \Sigma) = \frac{1}{2\pi\sqrt{1-\rho_{xy}^2}} \exp\left[-\frac{1}{2(1-\rho_{xy}^2)}\left(Z_y^2 + Z_x^2 - 2\rho_{xy} Z_y Z_x\right)\right] \tag{E.12}$$

Then put the expressions (E.11) and (E.12) into the transition density (E.7), we have

$$f\left(Y(t_i),\; X(t_i) \mid Y(t_{i-1}),\; X(t_{i-1}); \Theta\right) = \frac{1}{2\pi\sigma_y\sigma_x\Delta t\sqrt{1-\rho_{xy}^2}} \exp\left[-\frac{1}{2(1-\rho_{xy}^2)}\left(Z_y^2 + Z_x^2 - 2\rho_{xy} Z_y Z_x\right)\right] \tag{E.13}$$

Therefore, the parameters can be estimated by using the method of maximum likelihood

$$\Theta^* = argmax \left\{ \prod_{i=1}^{n} f\left(Y(t_i),\; X(t_i) \mid Y(t_{i-1}),\; X(t_{i-1}); \Theta\right) \right\} \tag{E.14}$$

where the joint likelihood is

$$\begin{aligned} \mathcal{L}(\Theta) &= \prod_{i=1}^{n} f\left(Y(t_i),\; X(t_i) \mid Y(t_{i-1}),\; X(t_{i-1}); \Theta\right) \\ &= (2\pi\sigma_y\sigma_x\Delta t)^{-n} \cdot (1-\rho_{xy}^2)^{-n/2} \cdot \exp\left[-\frac{1}{2(1-\rho_{xy}^2)} \sum_{i=1}^{n} \left(Z_y^2 + Z_x^2 - 2\rho_{xy} Z_y Z_x\right)\right] \end{aligned}$$

In this case, the natural logarithm of the likelihood function is

$$\ln \mathcal{L}(\Theta) = -n \ln(2\pi\sigma_y\sigma_x\Delta t) - \frac{n}{2}\ln(1-\rho_{xy}^2) - \frac{1}{2(1-\rho_{xy}^2)} \sum_{i=1}^{n} \left(Z_y^2 + Z_x^2 - 2\rho_{xy} Z_y Z_x\right) \tag{E.15}$$

therefore we get

$$\hat{\mu_y} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Y(t_i)-Y(t_{i-1})}{\Delta t} - \hat{\lambda_1}(Y(t_{i-1})-X(t_{i-1})) + \frac{1}{2}\hat{\sigma}_y^2\right) \tag{E.16}$$

$$\hat{\mu_x} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X(t_i)-X(t_{i-1})}{\Delta t} + \hat{\lambda_2}(Y(t_{i-1})-X(t_{i-1})) + \frac{1}{2}\hat{\sigma}_x^2\right) \tag{E.17}$$

$$\hat{\lambda_1} = \frac{2\sum_{i=1}^{n}\left(A(t_i)-\frac{1}{n}\sum_{i=1}^{n}A(t_i)\right)\left(z(t_{i-1})-\frac{1}{n}\sum_{i=1}^{n}z(t_{i-1})\right)}{\Delta t\cdot\sum_{i=1}^{n}\left(z(t_{i-1})-\frac{1}{n}\sum_{i=1}^{n}z(t_{i-1})\right)^2} \tag{E.18}$$

$$\hat{\lambda_2} = \frac{2\sum_{i=1}^{n}\left(B(t_i)-\frac{1}{n}\sum_{i=1}^{n}B(t_i)\right)\left(z(t_{i-1})-\frac{1}{n}\sum_{i=1}^{n}z(t_{i-1})\right)}{\Delta t\cdot\sum_{i=1}^{n}\left(z(t_{i-1})-\frac{1}{n}\sum_{i=1}^{n}z(t_{i-1})\right)^2} \tag{E.19}$$

$$\hat{\sigma}_y = \sqrt{\frac{1}{n\Delta t}\cdot\sum_{i=1}^{n}\left(\left(A(t_i)-\frac{1}{n}\sum_{i=1}^{n}A(t_i)\right)-\hat{\lambda_1}\Delta t\left(z(t_{i-1})-\frac{1}{n}\sum_{i=1}^{n}z(t_{i-1})\right)\right)^2} \tag{E.20}$$

$$\hat{\sigma}_x = \sqrt{\frac{1}{n\Delta t}\cdot\sum_{i=1}^{n}\left(\left(B(t_i)-\frac{1}{n}\sum_{i=1}^{n}B(t_i)\right)-\hat{\lambda_2}\Delta t\left(z(t_{i-1})-\frac{1}{n}\sum_{i=1}^{n}z(t_{i-1})\right)\right)^2} \tag{E.21}$$

$$\hat{\rho_{xy}} = \frac{1}{n}\sum_{i=1}^{n}(Z_yZ_x) \tag{E.22}$$

where $A(t_i) = Y(t_i) - Y(t_{i-1})$, $\quad B(t_i) = X(t_i) - X(t_{i-1})$, $\quad z(t_{i-1}) = Y(t_{i-1}) - X(t_{i-1})$.

