Pairs Trading and Relative Liquidity in the European Stock Market

Isabel Figuerola-Ferretti
Business Department, Universidad Carlos III de Madrid*

Ioannis Paraskevopoulos
Bankia.†

Tao Tang
Business Department, Universidad Carlos III de Madrid‡

November 12, 2014

Abstract
This paper uses an equilibrium demand and supply framework to describe the no-arbitrage relationship between two distinct but cointegrated assets, showing how adjustment dynamics between them can be exploited through pairs trading strategies to gain arbitrage profits out of temporary mispricings. In this framework, the two close substitutes are cointegrated meaning that they measure a common non-stationary factor. Price discovery is determined by the relative number of participants or relative liquidity in both markets. This theoretical model, which builds on Figuerola and Gonzalo (Journal of Econometrics 2010) is applied to all the EURO STOXX 50 traded equities and pairs equity portfolios, to explore the risk-return characteristics of pairs trading strategies that arise from cointegrated assets. Empirical results demonstrate that cointegration-based pairs strategies generate positive abnormal profits, and deliver superior Sharpe ratios relative to the correlation-based pair strategies proposed by Gatev, Goetzmann and Rouwenhorst (Review of Financial Studies 2006) and to the used benchmark market portfolios. This outperformance is enhanced in the recent financial crisis, and also robust out of sample and after accounting for transaction costs.

JEL classification: C58, G11, G12, G14

Keywords: pairs trading, stock price cointegration, relative liquidity

*Business Department, Universidad Carlos III de Madrid. E-mail: ifgarrig@emp.uc3m.es
†Treasury, Bankia Bank, Madrid. E-mail: iparaskevopoulos@bankia.com.
‡Business Department, Universidad Carlos III de Madrid. E-mail: ttang@emp.uc3m.es
1 Introduction

Short term price discrepancies are common across assets that are imperfectly integrated. Over the past twenty years, many hedge funds attempted to earn profits from the relative mispricing of closely related assets. This paper shows how pairs trading can be used to exploit temporary deviations from long term equilibrium relationships. Pairs trading belongs to the family of convergence trading strategies and 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 has opened a position he shorts the out-performer and longs the under-performer, until the mispricing is eliminated with converge to long run equilibrium level. We extend the Figuerola-Ferretti and Gonzalo [2010] (FG thereafter) demand and supply equilibrium framework to model bivariate equilibrium price dynamics for imperfectly integrated equity markets. In this context, two stocks that share common fundamentals are linked through pairs trading strategies that outperform benchmark trading strategies.

This paper is related to the literature of relative asset pricing, which suggests that two assets with the same payoff should be priced identically in any market state. Relative asset pricing is concerned with the price differential of two related assets. Finance academics and practitioners have paid vast interest in this type of market anomaly. For example, Froot and Dabola [1999] investigate “Siamese twin” companies whose stocks are traded in different location, while Lamont and Thaler [2003] study mispricing in tech stock carve-outs. Meulbroek [1992] and Grinblatt and Jegadeesh [1996] attempt to explain the mispricing of Eurodollar futures relative to forward contracts. While these papers tried to explain the phenomenon of market inefficiency in relative asset pricing, they do not propose any method to address the issue of price discrepancies and to gain profits out of these deviations. Given market inefficiency exists, one could construct an arbitrage portfolio by longing the “cheap” asset and shorting the “expensive” one, making profits from expected price convergence. In this regard, pairs trading as relative-value arbitrage exploits the violation of the law of one price. To this end, we model the evolution of mispricing and accordingly search for positive profit potentials.

Further, little is known about trading strategies, to the best of our knowledge, that arbitrageurs would use to optimally exploit relative mispricings. The most well-known paper is Gatev et al. [2006] (GGR thereafter), which examines the performance of pairs trading using daily U.S. stock return data from 1962 to December 2002. They use a pairs selection criteria based on the minimum distance method, that is, minimizes the sum of squared deviations between two normalized price series. Using this correlation-based algorithm, they find economically and statistically significant excess returns of around 11% per annum, and the Sharpe ratios four to six times larger than that of market premiums. Following GGR, Andrade et al. [2005] and Engelberg et al. [2009] also construct their pairs trading strategies adopting this correlation-based method. However, the results reported in this paper show that correlation can be an unreliable measure as it does not always guarantee future convergence of temporary devi-
ations. A frequently-cited case of this phenomenon is the near-collapse of Long Term Capital Management in 1998 (see for instance Edwards [1999], MacKenzie [2003], and Kondor [2009]).

In order to exploit relative mispricing, we put the focus on the concept of cointegration pointed out by Granger [1981] and Engle and Granger [1987]. Cointegration refers to a linear combination of non-stationary variables integrated of the same order could be stationary I(0), the most common case in which each series contains one unit root I(1). Cointegration in this context guarantees equilibrium price convergence that is represented in terms of a Vector Error Correction Model (VECM). The error correction term leads to mean reversion holding mispricing stationary, and guarantees that pricing error has a bounded deviation (stable variance) at all times relative to either of individual cointegrated price series with unexpected divergence risks (infinite variance). In this paper, we allow individual assets prices following a random walk, but pairwise assets are cointegrated. As such, cointegration is applied as an alternative measuring the comovement of two related assets to the traditional correlation method.

We use a theoretical model that depends on the measurement of a common non-stationary factor of two cointegrated equity markets. Under the existence of cointegration, mispricings are only short-lived in nature and therefore provide arbitrageurs opportunities to exploit future price corrections towards equilibrium via simultaneous long and short positions in cointegrated equity markets. We show how relative mispricing is contingent on the relative liquidity traded in the two related equity markets. Our model evolves around the speed by which arbitrageurs restore market disequilibria allowing measurement of price discovery and arbitrage profit determination. In this framework a market can be regarded as the determinant factor if it concentrates a higher number of participants, or higher relative liquidity. We therefore present an underlying lead-lag relation that is connected with the construction of pair strategies. Pairs construction requires the price leader to replicate the follower with a hedge ratio that is determined by the OLS coefficient. This differs from the prior literature (Shleifer and Vishny [1997], Liu and Longstaff [2004], Gatev et al. [2006] and Jurek and Yang [2007] etc.), which regard the standard (1, -1) strategy to be delta neutral.

Our empirical application uses all the EURO STOXX 50 traded equities to identify cointegration relationship with a sample ranging from 2000 to 2009, imposing the restriction that two cointegrated firms should belong to the same industry. This simplifies the search for common factor exposures driving stock prices of a given cointegrated pair. We analyze the performance of pairs strategy at the individual and portfolio level and compare its performance with the benchmark market portfolios and GGR’s pairs portfolio. We find that both pairs strategies outperform a simple buy-and-hold of market portfolios, as evidenced by significant abnormal profits and higher Sharpe ratios. More importantly, our cointegration-based pairs strategy provides superior performance relative to GGR’s correlation-based pairs trading in terms of Sharpe ratio that reflects risk-return profile during the sample period, and outperformance is enhanced.
in times of distress, the subperiod 2008-2009. These results are robust out of sample and to the existence of transaction costs.

We contribute to the finance literature in two main aspects. First, we exploit the mechanics of the VECM to construct pairs trading strategies within demand and supply framework in order to explore the risk-return characteristics of trading simultaneously cointegrated assets. The VECM equilibrium model used in this study demonstrates that pairs trading profitability is contingent on the size of cointegration error, and the elasticity of demand for arbitrage strategies. Positive profitability can be explained in terms of VECM parameters. Second, this theoretical model based on FG for equity markets is applied to all the EURO STOXX 50 traded assets. We document positive abnormal profits, and superior Sharpe ratios compared with the three proposed benchmark trading strategies and GGR’s pairs portfolio. Our results do therefore provide evidence supporting cointegration as requirement for maximizing risk-adjusted returns pursuing pairs trading strategies. These findings also suggest that unit hedge ratios commonly applied in the literature as delta neutrality is not always optimal.

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 result of the FG model applied to two distinct but cointegrated assets. Section 3 presents our data sample and empirical results on cointegration and price discovery. In section 4 we conduct an empirical analysis of pairs trading strategies and report their performance specifying the outcome with and without transaction costs. Section 5 performs an out-of-sample analysis to test for robustness of results. Section 6 concludes.

2 Theoretical Model

The aim of this section is to delineate the dynamics of two distinct but cointegrated assets in an equilibrium demand and supply framework, and to explicitly illustrate how arbitrage takes place through pairs trading strategies under the existence of short-lived mispricings. The price corrections of two cointegrated assets in price discovery depends on the relative liquidity traded in both markets. This allows us to specify which market is more efficient in incorporating new information, which will affect the portfolio allocations for hedging purposes.

We analyze pairs within the EURO STOXX 50. This provides a Blue-chip representation of supersector leaders in the Eurozone, regarded as a proxy of the overall Eurozone stock market. The constituents of EURO STOXX 50 index are large-cap stocks that can be traded in stock exchanges with complete access to market participants. In the case of cointegration, knowledge about the joint dynamics between pairwise assets is of critical importance to arbitrageurs who, will exploit short-lived deviations from equilibrium in quest of benefits from pairs strategies. Under imperfect integration, there is a finite elasticity of demand for
arbitrage strategies \((H)\) and relative prices may differ between market centers for short intervals of time by more than transaction costs. A diverse range of market frictions may impede the speed with which such price discrepancies are eliminated, such as existing communication technologies, differential tax treatment, capital constraints, and short-sale restrictions etc.

In what follows we extend FG to describe an equilibrium framework for imperfectly integrated markets. Let \(y_t\) and \(x_t\) be the price of pairwise assets in time \(t\), respectively. In order to find the non-arbitrage equilibrium condition, the following set of standard assumptions apply 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 pairwise assets are determined by the stationary process \(z_t\). These arise as a result of market imperfections that limit the communication between markets and leads to \(H > 0\) and \(z_t \neq 0\). In the limit, when arbitrage opportunities are exploited instantaneously \((H \to \infty)\) there is an immediate price adjustment to divergences between two imperfectly linked markets and \(z_t = 0\) as market frictions are eliminated. As a consequence, potential profits from pairs strategies become zero.
5. The series of price \(y_t\) and \(x_t\) are I(1), implying that their mean and auto covariances are different for every realization of \(t\).