Bibliography

Admati, A. R., Pfleiderer, P., 1988. A Theory of Intraday Patterns: Volume and Price Variability. Review of Financial studies 1 (1), 3–40.

Alexander, C., Giblin, I., III, W. W., 2002. Cointegration and Asset Allocation: A New Active Hedge Fund Strategy. Research in International Business and Finance 16 (5), 65–90.

Andrade, S., Di Pietro, V., Seasholes, M., 2005. Understanding The Profitability of Pairs Trading. Working Paper, UC Berkeley, and Northwestern University.

Ankirchner, S., Dimitroff, G., Heyne, G., Pigorsch, C., 2012. Futures Cross-Hedging With A Stationary Basis. Journal of Financial and Quantitative Analysis 47 (06), 1361–1395.

Asness, C. S., Moskowitz, T. J., Pedersen, L. H., 2013. Value and Momentum Everywhere. The Journal of Finance 68 (3), 929–985.

Avellaneda, M., Lee, J., 2008. Statistical Arbitrage in the US Equities Market. Quantitative Finance 10 (7), 761–782.

Bailey, D. H., Borwein, J. M., Lopez de Prado, M., Zhu, Q. J., 2014. The Probability of Backtest Overfitting. Available at SSRN 2326253.

Bailey, D. H., Lopez de Prado, M., 2014. The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting and Non-Normality. Journal of Portfolio Management, Forthcoming.

Bali, T. G., Brown, S. J., Demirtas, K. O., 2013. Do Hedge Funds Outperform Stocks and Bonds? Management Science 59 (8), 1887–1903.

Barroso, P., Santa-Clara, P., 2015. Momentum Has Its Moments. Journal of Financial Economics 116, 111–120.

Basak, S., Chabakauri, G., 2010. Dynamic Mean-Variance Asset Allocation. Review of Financial Studies, hhq028.

Black, F., Scholes, M., 1973. The Pricing of Options and Corporate Liabilities. The journal of political economy, 637–654.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. Journal of econometrics 31 (3), 307–327.

Bowen, D. A., Hutchinson, M. C., 2014. Pairs Trading in the UK Equity Market: Risk and Return. The European Journal of Finance, 1–25.

Brennan, M. J., Jegadeesh, N., Swaminathan, B., 1993. Investment Analysis and The Adjustment of Stock Prices to Common Information. Review of Financial Studies 6 (4), 799–824.

Brennan, M. J., Wang, A. W., 2010. The Mispricing Return Premium. Review of Financial Studies 23 (9), 3437–3468.

Brennan, M. J., Xia, Y., 2002. Dynamic Asset Allocation under Inflation. The Journal of Finance 57 (3), 1201–1238.

Brock, W., Lakonishok, J., LeBaron, B., 1992. Simple Technical Trading Rules and the Stochastic Properties of Stock Returns. The Journal of finance 47 (5), 1731–1764.

Broussard, J. P., Vaihekoski, M., 2012. Profitability of Pairs Trading Strategy in An Illiquid Market with Multiple Share Classes. Journal of International Financial Markets, Institutions and Money 22 (5), 1188–1201.

Brown, S. J., Fraser, T. L., Liang, B., 2008. Hedge Funds Due Diligence: A Source of Alpha in a Hedge Fund Portfolio Strategy. Journal of Investment Management 6, 23–33.

Campbell, J. Y., Thompson, S. B., 2008. Predicting Excess Stock Returns Out Of Sample: Can Anything Beat the Historical Average? Review of Financial Studies 21 (4), 1509–1531.

Cartea, A., Jaimungal, S., Penalva, J., 2015. Algorithmic and High-Frequency Trading. Cambridge University Press.

Chen, H., Chen, S., Li, F., 2012. Empirical Investigation of An Equity Pairs Trading Strategy. Working Paper, Texas A & M University, University of British Columbia, and University of Michigan.

Chhaochharia, V., Grinstein, Y., 2007. Corporate Governance and Firm Value: The Impact of the 2002 Governance Rules. The Journal of Finance 62 (4), 1789–1825.