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

\[
y_t = \gamma_0 + \gamma_1 x_t + z_t
\]

where \(\gamma_0\) is the (constant) cash amount invested (or borrowed) to buy \(\gamma_1\) units of asset \(x\) (required to replicate prices of asset \(y\)). Therefore \(\gamma_1\) is the hedge ratio as it reflects the size of the position that has to be taken in the portfolio with asset \(x\) to replicate the prices of asset \(y\). Equation 1, implies that \(y_t\) and \(x_t\) are cointegrated suggesting (imperfect) market integration.

We present the joint dynamics in both markets allowing for independent variations as well as common movements reflecting overlapping information. The pairwise assets are closely tied together via pairs trading strategies exploiting temporary price spreads that are stationary. When the spread between \(y_t\) and \(x_t\) widens, 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

---

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.

components measure a common non-stationary factor, the prices will restore equilibrium providing positive average (and cumulative) profits.

The model developed in Appendix A.1 describes the interaction between agents that trade two cointegrated assets when there is finite elasticity of demand for pursuing pairs strategies. The bivariate dynamics between \( y_t \) and \( x_t \) are represented as:

\[
\begin{pmatrix}
\Delta y_t \\
\Delta x_t
\end{pmatrix} = \frac{H}{d} \begin{pmatrix}
-N_x \\
N_y
\end{pmatrix} \begin{pmatrix}
1 & -\gamma_1 & -\gamma_0 \\
x_{t-1} & 1
\end{pmatrix} + \begin{pmatrix}
u^y_t \\
u^x_t
\end{pmatrix} \tag{2}
\]

with

\[
d = (H + AN_y) N_x + \gamma_1 H N_y \tag{3}
\]

Where there are \( N_y \) participants in the market for asset \( y \) and \( N_x \) participants in the market for asset \( x \) and, as previously specified, the elasticity of demand for pursuing pairs strategies is noted by \( H \). We rewrite the theoretical result in (2) as:

\[
\Delta P = \begin{pmatrix}
\Delta y_t \\
\Delta x_t
\end{pmatrix} = \begin{pmatrix}
\alpha_1 \\
\alpha_2
\end{pmatrix} z_{t-1} + u_t \tag{4}
\]

where \( u_t \) is a vector white noise with i.i.d shocks.\(^3\) In order to define the VECM well and ensure “pairs strategies” can be applied, the following conditions have to be satisfied:

1. If \( \alpha_1 \) and \( \alpha_2 \) are both statistically significant, they must have opposite signs, as predicted by the theoretical result in (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 with \( x_t (z_t > 0) \), 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, \( \alpha_1 \) will be negative and \( \alpha_2 \) positive, 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.

2. If \( z_t > 0 \) and the asset \( y_t \) were contributing significantly to price discovery, \( \alpha_2 \) 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 \( \alpha_1 \) would be negative and statistically significant. If both coefficients are significant then both markets contribute to price discovery. The existence of cointegration (and

---

\(^3\)Note that in the empirical part lags of \( \Delta P \) chosen in order to obtain white noise errors.
market integration) 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 cointegrating error ($\alpha_1 = \alpha_2 = 0$), the elasticity of demand for pairs strategies is infinite ($H \to \infty$), and there is no VECM, no price discovery, and no profit from pairs strategies.

3. In the VECM framework, the pairwise assets are modelled to converge to each other to restore equilibrium. The coefficients $\alpha_1$ and $\alpha_2$ 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 $x$ does not adjust to asset $y$ as it is essentially the common factor or efficient price. (The reverse is true when $\alpha_1 = 0$ and $\alpha_2 \neq 0$).

The analysis of price discovery and relative liquidity in the FG framework lies on a decomposition of cointegrated prices into a common permanent factor and a transitory component. The permanent component or common factor ($CF_t$) represents the fundamental factor and is a linear combination of $y_t$ and $x_t$ weighted by their corresponding price discovery metrics,

$$CF_t = PD_y y_t + PD_x x_t$$

It can be shown from VECM in (2) and (4), that the contribution to price discovery in the markets of asset $y$ and $x$ is:

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

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

If new information from both markets is incorporated into the common factor, $0 \leq PD_i \leq 1$ for $i = y, x$. If $PD_y = 1$ and $PD_x = 0$ then there is a predominance of asset $y$ in the price discovery process. If $PD_x = 0$ and $PD_y = 1$ there is a predominance of asset $x$ in terms of price discovery.

---

4In this case both markets are perfect substitutes and prices are “discovered” in both markets simultaneously. The model is not sustainable for this case.


6See Booth et al. [2002] and Blanco et al. [2005] for an equivalent representation of the price discovery parameters.

7Predominance in this context implies that the common factor is driven solely from the dominant price.
Equations (6) and (7) demonstrate that price discovery relies on the relative number of participants in both markets and therefore on the relative volumes traded. Changes in price discovery reflect variations in relative liquidity traded in both markets. An increase in relative liquidity of asset $x$ will be translated into higher difference in the adjustment parameters $(\alpha_1 - \alpha_2)$ and higher dominance of the market for trading asset $x$. Therefore this indicates the relative efficiency of both markets in reflecting common fundamentals.

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

$$z_t = y_t - \gamma_0 - \gamma_1 x_t$$  \hfill (8)

If $z_t > 0$, so that $y_t$ on the previous period was 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 with $x_t$ in order to profit from pairs strategies. Profits from this strategy may be defined as:

$$\Pi_t = M (-\Delta y_t + \gamma_1 \Delta x_t) = -M \Delta z_t$$  \hfill (9)

Where $\Pi_t$ are measured in €, $y_t$ and $x_t$ refer to equity price in €, and $M$ is the amount invested (in €). Delta neutrality in this framework requires the same allocation to $y_t$ and to its replicating portfolio with $x_t$. Given that portfolio replication is in price levels (and not returns) the delta or hedge ratio for a short position in asset $y$ will be $\gamma_1$ with respect to asset $x$. Portfolio allocations are therefore determined according to the cointegration regression coefficients. Substituting the result in equation (4), we get:

$$\Delta y_t = \alpha_1 z_{t-1}$$
$$\Delta x_t = \alpha_2 z_{t-1}$$
$$\Pi_t = M (-\alpha_1 + \gamma_1 \alpha_2) z_{t-1}$$  \hfill (10)

When $z_{t-1} > 0$, asset $y$ is overpriced in period $t-1$. This indicates that in time $t$, under VECM dynamics, $\alpha_1$ must be negative, $\frac{H}{d} (-N_x)$, and $\alpha_2$ positive, $\frac{H}{d} (N_y)$, as specified in our theoretical framework (2). Conversely when $z_{t-1} < 0$, asset $x$ is overpriced in $t-1$, which implies that $\alpha_2$ will be negative $\frac{H}{d} (-N_y)$ and $\alpha_1$ will be positive $\frac{H}{d} (N_x)$. Therefore we have:

$$\Pi_t = M * \frac{H}{d} (N_x + \gamma_1 N_y) z_{t-1}$$  \hfill (11)

In this framework expected theoretical profits from pairs strategies are always positive. For a given level of arbitrage opportunities ($z_t \neq 0$), the higher the elasticity of demand for arbitrage, the higher are the profits from pairs trading strategies. Profits are also increasing in the size of cointegration error given the fact that cointegration error is the source of profits in the model. In short,
under two conditions making the model work, \( H \) is finite and \( z_t \neq 0 \), a positive relationship exists between \( H \) and profits as well as \( z_t \) and profits.

Pairs trading profit dynamics can be defined in terms of the changes in the cointegrating error or stationary portfolio spread \( z_t \). These may be represented as:

\[
\Delta \Pi_t = M(\phi) \Delta z_{t-1}
\]

and,

\[
\phi = -\alpha_1 + \gamma_1 \alpha_2
\]

\[
\phi = \frac{H}{d} (N_x + \gamma_1 N_y)
\]

Equations (11) and (12) ensure that profits are stationary because the tracking error is stationary. Any drift of the replicating portfolio with asset \( x \) from its benchmark will be exploited by arbitrageurs and is therefore expected to be short-lived.

3 Cointegration and Price Discovery

In this paper we focus on the European equity market in order to identify potential profitable opportunities pursuing pairs trading strategy. All the EURO STOXX 50 traded equities are investigated. We use daily closing price data for the 50 stocks over the period dating from January 1st, 2000 to December 31st, 2009. The sample comprises 2608 trading days for most of these equities. The stock price data for five of the companies belonging to the index (Airbus Group, Anheuser-Busch, Deutsche Post, GDF Suez, and Inditex) are available for a shorter period and thus includes a smaller number of observations.\(^8\) The data source is Datastream. Our sample 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 or I(0). In what follows, we identify matching partners for each stock with the restriction that both stocks should belong to the same industry. The mechanism behind this cointegration relationship lies on the existence of an underlying common factor between them. In addition to the statistical sense, assets are linked via fundamentals. Therefore industry effect is considered to stress common factor exposures that drive prices to parity.

\(^8\)The data for Airbus Group, Anheuser-Busch, Deutsche Post, GDF Suez, and Inditex start from 2000/7/12, 2000/11/30, 2000/11/17, 2005/7/7 and 2001/5/22, respectively.
Our empirical analysis is based on the VECM specified in equation (4). Econometric details of the estimation and inference of (4) can be found in Johansen [1995] and Juselius [2006].

We start by conducting unit root tests, since unit root is a necessary condition for cointegration. Hence, Augmented Dickey-Fuller test is applied to find out which series are I(1). All I(1) series become stationary in first differences at the 1% level. Before performing the Johansen cointegration test, we determine the optimal VAR lag length required to obtain white noise error using the Schwarz information criterion, SIC.

Then Johansen cointegration test is performed to test the null hypothesis that the number of cointegration vectors is less than or equal to r against a general alternative. Empirical results are presented in Table 1. With the restriction of belonging to the same industry, we find evidence of cointegration on 12 equity pairs spreading across 7 different industries, that is, Chemicals, Insurance, Banks, Industrial Goods & Services, Utilities, Construction & Materials, and Oil & Gas. 8 out of 12 pairs are cointegrated at least at the 5% level, while the remaining 4 pairs are at the 10% level. Therefore, these pairwise equities are tied via a long-run arbitrage relationship under the imposed restriction that the error term is stationary. In addition to the statistical test, in Figure 1 we plot the normalized price paths for each pair. The pairwise equities tend to move in synchrony, which confirms that they are close substitutes to each other.

Under the existence of cointegration relationship, we then investigate the lead-lag relation for each pair to determine which asset dominates the price discovery process. Table 2 reports VECM estimates. We see that the adjustment coefficient $\alpha_1$ is significantly negative at the 1% level for all pairs, suggesting that the price of follower ($y_t$) is expected to drop by around $\alpha_1$ in response to one unit increase in the error correction term. The corresponding estimate of $\alpha_2$ is not significantly different from zero for 10 out of 12 pairs. As a result, an asymmetric lead-lag relation exists in most pairs, but with two exceptions: BBVA-Intesa and Saint Gobain-CRH. This implies that for these two pairs, both assets contribute to price discovery. But for the remaining pairs, there is a dominant asset ($x_t$) relative to its partner ($y_t$) in terms of price discovery, and thus the follower ($y_t$) does all the adjustments to restore long-term equilibrium.

Based on the reported lead-lag relation above, we use the leading asset ($x_t$) to replicate the follower ($y_t$) when constructing our pairs strategies. That is, we run an OLS regression of the follower against its leader to determine the portfolio composition, reflecting linear characteristics of cointegration relationship. This application makes our portfolio construction simple and uses an optimal hedge ratio obtained from static variance minimizing conditions (For optimality of $\gamma_1$ in Section 4.1). The estimated coefficients $\gamma_0$ and $\gamma_1$ are reported in Table 3. The (constant) cash amount, $\gamma_0$, required to replicate the follower is positive for all pairs, with the sole exception of the pair Saint Gobain and CRH. The positive sign of $\gamma_0$ suggests that long cash positions should be held to replicate the follower with $\gamma_1$ units of the leading asset. By contrast, the negative $\gamma_0$ on the pair Saint Gobain and CRH require investors to borrow $\gamma_0$ units of cash. Because most replicating portfolios require long positions in the risk-free assets interest
expenses are omitted from construction of arbitrage profits. Then we look at the estimated slope coefficient, which refers to $\gamma_1$ units of the leading asset taken in the replicating portfolio. This slope coefficient reflects the sensitivity of one asset to its matching partner, in essence the hedge ratio in our pairs trading strategy.

4 Profitability of Pairs Trading

4.1 Cointegration-based pairs trading strategies

In this section, we illustrate our pairs trading strategy based on cointegration relationship. The trading mechanism is described as follows. An arbitrageur opens a long-short position when price differential is more than one standard deviation calculated from the historical spreads, and unwind the initial position when price reversion eventually occurs. This trading mechanism, 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 (9) and (10). Given that pairs selection is given by cointegration, profits generated by our strategy are also expected to be stationary as outlined in Equations (11) and (12). Therefore, 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 2 illustrates how to perform the strategy using the cointegrated pair, Air Liquide and BASF. The fluctuating line in blue represents price spreads $z_t$, while the two straight lines in green indicate the border of one standard deviation (either positive or negative). The line in yellow, near to the x-axis, suggests the opening and closing of pairs strategy on a daily basis. We see that a position is initiated when price spread moves beyond the border and then closed when price deviation lies between the border lines. Notice that this pair opens 43 times during the sample period, and the longest trading interval lasts for 134 days.

We require two consecutive signals of "open/close" as trading criterion, meaning that the trade signal at time $t$ has to be confirmed in the following trading day $t+1$. More specifically, we will initiate a position whenever the absolute value of mispricing diverges, in two consecutive days, by more than one standard deviation. The other way around: two consecutive "close" signals are needed to liquidate the initial position. The objective of this criterion is to make perceived arbitrage opportunities credible enough and thus avoid unnecessary transaction costs arising from frequent trading. In addition, the data are closing prices on a daily basis, such that arbitrageurs will establish a position at time $t+1$ if the price spread is greater than one standard deviation at time $t$. But the opening price in general is not the same as the closing price, and thereby arbitrageurs may wait for one more signal until the end of time $t+1$ to place a bet. Note that strategy profit at time $t+1$ is contingent on the pricing error at
time \( t \) as outlined in Equation (10). In this sense, our double-signal requirement appears to reflect that “observing sufficient mispricing today, opening a trade and earning profits tomorrow”. After a pair has completed a round-trip trade, it will be subject to the identical trading rule again.

The hedge ratio in our pairs trading strategy is empirically determined as the slope coefficient by running an OLS regression between pairwise equities, as in Schaefer and Strebulaev [2008] and Kapadia and Pu [2012]. Schaefer and Strebulaev [2008] employ the empirical hedge ratio from linear regression as the benchmark, to justify the ability of structural models to predict the hedge ratio of bond returns against stock returns. Our works is more closely related to Kapadia and Pu [2012]. They exploit price discrepancies across firm’s equity and credit markets using convergence trade, in which hedge ratio is determined by the OLS regression. It is therefore commonly accepted in the statistical arbitrage literature that the linear OLS regression is applied to determine hedge ratio. In our analysis, the existence of cointegration allows simple least squares regression to reliably determine the composition of the replicating portfolio. Given that our cointegrated price series require a long-term arbitrage relationship, we assume that dynamic hedging is not required. As we know, simultaneous long-short positions can be used to hedge similar types of risks between close substitutes. Therefore, we can further show using the minimum variance (MV) method, that the OLS estimation provides an optimal static hedge ratio. In what follows, we use the MV hedge ratio to minimize the variance of portfolio value.

The hedge portfolio pursuing pairs trading can be described as longing the underpriced asset and selling the overpriced one:

\[
y_t = \gamma_0 - \gamma_1 x_t
\]

The hedger wants to minimize the variance of portfolio value so as to find the value of \( \gamma_1 \). To achieve this, we have:

\[
\min_{\gamma_1} \text{Variance} = \sigma_y^2 + \gamma_1^2 \sigma_x^2 - 2\gamma_1 \sigma_{xy}
\]

\[
\text{F.O.C.} \to 2\gamma_1 \sigma_x^2 - 2\sigma_{xy} = 0
\]

\[
\gamma_1 = \frac{\sigma_{xy}}{\sigma_x^2}
\]

Therefore the value of \( \gamma_1 \) that minimizes the variance is \( \frac{\sigma_{xy}}{\sigma_x^2} \).

The hedge ratio derived from OLS estimates is therefore identical to the one from the MV method. We contend that the optimal hedge ratio in a static framework is the OLS slope coefficient, in that cointegration ensures that the price departure between the replicating portfolio and its close substitute follows a mean-reverting process with minimum variance.

### 4.2 Statistical summary of price spread and daily excess return to equity pairs

As was discussed in section 4.1, the price spread serves as the threshold of triggering pairs trading. Given its importance, we first report the mean and standard deviation of price spreads for each pair in Table 4. Because the 2008
financial crisis is a systematic event greatly affecting the whole market, we split
the sample period into two subperiods: the tranquil period over January 2000
In general, the average pricing error is smaller in the tranquil market relative to
the mispricing in the crisis period. But we do not find evident effect of financial
crisis on the standard deviation of price spreads. This valuable evidence seems
to suggest that a non-unit hedge ratio is appropriate to be used and therefore a
linear model is valid for constructing across-asset trading. Figure 3 shows the
time series plot of the pricing error for each pair. We see that price discrepancies
can be substantial, and mispricings tend to decrease once the distance from
equilibrium is maximized at some point. This evidence supports mean reversion
under cointegration, with slow adjustment speed towards equilibrium.

Then in Table 5 we present statistical summary of daily excess return for
each equity pair. Given higher mean pricing error between 2008 and 2009, high
frequency of price deviations is expected to provide abundant opportunities to
execute pairs strategies. One piece of evidence supporting this argument is that,
for 9 out of 12 selected pairs, a larger annualized mean number of trades is doc-
dumented over the 2008-2009 period, compared to the normal market conditions
2000-2007. Our results show that pairs trading does yield statistically signifi-
cant positive abnormal profits. Importantly, 7 out of 12 equity pairs generate
higher mean daily excess returns in the crisis period. The remaining 5 pairs
that outperform in the non-crisis period seems to show huge pricing errors be-
tween 2000-2002, which are greater than or close to price spreads documented
during the 2008 crisis. The association between pricing error and profitability
deserves our enough attention. Besides, we find that the returns to pairs trading
are, in most cases, positively skewed within both subperiods. This implies that
positively skewed distribution of excess returns is common, which would lead
to downward bias of our excess returns and thus Sharpe ratios. 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]
which reports that arbitrageurs’ total return is negatively skewed as total losses
are large with small probability due to a large price divergence. Moreover, the
longest trading interval is also reported for each pair in a context where longer
trading intervals are expected to be associated with higher profits and lower
speed of equilibrium convergence.

In addition to summarize statistical characteristics pair by pair, an equal-
weighted portfolio with selected equity pairs is constructed to describe our stra-
tegy performance at the portfolio level (see details in Section 4.3). In line with
our findings at the individual level, the return to pairs portfolio is positively
skewed. In Panel A of Table 5, we see that our pairs portfolio earns an excess
return of 0.273 in the pre-crisis period. But the more attractive performance is
documented with higher magnitude of excess return during the crisis, which is
0.304 on a daily basis. Also, lower standard deviation is achieved between 2008
and 2009. This is of great importance in mirroring the ability of pair strategies
to capture arbitrage opportunities and thus generate stable profits.