Chordia, T., Swaminathan, B., 2000. Trading Volume and Cross-autocorrelations in Stock Returns. The Journal of Finance 55 (2), 913–935.

Cumming, D., Johan, S., Li, D., 2011. Exchange Trading Rules and Stock Market Liquidity. Journal of Financial Economics 99 (3), 651–671.

Daniel, K., Moskowitz, T., 2016. Momentum Crashes. Journal of Financial Economics 122 (2), 221–247.

DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal Versus Naive Diversification: How Inefficient is the 1/n Portfolio Strategy? Review of Financial Studies 22 (5), 1915–1953.

Do, B., Faff, R., 2010. Does Simple Pairs Trading Still Work? Financial Analysts Journal 66 (4), 83–95.

Do, B., Faff, R., Hamza, K., 2006. A New Approach to Modeling and Estimation for Pairs Trading. In: Proceedings of 2006 Financial Management Association European Conference. Citeseer, pp. 87–99.

Duan, J.-C., Pliska, S. R., 2004. Option Valuation with Co-integrated Asset Prices. Journal of Economic Dynamics and Control 28 (4), 727–754.

Dunis, C., Ho, R., 2016. Asset Management. Springer International Publishing, Ch. Cointegration Portfolios of European Equities for Index Tracking and Market Neutral Strategies, pp. 191–223.

Dunis, C., Lequeux, P., 2000. Intraday Data and Hedging Efficiency in Interest Spread Trading. The European Journal of Finance 6 (4), 332–352.

Elliott, R. J., Van Der Hoek, J., Malcolm, W. P., 2005. Pairs Trading. Quantitative Finance 5 (3), 271–276.

Engelberg, J., Gao, P., Jagannathan, R., 2009. An Anatomy of Pairs Trading: The Role of Idiosyncratic News, Common Information and Liquidity. Working Paper, University of North Carolina, University of Notre Dame, and Northwestern University.

Engle, R. F., Granger, C. W., 1987. Cointegration and Error Correction: Representation, Estimation, and Testing. Econometrica: Journal of the Econometric Society, 251–276.

Fama, E. F., French, K. R., 1992. The Cross-Section of Expected Stock Returns. Journal of Finance 47 (2), 427–465.

Fama, E. F., French, K. R., 1993. Common Risk Factors in the Returns on Stocks and Bonds. Journal of financial economics 33 (1), 3–56.

Fama, E. F., French, K. R., 1996. Multifactor Explanations of Asset Pricing Anomalies. The journal of finance 51 (1), 55–84.

Figuerola-Ferretti, I., Gonzalo, J., 2010. Modelling and Measuring Price Discovery in Commodity Markets. Journal of Econometrics 158 (1), 95–107.

Figuerola-Ferretti, I., Paraskevopoulos, I., Tang, T., 2016. Pairs Trading and Spread persistence in the European Stock Market. Working Paper.

Gatev, E., Goetzmann, W., Rouwenhorst, K., 2006. Pairs Trading: Performance of A Relative-Value Arbitrage Rule. Review of Financial Studies 19 (3), 797–827.

Goetzmann, W., Ingersoll, J., Spiegel, M., Welch, I., 2007. Portfolio Performance Manipulation and Manipulation-Proof Performance Measures. Review of Financial Studies 20 (5), 1503–1546.

Goetzmann, W. N., Kumar, A., 2008. Equity Portfolio Diversification. Review of Finance 12 (3), 433–463.

Gonzalo, J., Granger, C., 1995. Estimation of Common Long-Memory Components in Cointegrated Systems. Journal of Business & Economic Statistics 13 (1), 27–35.

Gromb, D., Vayanos, D., 2002. Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs. Journal of Financial Economics 66 (2), 361–407.

Harvey, C. R., Liu, Y., 2014. Backtesting. Available at SSRN 2345489.

Hasbrouck, J., 1995. One Security, Many Markets: Determining the Contributions to Price discovery. The Journal of Finance 50 (4), 1175–1199.

Jackwerth, J. C., Slavutskaya, A., 2016. The Total Benefit of Alternative Assets to Pension Fund Portfolios. Journal of Financial Markets 31, 25–42.

Johansen, S., 1991. Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. Econometrica: Journal of the Econometric Society, 1551–1580.

Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. New York.

Jurek, J., Yang, H., 2007. Dynamic Portfolio Selection in Arbitrage. In: EFA 2006 Meetings Paper.

Juselius, K., 2006. The Cointegrated VAR Model: Methodology and Applications. Oxford University Press.