In short, the mean pricing errors are substantially greater in a time of market
turmoil. These discrepancies occur frequently and take on some short-term patterns, so that mispricings are eliminated along with the passage of time. Most importantly, this preliminary analysis suggests that pairs trading may be more profitable on hard times, which we confirm in the following subsection.

4.3 Performance comparison: cointegration-based vs. correlation-based pairs strategies vs. market portfolios

To further understand the performance of cointegration-based pairs strategies, we construct an equally weighted portfolio with the 12 cointegrated pairs, which is called "Portfolio A". Notice that the restriction of pairwise equities belonging to identical industry is imposed. This pairs portfolio is formed in order to compare our strategy performance with the benchmark market portfolios. This analysis helps to uncover whether our simple and active investment strategy holds relative strength to a buy-and-hold strategy of market portfolios as passive investment. More importantly, we relate our trading strategy to the correlation-based pairs trading proposed by GGR (2006). We follow GGR to identify pairs from all the EURO STOXX 50 traded equities, by minimizing the sum of squared spreads between two normalized price series. To make these two portfolios comparable, the top 12 pairs ranked by correlation are selected to form another equal-weighted portfolio, called "Portfolio B." We do not impose the industry restriction when applying GGR’s pairs strategy. The reason is that GGR selecting pairs by the degree of correlation implies that a higher correlated pair may perform better. This underlying assumption is imposed when they form pairs portfolios across and within industries. As such, Portfolio B consisting of the 12 best pairs may provide the largest benefits based on their trading philosophy.

Panel A in Table 6 reports Sharpe ratios and excess returns to Portfolio A and B over the in-sample period, before transaction costs. Columns 1-4 present daily and annual average profits and volatilities, respectively. On this basis, we provide the simplified Sharpe ratio in column 5, as a proxy of risk-adjusted return. To unfold the economic significance from arbitrageurs’ perspective, we present annualized cumulative profits in the last column. Reported results show that both Portfolio A and B produce higher Sharpe ratios and cumulative profits relative to those earned by investing in the benchmark market portfolios such as DAX, EURO STOXX 50 and S&P 500. This finding reflects that pairs trading strategies, as a representative of active investment, are likely to beat the market. To see under which market conditions pairs strategies perform better, we plot in Figure 4 the cumulative excess returns of our pairs portfolio against the movement of EURO STOXX 50. When EURO STOXX 50 experienced a downward trend over the periods 2000-2002 and 2008-2009, pairs portfolio profits achieved a dramatic rise as evidenced by a steeper slope. By contrast,
our pairs portfolio had a interval of moderate performance whereas the market
performed particularly well. In what follows, we will further show that pairs
trading is more likely to deliver superior performance in bad times, which is
in line with GGR (2006). We additionally display the correlation between our
portfolio profits and market indices’ daily returns in Table 7. The results tell
us that our pairs portfolio is almost neutral (slightly positively correlated with)
to market indices.

Table 6 exhibits that our pairs portfolio (Portfolio A) generates annualized
cumulative profits of 72.70 Euros during the in-sample period 2000-2009, which
implies that each invested Euro will earn around 6.06 Euros per year. In con-
trast, Portfolio B yields higher annualized profits of 92.69 Euros in the same
period. We see that both pairs portfolios deliver significantly positive excess
returns in absolute terms, even with different portfolio components. However,
profit generation is inevitably accompanied by risk-taking. We therefore con-
centrate on the risk-adjusted indicator to fairly reflect both sides of the coin:
risk-return profile. Standard deviation is a commonly accepted indicator of risk,
such that we use the Sharpe Ratio to evaluate pairs strategies’ performance. Our
Portfolio A provides the Sharpe ratio of 1.63, relative to 1.45 from Portfolio B.
This indicates that Portfolio A produces excess return per unit of deviation 0.18
Euros higher than that from Portfolio B. In other words, abnormal profits from
pursuing Portfolio B is not high enough to compensate arbitrageurs for greater
risk taking.

The next step is to introduce transaction cost to assess its impact on arbit-
grage profits. For the reason that prices used to compute abnormal returns are
closing prices, with the identical probability of being at bid or ask, corrections
of these profits are needed to reflect that in practice we long at the ask and sell
at the bid prices. To this end, we collect bid and ask prices for each equity pair.
Suppose the closing prices are equally likely to be at bid or ask, then transaction
costs will subtract abnormal profits of each pair by one-half of the sum of the
bid-ask spreads of both assets, whenever we open or close a position. Panel B
in Table 6 demonstrates Sharpe ratios and cumulative profits net of transaction
costs. For our Portfolio A, reported Sharpe ratio and cumulative profits drops
by around 10.4% and 10%, respectively. However, they remain positive, and
our portfolio outperforms the Sharpe ratio of Portfolio B as was the case be-
fore transaction cost. Reported results therefore show that our pairs portfolio
survives transaction costs and provides better risk-adjusted returns than the
benchmark portfolios and GGR’s correlation-based pairs strategy.

In order to compare pairs portfolios’ performance in two distinct market
states, Table 8 presents Sharpe ratios of both Portfolio A and B over 2000-2007
and 2008-2009 respectively. Consistent with results at the individual firm level,
both portfolios produce much higher Sharpe ratios during the crisis period. A
possible interpretation is that market imperfections were higher, reflected by
higher $z_i$ in our model as arbitrage opportunities and also the source of returns.
We also find interesting evidence that, the difference of the magnitude in the
Sharpe ratios between Portfolio A and B is larger over the 2008-2009 period.
This finding implies that, although Portfolio A always outperforms Portfolio B
the proposed cointegration-based strategy benefits from arbitrage opportunities to higher extent during turbulent times. It is also worth noting that, higher Sharpe ratio of Portfolio A during the crisis arises as a consequence of higher average return and lower volatility of arbitrage profit. Cointegration-based pairs strategies can therefore be used for hedging away market shocks and simultaneously yield considerable profits under volatile market conditions.

We further examine the percentage of winning bets in both long and short positions. Reported results in Table 9 reveal that both portfolios have over 50% of winning trades from their long as well as short positions. Given that daily net returns are equal to the sum of profit/loss from both positions, we extend our analysis examining the percentage of negative returns and annualized total losses over the whole period. Reported results are favorable to Portfolio A, which provides lower percentage of losing trades and smaller amount of total losses than Portfolio B. The proposed trading framework is therefore more likely to give correct signals of price movements in the following day.

All in all, pairs strategies yield significant abnormal profits. This suggests price convergence after the onset of a pricing anomaly, as evidence to claim that pricing error between pairwise assets are anomalous. Given the high frequency of arbitrage opportunities in turbulent markets, pairs trading performs well under adverse market conditions. By comparison, we find that our cointegration-based strategy outperforms GGR’s strategy as well as market portfolios in terms of Sharpe ratio, and outperformance is enhanced on hard times.

4.4 Why cointegration-based strategy performs better?

In this subsection, we describe three potential drivers that explain the reported outperformance of the cointegration-based strategy. We believe that one of the key explanations is that our portfolio is relatively more diversified, supported by the fact that Portfolio A contains 20 distinct equities across 7 industries. The larger the number of assets from diverse industries, the greater ability a portfolio holds to generate diversification effect protecting investors from market shocks. Although Portfolio B also covers 7 industries, it only consists of 12 distinct equities. It means that a given asset is more likely to be used repeatedly when forming pairs. Portfolio B is hence concentrated on a limited number of assets, leading to less diversification and thus more likely to be exposed to market risks.

Another possible reason is associated with the difference between cointegration and correlation. High correlation does not necessarily indicate high cointegration. Similarly, cointegrated series can move in opposite directions showing temporal low correlation. Correlation is in fact a short-term measure, unable to reflect the long-run behavior between price series. Since correlation does not guarantee the stationarity of price spreads, correlation-based strategies can lead to non-stationary pricing errors and non-temporary deviations from equilibrium. Correlation is therefore not sufficient to model price comovements in the long run. In contrast, cointegration ensures mean reversion of price spreads because two non-stationary price series share a common stochastic trend. Under cointegration, deviation occurs only in the short term but disappears with time.
to reflect the long-term equilibrium relationship. This is why price discrepancies between cointegrated assets are temporary and can also be well exploited through pairs trading strategies.

Last but not least, the hedge ratio applied to construct long-short positions may also contribute to the improved performance. Hedge ratios in this analysis are determined empirically by an OLS regression exploiting the consistency property between OLS and maximum-likelihood estimates under cointegration. As was discussed above, the minimum variance (MV) method confirms that our portfolio composition minimizes the variance of portfolio value.

Furthermore, the standard (1, -1) arbitrage strategy is applied to portfolio holdings for pairs trading purpose in GGR. This seems reasonable and achieve a delta neutral position by using unit hedge ratio. However it leads to greater volatility of portfolio profits compared to our portfolio based upon linear relations. As a consequence, greater volatility exerts a negative effect on the risk-adjusted return to GGR’s pairs portfolio, represented by the Sharpe ratio.

4.5 Is one standard deviation the optimal threshold?

In this subsection, we investigate whether one standard deviation is in general the optimal threshold triggering pairs trading. Table 10 exhibits performance results from investments in our pairs portfolio using 1, 1.65, 2 and 2.58 standard deviations as threshold, respectively. In addition to 1 standard deviation, the remaining three scenarios correspond to statistical significance at the 10%, 5% and 1% level, respectively. Reported figures show that, cumulative profits and the Sharpe ratio are maximized when one historical standard deviation is applied to initial pairs trading. By contrast, using higher standard deviation as threshold might result in lower frequency of trading. As expected, more trading opportunities arise with the 1 standard deviation threshold which leads to better performance of pairs trading strategies.

In addition, Table 10 also exhibits the source of profits, from either the long or short positions in the round-trip trades. We find that, in all scenarios, over 50% of excess returns are from short positions. This finding is in line with the empirical evidence documented in GGR (2006). We believe that the asymmetric source of profit generation, may be associated with limits of arbitrage. For instance, the presence of restrictions to arbitrage reflected in $H$ and $z_t$, such as limits in short positions, may explain the magnitude and persistence of excess returns.

5 Robustness check: Out of sample performance analysis

In this section, we conduct an out of sample analysis over the 2010-2013 period. This is a quite important test of performance reliability to state that the relative strength of our strategy is not sample dependent. Performance results
are reported in Table 11, before and after transaction costs in Panel A and B respectively. Both cointegration-based and correlation-based pairs portfolios produce positive cumulative returns and Sharpe ratios that outperform those earned from investing in the benchmark market portfolios. By subtracting the round-trip trading costs to get portfolio profits net of transaction costs, reported positive figures confirm robustness of pairs trading strategy out of sample. Consistent with what we find over the 2000-2009 period, our proposed Portfolio A outperforms Sharpe ratio of Portfolio B based on GGR (2006), even after accounting for transaction costs. On the other hand, we also see that Sharpe ratios go down out of sample, as documented by Chen et al. [2012]. We think this is not a serious issue because the financial crisis did contribute a lot to driving up the profitability of pairs trading strategies. When we compare the out-of-sample performance with the pre-crisis one, Sharpe ratios just decrease slightly and hence indirectly confirms the evidence of better performance in difficult times. Moreover, results in Table 12 reflect that although both pairs strategies exhibit good ability to signal price dynamics, our proposed trading rule suffers less losses in absolute as well as percentage terms. As shown in Table 13, we find again that the Sharpe ratio and cumulative profits are maximized when using one standard deviation as the threshold to place a position. It is therefore optimal to use a relatively small standard deviation to trigger trading. Higher standard deviation leads to lower frequency of trading and lower Sharpe Ratios.

The above analysis indicates that our pairs strategy remains profitable when employing the in-sample cointegration relationship to identify arbitrage opportunities over the extended sample period. More importantly, our cointegration-based portfolio yields better Sharpe ratio relative to that obtained from correlation-based pairs strategy and market portfolios. We therefore conclude that our strategy’s performance is not an artifact of the sample period.

6 Conclusion

In this paper we work on a linear price discovery equilibrium model illustrating that two cointegrated assets are closely linked through price spread under the existence of pairs trading strategy. Price discovery process is of critical importance and we relate this process to the relative level of liquidity in both markets. In particular, this simple equilibrium framework shows how smooth adjustment of imperfectly integrated markets can be exploited to obtain arbitrage profits. The VECM model is extended to demonstrate how long and short positions in cointegrated assets lead to profits that are stationary and dependent on the size of cointegration error and the elasticity of demand for pairs strategies.

Our study of pairs trading strategies in a non-arbitrage framework documents significant abnormal profits, suggesting that price discrepancy is short-lived mispricing instead of permanent deviation. Empirical analysis also shows that cointegration-based pairs portfolio delivers superior Sharpe ratio relative to pairs portfolio with highly correlated assets as well as the benchmark market.
portfolios, even after transaction costs are accounted for. This finding is particularly evident during the crisis period 2008-2009, indicating that turbulent market conditions may favor pursuing pairs strategy with cointegrated assets. In addition, our proposed pairs strategy remains profitable over the extended sample period, which implies the robustness of strategy performance that is not an artifact of the sample period.
References


Part I
Appendix A.1 A theoretical model for the dynamics of imperfectly integrated markets

The following model presents bivariate equilibrium price dynamics for imperfectly integrated markets with inter-market restrictions for cross communication and trading. Let $y_t$ and $x_t$ be the price of the pairwise assets in time $t$, respectively. We assume that there are $N_y$ participants in the market for asset $y$ and $N_x$ participants in the market for asset $x$. These participants will take positions in asset $y$ or/and $x$. They will also pursue arbitrage strategies exploiting temporary mispricings from the existing long-term relationship in the two markets:

$$y_t = \gamma_0 + \gamma_1 x_t + z_t$$  \hspace{1cm} (13)

Let $P_{i,t}$ be the net position of the $i^{th}$ participant immediately prior to period $t$ and $B_{i,t}$ the bid price at which that participant is willing to hold the position $P_{i,t}$. Then the demand schedule of the $i^{th}$ participant in the market for asset $y$ in period $t$ is

$$P_{i,t} - A(y_t - B_{i,t})$$  \hspace{1cm} (14)

with $i = 1, ..., N_y$ where $A > 0$, is the elasticity of demand, assumed to be the same for all participants. Note that due to the dynamic structure to be imposed to the bid price, $B_{i,t}$.

The demand schedule for the $j^{th}$ participant in the market for asset $x$ is

$$P_{j,t} - A(x_t - B_{j,t}) , A > 0, j = 1, ..., N_x$$  \hspace{1cm} (15)

The aggregate market demand schedule for agents that perform pairs trading strategies in $y_t$ and $x_t$ exploiting period $t$ temporary mispricings:

$$H((\gamma_1 x_t + \gamma_0) - y_t) , H > 0$$

$$H(z_t) , H > 0$$  \hspace{1cm} (16)

This implies that $y_t$ can be replicated by a portfolio with $x_t$. $z_t$ represents
the stationary arbitrage opportunities in two imperfectly integrated markets, and $H$ is the elasticity of market demand for pair strategies. When transaction costs are negligible, inter-market communications improve and markets become in the limit perfectly integrated. However transaction costs can be significant for providing arbitrage services. Capital constrain and short-sale restriction also impede inter-market communications. The elasticity of demand for pairs strategies is therefore finite because buying $y_t$ and selling $x_t$ or vice versa are not risk-less strategies.

The market for asset $y$ will clear at the value of $y_t$ that solves,

$$
\sum_{i=1}^{N_y} P_{i,t} = \sum_{i=1}^{N_y} (P_{i,t} - A (y_t - B_{i,t})) + H ((\gamma_1 x_t + \gamma_0) - y_t) \tag{17}
$$

with $H > 0$.

The market for asset $x$ will clear at the value of $x_t$ such that

$$
\sum_{j=1}^{N_x} P_{j,t} = \sum_{j=1}^{N_x} (P_{j,t} - A (x_t - B_{j,t})) + H ((\gamma_1 x_t + \gamma_0) - y_t) \tag{18}
$$

Solving equations (17) and (18) for $y_t$ and $x_t$ as a function of the mean bid price set by market participants in $y_t \left( B^y_t = N_y^{-1} \sum_{i=1}^{N_y} B_{i,t}\right)$ and the mean bid price $\left( B^x_t = N_x^{-1} \sum_{j=1}^{N_x} B_{j,t}\right)$ for market participants in $x_t$, we obtain

$$
y_t = \frac{(AN_y + H \gamma_1) N_y B^y_t + H N_y \gamma_1 B^x_t + H N_y \gamma_0}{(H + AN_y) N_y + H N_y \gamma_1 + H N_y \gamma_0}
$$

$$
x_t = \frac{H N_x B^x_t + (AN_x + H) N_x B^x_t - H N_x \gamma_0}{(H + AN_y) N_x + H N_y \gamma_1} \tag{19}
$$

To derive the dynamic price relationships, the model in equation (19) must be characterized with a description of the evolution of bid prices. It is assumed that immediately after the market clearing period $t - 1$ the $i^{th}$ participant in $y_t$ was willing to hold a position $P_{i,t}$ at a price $y_{t-1}$. Following FG, 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

$$
B_{i,t} = y_{t-1} + e_t + w_{i,t}
$$

$$
B_{j,t} = x_{t-1} + e_t + w_{j,t} \tag{20}
$$

$$
cov(e_t, w_{i,t}) = 0, \forall i
$$

$$
cov(w_{i,t}, w_{j,t}) = 0, \forall i \neq j
$$

23
with \( i = 1, \ldots, N_y \) and \( j = 1, \ldots, N_x \). Where the vector \( (e_1, w_{i,t}, w_{j,t}) \) 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 the position \( P_{i,t} \) in the market \( y_t \). This price change has a component common to all participants \( (e_t) \) and a component idiosyncratic to the \( i^{th} \) participant \( (w_{i,t}) \).

The equations in (20) imply that the mean bid price in each market in period \( t \) will be

\[
B^y_t = y_{t-1} + e_t + w^y_{i,t} \\
B^x_t = x_{t-1} + e_t + w^x_{i,t}
\]

(21)

where, \( w^y_{i,t} = \sum_{i=1}^{N_y} w^y_{i,t} \) and \( w^x_{i,t} = \sum_{j=1}^{N_x} w^x_{j,t} \). Substituting expressions (21) into (19) yields the following vector model

\[
\begin{pmatrix}
  y_t \\
  x_t
\end{pmatrix} = \frac{H\gamma_0}{d} \begin{pmatrix}
  N_x \\
  -N_y
\end{pmatrix} + M \begin{pmatrix}
  y_{t-1} \\
  x_{t-1}
\end{pmatrix} + \begin{pmatrix}
  u^y_t \\
  u^x_t
\end{pmatrix}
\]

(22)

where

\[
\begin{pmatrix}
  u^y_t \\
  u^x_t
\end{pmatrix} = M \begin{pmatrix}
  e_t + w^y_{i,t} \\
  e_t + w^x_{i,t}
\end{pmatrix}
\]

(23)

\[
M = \frac{1}{d} \begin{bmatrix}
  N_y (\gamma_1 H + AN_x) & \gamma_1 H N_x \\
  H N_y & (H + AN_y) N_x
\end{bmatrix}
\]

(24)

And

\[
d = (H + AN_y) N_x + \gamma_1 H N_y
\]

(25)

We now convert (22) into a Vector Error Correction Model (VECM) by subtracting \((y_{t-1}, x_{t-1})\)' from both sides, with

\[
\begin{pmatrix}
  \Delta y_t \\
  \Delta x_t
\end{pmatrix} = \frac{H\gamma_0}{d} \begin{pmatrix}
  N_x \\
  -N_y
\end{pmatrix} + (M - I) \begin{pmatrix}
  y_{t-1} \\
  x_{t-1}
\end{pmatrix} + \begin{pmatrix}
  u^y_t \\
  u^x_t
\end{pmatrix}
\]