Kapadia, N., Pu, X., 2012. Limited Arbitrage between Equity and Credit Markets. Journal of Financial Economics 105 (3), 542–564.

Kim, T. S., Omberg, E., 1996. Dynamic Nonmyopic Portfolio Behavior. Review of financial studies 9 (1), 141–161.

Kondor, P., 2009. Risk in Dynamic Arbitrage: The Price Effects of Convergence Trading. The Journal of Finance 64 (2), 631–655.

Kondor, P., Vayanos, D., 2014. Liquidity Risk And The Dynamics of Arbitrage Capital. Tech. rep., National Bureau of Economic Research.

Lehmann, B. N., 2002. Some Desiderata for the Measurement of Price Discovery Across Markets. Journal of Financial Markets 5 (3), 259–276.

Lei, Y., Xu, J., 2015. Costly Arbitrage Through Pairs Trading. Journal of Economic Dynamics and Control 56, 1–19.

Lin, Y.-X., McCrae, M., Gulati, C., 2006. Loss Protection in Pairs Trading through Minimum Profit Bounds: A Cointegration Approach. Journal of Applied Mathematics and Decision Sciences 2006.

Liu, J., 2007. Portfolio Selection in Stochastic Environments. Review of Financial Studies 20 (1), 1–39.

Liu, J., Longstaff, F. A., 2004. Losing Money on Arbitrage: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities. Review of Financial Studies 17 (3), 611–641.

Liu, J., Timmermann, A., 2013. Optimal Convergence Trade Strategies. Review of Financial Studies 26 (4), 1048–1086.

Llorente, G., Michaely, R., Saar, G., Wang, J., 2002. Dynamic Volume-Return Relation of Individual Stocks. Review of Financial studies 15 (4), 1005–1047.

Lo, A. W., Wang, J., 1995. Implementing Option Pricing Models When Asset Returns Are Predictable. The Journal of Finance 50 (1), 87–129.

Margrabe, W., 1978. The Value of An Option to Exchange One Asset for Another. The journal of finance 33 (1), 177–186.

Marquering, W., Verbeek, M., 2004. The Economic Value of Predicting Stock Index Returns And Volatility. Journal of Financial and Quantitative Analysis 39 (02), 407–429.

McMillan, D. G., Ülkü, N., 2009. Persistent Mispricing in A Recently Opened Emerging Index Futures Market: Arbitrageurs Invited. Journal of Futures Markets 29 (3), 218–243.

Merton, R. C., 1971. Optimum Consumption and Portfolio Rules in a Continuous-time Model. Journal of Economic Theory 3 (4), 373–413.

Papadakis, G., Wysocki, P., 2007. Pairs Trading and Accounting Information. Working Paper, Boston University and MIT.

Park, T. H., Switzer, L. N., 1996. Mean Reversion of Interest-Rate Term Premiums and Profits from Trading Strategies with Treasury Futures Spreads. Journal of Futures Markets 16 (3), 331–352.

Puspaningrum, H., Lin, Y.-X., Gulati, C. M., 2010. Finding the Optimal Pre-set Boundaries for Pairs Trading Strategy Based on Cointegration Technique. Journal of Statistical Theory and Practice 4 (3), 391–419.

Rad, H., Low, R. K. Y., Faff, R., 2016. The Profitability of Pairs Trading Strategies: Distance, Cointegration and Copula Methods. Quantitative Finance, 1–18.

Schaefer, S. M., Strebulaev, I. A., 2008. Structural Models of Credit Risk are Useful: Evidence from Hedge Ratios on Corporate Bonds. Journal of Financial Economics 90 (1), 1–19.

Shleifer, A., Vishny, R. W., 1997. The Limits of Arbitrage. The Journal of Finance 52 (1), 35–55.

Stulz, R., 1982. Options on the Minimum Or the Maximum of Two Risky Assets: Analysis and Applications. Journal of Financial Economics 10 (2), 161–185.

Suen, R. M., 2009. Bounding The CRRA Utility Functions.

Tu, J., Zhou, G., 2011. Markowitz Meets Talmud: A Combination of Sophisticated and Naive Diversification Strategies. Journal of Financial Economics 99 (1), 204–215.

Uhlenbeck, G. E., Ornstein, L. S., 1930. On the Theory of the Brownian Motion. Physical review 36 (5), 823.

Vidyamurthy, G., 2004. Pairs Trading: Quantitative Methods and Analysis. Vol. 217. John Wiley & Sons.

Xiong, W., 2001. Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets. Journal of Financial Economics 62 (2), 247–292.