(26)

\[
M - I = \frac{1}{d} \begin{bmatrix}
  -H N_x & \gamma_1 H N_x \\
  H N_y & -H N_y \gamma_1
\end{bmatrix}
\]

(27)
Rearranging terms,

\[
\begin{pmatrix}
\Delta y_t \\
\Delta x_t
\end{pmatrix}
= \frac{H}{d} \begin{pmatrix}
-N_x \\
N_y
\end{pmatrix}
\begin{pmatrix}
1 & -\gamma_1 & -\gamma_0
\end{pmatrix}
\begin{pmatrix}
y_{t-1} \\
x_{t-1}
\end{pmatrix}
+ \begin{pmatrix}
u^y_t \\
u^x_t
\end{pmatrix}
\]

(28)
Part II
Appendix A.2 Empirical Results

Table 1: Johansen cointegration test on equity pairs
(January 2000-December 2009)

<table>
<thead>
<tr>
<th>Industry</th>
<th>Pairs: $y_t \sim x_t$</th>
<th>Number of Cointegration Vectors</th>
<th>Optimal lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemicals</td>
<td>Air Liquide-BASF</td>
<td>19.643*</td>
<td>2</td>
</tr>
<tr>
<td>Insurance</td>
<td>Assicurazioni-AXA</td>
<td>30.554***</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Muenchener-Allianz</td>
<td>18.110*</td>
<td>1</td>
</tr>
<tr>
<td>Banks</td>
<td>BBVA-Deutsche Bank</td>
<td>21.786**</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>BBVA-Intesa</td>
<td>25.646***</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Deutsche Bank-Unicredit</td>
<td>21.454**</td>
<td>1</td>
</tr>
<tr>
<td>Industrial Goods &amp; Services</td>
<td>Philips-Airbus</td>
<td>22.537**</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Philips-Deutsche Post</td>
<td>18.123*</td>
<td>1</td>
</tr>
<tr>
<td>Utilities</td>
<td>GDF Suez-E.ON</td>
<td>18.430*</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E.ON-Iberdrola</td>
<td>30.591***</td>
<td>1</td>
</tr>
<tr>
<td>Construction &amp; Materials</td>
<td>Saint Gobain-CRH</td>
<td>31.577***</td>
<td>1</td>
</tr>
<tr>
<td>Oil &amp; Gas</td>
<td>Total-Repsol</td>
<td>26.595***</td>
<td>1</td>
</tr>
</tbody>
</table>

The first two columns of Table 2 present Johansen trace test statistics for the number of cointegrating relations between selected pairwise equities. A constant is included in the long-term statistical relation. The number of lags for each pair is reported in the third column, which is optimized according to the Schwarz criterion. *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.

Table 2: VECM estimates of equity pairs
(January 2000-December 2009)

<table>
<thead>
<tr>
<th>Pairs: $y_t \sim x_t$</th>
<th>$\alpha_1$</th>
<th>t-Statistics</th>
<th>$\alpha_2$</th>
<th>t-Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Liquide-BASF</td>
<td>-0.014***</td>
<td>-2.944</td>
<td>0.001</td>
<td>0.440</td>
</tr>
<tr>
<td>Assicurazioni-AXA</td>
<td>-0.014***</td>
<td>-4.404</td>
<td>-0.002</td>
<td>-0.540</td>
</tr>
<tr>
<td>Muenchener-Allianz</td>
<td>-0.010***</td>
<td>-3.356</td>
<td>-0.004</td>
<td>-1.325</td>
</tr>
<tr>
<td>BBVA-Deutsche Bank</td>
<td>-0.011***</td>
<td>-2.735</td>
<td>0.023</td>
<td>0.842</td>
</tr>
<tr>
<td>BBVA-Intesa</td>
<td>-0.010***</td>
<td>-2.930</td>
<td>0.016</td>
<td>1.667*</td>
</tr>
<tr>
<td>Deutsche Bank-Unicredit</td>
<td>-0.012***</td>
<td>-3.034</td>
<td>0.001</td>
<td>0.953</td>
</tr>
<tr>
<td>Philips-Airbus</td>
<td>-0.007***</td>
<td>-3.958</td>
<td>0.001</td>
<td>0.514</td>
</tr>
<tr>
<td>Philips-Deutsche Post</td>
<td>-0.011***</td>
<td>-3.692</td>
<td>-9.10E-05</td>
<td>-0.057</td>
</tr>
<tr>
<td>GDF Suez-E.ON</td>
<td>-0.029***</td>
<td>-3.557</td>
<td>-0.010</td>
<td>-1.223</td>
</tr>
<tr>
<td>E.ON-Iberdrola</td>
<td>-0.018***</td>
<td>-4.863</td>
<td>-0.001</td>
<td>-0.763</td>
</tr>
<tr>
<td>Saint Gobain-CRH</td>
<td>-0.011***</td>
<td>-3.303</td>
<td>0.003</td>
<td>2.057**</td>
</tr>
<tr>
<td>Total-Repsol</td>
<td>-0.013***</td>
<td>-3.492</td>
<td>0.0004</td>
<td>0.311</td>
</tr>
</tbody>
</table>

This table presents estimates of the adjustment vector in the VECM specified in (4). *, ** and *** refers to statistical significance at the 10%, 5% and 1% level.
Table 3: The OLS estimates of equity pairs (January 2000-December 2009)

<table>
<thead>
<tr>
<th>Pair</th>
<th>$\gamma_0$</th>
<th>$\gamma_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Liquide-BASF</td>
<td>17.170</td>
<td>1.252</td>
</tr>
<tr>
<td>Assicurazioni-AXA</td>
<td>9.064</td>
<td>0.684</td>
</tr>
<tr>
<td>Muenchener-Allianz</td>
<td>15.015</td>
<td>0.915</td>
</tr>
<tr>
<td>BBVA-Deutsche Bank</td>
<td>2.484</td>
<td>0.137</td>
</tr>
<tr>
<td>BBVA-Intesa</td>
<td>2.762</td>
<td>2.468</td>
</tr>
<tr>
<td>Deutsche Bank-Unicredit</td>
<td>4.198</td>
<td>2.382</td>
</tr>
<tr>
<td>Philips-Airbus</td>
<td>13.094</td>
<td>0.467</td>
</tr>
<tr>
<td>Philips-Deutsche Post</td>
<td>5.111</td>
<td>0.938</td>
</tr>
<tr>
<td>GDF Suez-E.ON</td>
<td>10.879</td>
<td>0.645</td>
</tr>
<tr>
<td>E.ON-Iberdrola</td>
<td>4.171</td>
<td>4.868</td>
</tr>
<tr>
<td>Saint Gobain-CRH</td>
<td>-2.357</td>
<td>2.317</td>
</tr>
<tr>
<td>Total-Repsol</td>
<td>14.012</td>
<td>1.705</td>
</tr>
</tbody>
</table>

This table shows the estimated OLS coefficients $\gamma_0$ and $\gamma_1$ specified in (1).

Table 4: Mean and standard deviation of price spreads (January 2000-December 2009)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average price spread</td>
<td>Std. error</td>
<td>Average price spread</td>
<td>Std. error</td>
</tr>
<tr>
<td>Air Liquide-BASF</td>
<td>2.68</td>
<td>3.42</td>
<td>3.19</td>
<td>3.26</td>
</tr>
<tr>
<td>Assicurazioni-AXA</td>
<td>1.60</td>
<td>2.13</td>
<td>2.35</td>
<td>2.66</td>
</tr>
<tr>
<td>Muenchener-Allianz</td>
<td>22.84</td>
<td>26.81</td>
<td>16.34</td>
<td>15.56</td>
</tr>
<tr>
<td>BBVA-Deutsche Bank</td>
<td>0.69</td>
<td>0.87</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>BBVA-Intesa</td>
<td>0.75</td>
<td>0.96</td>
<td>1.23</td>
<td>1.11</td>
</tr>
<tr>
<td>Deutsche Bank-Unicredit</td>
<td>5.44</td>
<td>7.02</td>
<td>4.10</td>
<td>4.70</td>
</tr>
<tr>
<td>Philips-Airbus</td>
<td>5.05</td>
<td>6.94</td>
<td>3.92</td>
<td>3.44</td>
</tr>
<tr>
<td>Philips-Deutsche Post</td>
<td>3.14</td>
<td>4.22</td>
<td>1.89</td>
<td>1.03</td>
</tr>
<tr>
<td>GDF Suez-E.ON</td>
<td>1.86</td>
<td>1.96</td>
<td>2.01</td>
<td>2.13</td>
</tr>
<tr>
<td>E.ON-Iberdrola</td>
<td>1.89</td>
<td>2.35</td>
<td>2.42</td>
<td>2.98</td>
</tr>
<tr>
<td>Saint Gobain-CRH</td>
<td>2.82</td>
<td>3.62</td>
<td>5.87</td>
<td>5.46</td>
</tr>
<tr>
<td>Total-Repsol</td>
<td>3.09</td>
<td>3.95</td>
<td>2.25</td>
<td>2.28</td>
</tr>
</tbody>
</table>

This table shows the average price spread and standard deviation of the price spread. The price spread is defined as the price of asset $y$ minus its replicating portfolio with asset $x$, which is specified in (8).
Table 5: Summary statistics of daily excess return to equity pairs (January 2000-December 2009)

Panel A: January 2000-December 2007

<table>
<thead>
<tr>
<th>Pairs: $y_t \sim x_t$</th>
<th>Annualized num of trades</th>
<th>Longest trading interval</th>
<th>Mean</th>
<th>Median</th>
<th>Stdev</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Liquide-BASF</td>
<td>4.3</td>
<td>134</td>
<td>0.079</td>
<td>0.058</td>
<td>0.79</td>
<td>0.04</td>
<td>1.47</td>
<td>-3.25</td>
<td>3.16</td>
</tr>
<tr>
<td>Assicurazioni-AXA</td>
<td>2.4</td>
<td>111</td>
<td>0.034</td>
<td>0.000</td>
<td>0.50</td>
<td>0.14</td>
<td>0.78</td>
<td>-1.87</td>
<td>1.71</td>
</tr>
<tr>
<td>Muenchener-Allianz</td>
<td>3.3</td>
<td>239</td>
<td>0.266</td>
<td>0.006</td>
<td>4.06</td>
<td>0.04</td>
<td>7.10</td>
<td>-25.9</td>
<td>24.4</td>
</tr>
<tr>
<td>BBVA-Deutsche Bank</td>
<td>3.1</td>
<td>123</td>
<td>0.019</td>
<td>0.008</td>
<td>1.21</td>
<td>1.21</td>
<td>9.40</td>
<td>-0.64</td>
<td>1.54</td>
</tr>
<tr>
<td>BBVA-Intesa</td>
<td>4.0</td>
<td>137</td>
<td>0.030</td>
<td>0.012</td>
<td>0.22</td>
<td>0.88</td>
<td>5.13</td>
<td>-0.82</td>
<td>1.52</td>
</tr>
<tr>
<td>Deutsche Bank-Unicredit</td>
<td>4.3</td>
<td>179</td>
<td>0.128</td>
<td>0.104</td>
<td>1.42</td>
<td>0.37</td>
<td>11.75</td>
<td>-10.5</td>
<td>11.8</td>
</tr>
<tr>
<td>Philips-Airbus</td>
<td>1.9</td>
<td>178</td>
<td>0.069</td>
<td>0.047</td>
<td>0.99</td>
<td>0.04</td>
<td>0.77</td>
<td>-5.23</td>
<td>4.46</td>
</tr>
<tr>
<td>Philips-Deutsche Post</td>
<td>2.8</td>
<td>196</td>
<td>0.053</td>
<td>0.017</td>
<td>0.80</td>
<td>0.11</td>
<td>2.96</td>
<td>-2.98</td>
<td>4.23</td>
</tr>
<tr>
<td>GDF Suez-E.ON</td>
<td>1.4</td>
<td>87</td>
<td>0.052</td>
<td>0.002</td>
<td>0.39</td>
<td>0.47</td>
<td>1.14</td>
<td>-1.03</td>
<td>1.33</td>
</tr>
<tr>
<td>E.ON-Iberdrola</td>
<td>2.6</td>
<td>233</td>
<td>0.045</td>
<td>0.015</td>
<td>0.42</td>
<td>0.57</td>
<td>3.84</td>
<td>-1.48</td>
<td>2.71</td>
</tr>
<tr>
<td>Saint Gobain-CRH</td>
<td>3.4</td>
<td>115</td>
<td>0.115</td>
<td>0.070</td>
<td>0.93</td>
<td>0.33</td>
<td>1.58</td>
<td>-3.07</td>
<td>3.90</td>
</tr>
<tr>
<td>Total-Repsol</td>
<td>3.5</td>
<td>179</td>
<td>0.070</td>
<td>0.055</td>
<td>0.59</td>
<td>0.14</td>
<td>0.77</td>
<td>-1.86</td>
<td>2.42</td>
</tr>
<tr>
<td>Pairs Portfolio</td>
<td>0.273</td>
<td>0.034</td>
<td>2.910</td>
<td>0.145</td>
<td>26.05</td>
<td>-35.46</td>
<td>30.84</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel B: January 2008-December 2009

<table>
<thead>
<tr>
<th>Pairs: $y_t \sim x_t$</th>
<th>Annualized num of trades</th>
<th>Longest trading interval</th>
<th>Mean</th>
<th>Median</th>
<th>Stdev</th>
<th>Skew</th>
<th>Kurtosis</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air Liquide-BASF</td>
<td>4.5</td>
<td>102</td>
<td>0.068</td>
<td>0.050</td>
<td>0.914</td>
<td>0.188</td>
<td>0.746</td>
<td>-2.48</td>
<td>2.89</td>
</tr>
<tr>
<td>Assicurazioni-AXA</td>
<td>3.0</td>
<td>122</td>
<td>0.023</td>
<td>0.017</td>
<td>0.242</td>
<td>0.360</td>
<td>1.288</td>
<td>-0.69</td>
<td>0.96</td>
</tr>
<tr>
<td>Muenchener-Allianz</td>
<td>2.0</td>
<td>85</td>
<td>0.347</td>
<td>0.035</td>
<td>2.534</td>
<td>-0.032</td>
<td>0.557</td>
<td>-6.82</td>
<td>7.63</td>
</tr>
<tr>
<td>BBVA-Deutsche Bank</td>
<td>4.5</td>
<td>112</td>
<td>0.004</td>
<td>-0.001</td>
<td>0.146</td>
<td>1.939</td>
<td>0.60</td>
<td>-0.60</td>
<td>0.55</td>
</tr>
<tr>
<td>BBVA-Intesa</td>
<td>5.0</td>
<td>94</td>
<td>0.015</td>
<td>0.006</td>
<td>0.156</td>
<td>0.216</td>
<td>1.761</td>
<td>-0.55</td>
<td>0.68</td>
</tr>
<tr>
<td>Deutsche Bank-Unicredit</td>
<td>3.5</td>
<td>28</td>
<td>0.154</td>
<td>0.000</td>
<td>1.105</td>
<td>0.432</td>
<td>0.969</td>
<td>-2.83</td>
<td>3.98</td>
</tr>
<tr>
<td>Philips-Airbus</td>
<td>4.0</td>
<td>27</td>
<td>0.046</td>
<td>0.000</td>
<td>0.357</td>
<td>0.616</td>
<td>1.170</td>
<td>-0.83</td>
<td>1.31</td>
</tr>
<tr>
<td>Philips-Deutsche Post</td>
<td>1.5</td>
<td>25</td>
<td>0.281</td>
<td>0.230</td>
<td>0.508</td>
<td>0.955</td>
<td>0.136</td>
<td>-0.33</td>
<td>1.39</td>
</tr>
<tr>
<td>GDF Suez-E.ON</td>
<td>5.5</td>
<td>44</td>
<td>0.069</td>
<td>0.057</td>
<td>0.807</td>
<td>0.026</td>
<td>2.333</td>
<td>-2.90</td>
<td>3.39</td>
</tr>
<tr>
<td>E.ON-Iberdrola</td>
<td>4.5</td>
<td>79</td>
<td>0.077</td>
<td>0.064</td>
<td>0.780</td>
<td>-0.310</td>
<td>3.833</td>
<td>-3.81</td>
<td>3.19</td>
</tr>
<tr>
<td>Saint Gobain-CRH</td>
<td>4.5</td>
<td>201</td>
<td>0.115</td>
<td>0.143</td>
<td>1.149</td>
<td>-0.001</td>
<td>0.346</td>
<td>-3.34</td>
<td>3.71</td>
</tr>
<tr>
<td>Total-Repsol</td>
<td>4.5</td>
<td>29</td>
<td>0.126</td>
<td>0.100</td>
<td>0.834</td>
<td>-0.429</td>
<td>1.181</td>
<td>-2.94</td>
<td>2.08</td>
</tr>
<tr>
<td>Pairs Portfolio</td>
<td>0.304</td>
<td>0.094</td>
<td>1.750</td>
<td>0.727</td>
<td>4.335</td>
<td>-6.88</td>
<td>9.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table presents descriptive statistics for the mean excess return for each equity pair. We trade according to the rule that establish a position in a pair when price spreads diverge, in two consecutive days, by more than one historical standard deviation. Panel A and B report results over the 2000-2007 period and the crisis period of 2008-2009, respectively.
Table 6: Annual simplified Sharpe ratio and annual cumulative arbitrage profits for three market portfolios and two pairs portfolios (January 2000-December 2009)

<table>
<thead>
<tr>
<th></th>
<th>Average Profit</th>
<th>Average Volatility</th>
<th>Sharpe ratio</th>
<th>Cumulative Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily</td>
<td>Annual</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(i) DAX</td>
<td>0.0001</td>
<td>0.0226</td>
<td>0.0166</td>
<td>0.2635</td>
</tr>
<tr>
<td>(ii) EURO STOXX50</td>
<td>-0.0001</td>
<td>-0.0165</td>
<td>0.0160</td>
<td>0.2541</td>
</tr>
<tr>
<td>(iii) S&amp;P 500</td>
<td>-0.00001</td>
<td>0.0020</td>
<td>0.0140</td>
<td>0.2223</td>
</tr>
</tbody>
</table>

Panel A. Excess returns before transaction costs

|                      | Daily | Annual |              |                    |
| (iv) Portfolio A     | 0.28  | 70.27  | 2.72         | 43.13              |
| (v) Portfolio B      | 0.36  | 89.56  | 3.88         | 61.60              |

Panel B. Excess returns after transaction costs

|                      | Daily | Annual |              |                    |
| (iv) Portfolio A     | 0.25  | 63.18  | 2.72         | 43.19              |
| (v) Portfolio B      | 0.33  | 83.92  | 3.88         | 61.61              |

This table reports mean profits, volatility (measured by the standard deviation), simplified Sharpe Ratios and annualized cumulative profits gained with the five proposed strategies. (i), (ii) and (iii) are buy-and-hold strategies of market portfolios. That is, €1 investment in DAX index, €1 in EURO STOXX50 index and $1 in S&P500 index, respectively. And (iv) and (v) are equal-weighted portfolios performing pairs trading strategies. Portfolio A is constructed with 12 cointegrated pairs, whereas Portfolio B is formed with the top 12 pairs ranked by the minimum distance method. For each pair in either portfolio, €1 investment (with opposite positions) in pairwise equities is conducted. Zero interest rates are assumed. Panel A reports figures before transaction costs. Panel B reduces profits in A by one half of the sum of the bid-ask spreads.

Table 7: Correlations between pairs portfolio profits and market indices (January 2000-December 2009)

<table>
<thead>
<tr>
<th>Correlations</th>
<th>EURO STOXX 50</th>
<th>S&amp;P 500</th>
<th>Portfolio A</th>
</tr>
</thead>
<tbody>
<tr>
<td>EURO STOXX 50</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>0.547</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Portfolio A</td>
<td>0.032</td>
<td>0.014</td>
<td>1</td>
</tr>
</tbody>
</table>

This table describes the correlation between pairs strategy’s (Portfolio A) and market indices’ daily excess returns.
### Table 8: Annual simplified Sharpe ratio for two pairs portfolios: Portfolio A and B (January 2000-December 2009)

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Annual Average Volatility</th>
<th>Annual Sharpe ratio</th>
<th>Annual Average Volatility</th>
<th>Annual Sharpe ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel A. before transaction costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000-2009</td>
<td>70.27</td>
<td>43.13</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>2000-2007</td>
<td>68.71</td>
<td>46.20</td>
<td>1.49</td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>76.50</td>
<td>27.77</td>
<td>2.75</td>
<td></td>
</tr>
<tr>
<td>Panel A. after transaction costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000-2009</td>
<td>63.18</td>
<td>43.19</td>
<td>1.46</td>
<td></td>
</tr>
<tr>
<td>2000-2007</td>
<td>60.88</td>
<td>46.26</td>
<td>1.32</td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>72.36</td>
<td>27.77</td>
<td>2.61</td>
<td></td>
</tr>
<tr>
<td>Panel B. before transaction costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000-2009</td>
<td>89.56</td>
<td>61.60</td>
<td>1.45</td>
<td></td>
</tr>
<tr>
<td>2000-2007</td>
<td>87.34</td>
<td>64.11</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>98.41</td>
<td>50.41</td>
<td>1.95</td>
<td></td>
</tr>
<tr>
<td>Panel B. after transaction costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000-2009</td>
<td>83.92</td>
<td>61.61</td>
<td>1.36</td>
<td></td>
</tr>
<tr>
<td>2000-2007</td>
<td>81.52</td>
<td>64.12</td>
<td>1.27</td>
<td></td>
</tr>
<tr>
<td>2008-2009</td>
<td>93.45</td>
<td>50.44</td>
<td>1.85</td>
<td></td>
</tr>
</tbody>
</table>

This table shows the Sharpe ratios for two pairs strategies, Portfolio A and B, for comparison purpose. Panel A and B report figures without and with transaction costs, respectively.

### Table 9: Ability to signal price dynamics (January 2000-December 2009)

<table>
<thead>
<tr>
<th></th>
<th>Long positions</th>
<th>Short positions</th>
<th>% of negative returns</th>
<th>Annual total negative returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of winner</td>
<td>% of loser</td>
<td>% of winner</td>
<td>% of loser</td>
</tr>
<tr>
<td>Portfolio A</td>
<td>51.46%</td>
<td>48.54%</td>
<td>51.89%</td>
<td>48.11%</td>
</tr>
<tr>
<td>Portfolio B</td>
<td>52.42%</td>
<td>47.37%</td>
<td>51.97%</td>
<td>48.03%</td>
</tr>
</tbody>
</table>

This table shows the ability of pairs strategies to predict price evolution. For either long or short positions, the percentage of winning/losing bets are calculated. The percentage of trades with negative return is also computed. Both calculations only take into account trading days with opening positions. In absolute terms, we provide the average of total losses on a yearly basis.
Table 10: Cumulative profits and Sharpe ratios of cointegration-based pairs strategy (Portfolio A) in four scenarios (January 2000-December 2009)

<table>
<thead>
<tr>
<th>Threshold Rule</th>
<th>Cumulative Profits</th>
<th>Sharpe Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long Positions</td>
<td>Short Positions</td>
</tr>
<tr>
<td>1 standard deviation</td>
<td>142.02</td>
<td>584.97</td>
</tr>
<tr>
<td>1.65 standard deviations</td>
<td>15.39</td>
<td>361.66</td>
</tr>
<tr>
<td>2 standard deviations</td>
<td>38.58</td>
<td>357.22</td>
</tr>
<tr>
<td>2.58 standard deviations</td>
<td>56.76</td>
<td>92.87</td>
</tr>
</tbody>
</table>

This table presents cumulative profits and Sharpe ratios of cointegration-based pairs strategy, represented as Portfolio A, in four distinct scenarios. Cumulative profits are also separated to show the returns to the long and short positions.

Table 11: Annual simplified Sharpe ratio and annual cumulative arbitrage profits for three market portfolios and two pairs portfolios (January 2010-December 2013)

<table>
<thead>
<tr>
<th></th>
<th>Average Profit</th>
<th>Average Volatility</th>
<th>Sharpe ratio</th>
<th>Cumulative Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daily</td>
<td>Annual</td>
<td>Daily</td>
<td>Annual</td>
</tr>
<tr>
<td>(i) DAX</td>
<td>0.0005</td>
<td>0.1353</td>
<td>0.0129</td>
<td>0.2034</td>
</tr>
<tr>
<td>(ii) EURO STOXX50</td>
<td>0.0001</td>
<td>0.0369</td>
<td>0.0142</td>
<td>0.2259</td>
</tr>
<tr>
<td>(iii) S&amp;P 500</td>
<td>0.0005</td>
<td>0.1362</td>
<td>0.0105</td>
<td>0.1667</td>
</tr>
</tbody>
</table>

Panel A. Excess returns before transaction costs

|                       | Daily          | Annual             | Daily        | Annual             | Daily          | Annual             |                  |
|(iv) Portfolio A       | 0.10           | 25.00              | 1.25         | 19.77              | 1.27           | 103.75             |
| (v) Portfolio B       | 0.17           | 42.36              | 2.56         | 40.68              | 1.04           | 174.97             |

Panel B. Excess returns after transaction costs

|                       | Daily          | Annual             | Daily        | Annual             | Daily          | Annual             |                  |
|(iv) Portfolio A       | 0.09           | 23.00              | 1.24         | 19.75              | 1.16           | 95.09              |
| (v) Portfolio B       | 0.15           | 39.03              | 2.56         | 40.62              | 0.96           | 161.38             |

This table reports mean profits, volatility (measured by the standard deviation), simplified Sharpe Ratios and annualized cumulative profits gained with the five proposed strategies over the extended sample period. (i), (ii) and (iii) are buy-and-hold strategies of market portfolios. That is, €1 investment in DAX index, €3 in EURO STOXX50 index and $1 in S&P500 index, respectively. And (iv) and (v) are equal-weighted portfolios performing pairs trading strategies. Portfolio A is constructed with 12 cointegrated pairs, whereas Portfolio B is formed with the top 12 pairs ranked by the minimum distance method. For each pair in either portfolio, €1 investment [with opposite positions] in pairwise equities is conducted. Zero interest rates are assumed. Panel A reports figures before transaction costs. Panel B reduces profits in A by one half of the sum of bid-ask spreads.
Table 12: Ability to signal price dynamics (January 2010-December 2013)

<table>
<thead>
<tr>
<th></th>
<th>Long Positions</th>
<th>Short Positions</th>
<th>% of negative returns</th>
<th>Annual total negative returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of winner</td>
<td>% of loser</td>
<td>% of winner</td>
<td>% of loser</td>
<td></td>
</tr>
<tr>
<td>Portfolio A</td>
<td>50.64%</td>
<td>49.36%</td>
<td>50.53%</td>
<td>49.47%</td>
</tr>
<tr>
<td>Portfolio B</td>
<td>53.02%</td>
<td>47.14%</td>
<td>49.23%</td>
<td>50.77%</td>
</tr>
</tbody>
</table>

This table shows the ability of pairs strategies to predict price evolution. For either long or short positions, the percentage of winning/losing bets are calculated. The percentage of trades with negative return is also computed. Both calculations only take into account trading days with opening positions. In absolute terms, we provide the average of total losses on a yearly basis.

Table 13: Cumulative profits and Sharpe ratios of cointegration-based pairs strategy (Portfolio A) in four scenarios (January 2010-December 2013)

<table>
<thead>
<tr>
<th>Threshold Rule</th>
<th>Long Positions</th>
<th>Short Positions</th>
<th>Total Cumulative Profits</th>
<th>Sharpe Ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 standard deviation</td>
<td>33.62</td>
<td>70.13</td>
<td>103.75</td>
<td>1.27</td>
</tr>
<tr>
<td>1.65 standard deviations</td>
<td>-22.40</td>
<td>79.97</td>
<td>57.57</td>
<td>0.94</td>
</tr>
<tr>
<td>2 standard deviations</td>
<td>-0.69</td>
<td>40.93</td>
<td>40.24</td>
<td>0.77</td>
</tr>
<tr>
<td>2.58 standard deviations</td>
<td>2.87</td>
<td>36.56</td>
<td>39.42</td>
<td>0.91</td>
</tr>
</tbody>
</table>

This table presents cumulative profits and Sharpe ratios of cointegration-based pairs strategy, represented as Portfolio A, in four distinct scenarios. Cumulative profits are also separated to show the returns to the long and short positions.
Figure 1: Time-series plots of daily normalized prices of equity pairs (January 2000-December 2009)

This figure plots normalized prices for pairwise equities identified from the constituents of EURO STOXX 50 index.
Figure 2: Price spreads between Air Liquide and BASF and pairs trading establishment (2000-2006 and 2007-2009, respectively)

This figure illustrates how to perform pairs trading strategy using the cointegrated pair, Air Liquide and BASF.
This figure plots the price spread between pairwise equities from the constituents of EURO STOXX 50 index.
Figure 4: Cumulative excess returns of pairs portfolio (Portfolio A) VS. EURO STOXX 50 (January 2000-December 2009)

This figure plots cumulative excess returns of cointegration-based pairs strategy (Portfolio A) against the movements of EURO STOXX 50 index.