

# Arbitragem Estatística, Estratégia Long-Short Pairs Trading, Abordagem com Cointegração Aplicada ao Mercado de Ações Brasileiro

João F. Caldeira

*Departamento de Economia, Universidade Federal do Rio Grande do Sul (UFRGS),  
Brasil*

---

## Resumo

Este artigo aplica testes de cointegração para identificar pares de ações a serem usados numa estratégia *pairs trading*. Além de estimar as relações de equilíbrio de longo prazo e modelar os resíduos resultantes emprega-se um indicador de rentabilidade em simulações dentro da amostra para selecionar pares que irão compor uma carteira *pairs trading*. A lucratividade da estratégia é avaliada usando dados do mercado de ações brasileiro no período de Janeiro de 2005 e Dezembro de 2009. Os resultados obtidos mostram que a estratégia alcança rentabilidade média de 17,49% ao ano, índice de Sharpe de 1.29 e baixa correlação com o mercado, reforçando o uso da metodologia de cointegração no desenvolvimento de estratégias quantitativas. Arbitragem estatística é também uma forma de testar a hipótese de eficiência de mercado que elimina o dilema tradicional relativo ao conjunto de hipóteses envolvidos nos testes e sua implementação é independente de um modelo de equilíbrio para os retornos dos ativos.

*Palavras-chave:* *Pairs Trading*, Arbitragem Estatística, Cointegração, Estratégia Neutra ao Mercado, Eficiência de Mercado

*Classificação JEL:* C58, G11, G17

---

## Abstract

This article applies cointegration tests to identify stocks to be used in pairs trading strategy. In addition to estimating long-term equilibrium and model the resulting residual, we use an indicator of profitability in in-sample simulations to select stock pairs to compose a pairs trading portfolio. We assessed the profitability of the strategy with data from the Brazilian market between January 2005 and December 2009. Considering we reached an average return of 17.49% per year, Sharpe Ratio of 1.29 and low correlation with the market, the results reinforce the use of cointegration in quantitative management. Statistical arbitrage circumvents the joint hypothesis dilemma of traditional market efficiency tests because its definition is independent of any equilibrium model and its existence is incompatible with market efficiency.

## 1. Introdução

O tema central para investir no mercado de ações, do ponto de vista da precificação de ativos consiste em vender ativos sobrevalorizados e comprar ativos subvalorizados. No entanto, para afirmar que o preço de um ativo está elevado ou baixo é preciso conhecer o valor do ativo em termos absolutos. Estratégias *pairs trading* constituem uma tentativa de resolver esta questão através de precificação relativa. O pressuposto subjacente da estratégia *pairs trading* é que dois ativos com características similares devem exibir preços próximos e devem compartilhar a uma tendência estocástica comum. Neste caso, o preço específico de cada ativo não é o mais importante. Os preços individuais podem estar errados, o que importa é que os preços dos dois ativos sejam iguais ou bastante próximos. Se em algum momento os preços se distanciam pode ser que um dos ativos esteja sobreprecificado, o outro esteja subprecificado, ou que a distorção nos preços seja uma combinação das duas coisas.

Em estratégias *pairs trading* o objetivo não é definir tendências para ativos específicos, mas sim estabelecer uma relação de equilíbrio de longo prazo entre um par de ações. A ideia por trás da estratégia *pairs trading* envolve inicialmente identificar pares de ações com histórico de movimentos similares nos preços. Então, sempre que ocorrer divergência suficiente entre os preços das ações que compõem o par assume-se uma posição comprada e outra vendida, simultaneamente, apostando que a divergência nos preços do par é temporária e que convergirá ao longo do tempo. Um dos objetivos com estratégias *pairs trading* é a neutralidade ao mercado, de forma que a posição *long-short* assumida nos ativos envolvidos tenha beta próximo a zero ou estatisticamente não-significante, e, portanto, exposição mínima ao mercado. Consequentemente, os retornos obtidos com a estratégia devem exibir baixa correlação com o mercado, característica típica de estratégias neutras ao mercado.

Estratégias envolvendo pares de ações são amplamente utilizadas na indústria financeira desde meados da década de 1980. A essência de uma estratégia *pairs trading* é explorar desvios de curto prazo nas relações de equilíbrio de longo prazo entre os preços de ações que apresentam movimentos relacionados. Trata-se, portanto, de uma estratégia de valor relativo, uma vez que o *spread* mede o grau do erro de precificação de um ativo em relação ao outro. O termo arbitragem estatística engloba uma variedade de estratégias de investimento, as quais tem como característica principal o uso de ferramentas estatísticas para gerar excessos de retornos. Outra característica dessa classe de estratégias é que os sinais para as operações são sistemáticos, ou baseados em regras, ao invés de fundamentos.

---

\* Recebido em dezembro de 2010, aprovado em setembro de 2012.  
E-mail address: joao.caldeira@ufrgs.br

No Brasil, a utilização desse tipo de estratégia em maior escala e de forma regulamentada pela Comissão de Valores Mobiliários (CVM) só foi possível com a criação do Banco de Títulos CBLC (BTC) em 1996<sup>1</sup> e com o crescimento da indústria de fundos multimercado e fundos *long-short*. Os fundos multimercados são veículos de investimento que podem aplicar seus recursos em diversas classes de ativos e usar alavancagem para auferir ganhos. Os fundos de investimento *long-short* aplicam seus recursos em renda variável e buscam obter ganhos através de operações com ações assumindo posições compradas e vendidas.

Os métodos tradicionalmente empregados no desenvolvimento de estratégias *pairs trading* com ações buscam identificar os pares com base em correlação ou em outras regras de decisão não-paramétricas. A presença de cointegração possibilita obter uma combinação linear entre os preços de duas ações de forma que o portfólio combinado seja um processo estacionário. O portfólio é formado comprando a ação relativamente sub-valorizada e vendendo a ação relativamente sobrevalorizada. Se as séries de preços de duas ações apresentam uma relação de equilíbrio de longo prazo, então os desvios desse equilíbrio devem ser temporários e espera-se que convirjam, assegurando reversão à média.

Algumas questões básicas precisam ser consideradas no desenvolvimento de estratégias *pairs trading*, por exemplo:

- (i) como identificar os pares;
- (ii) quando o portfólio combinado se distancia suficientemente da relação de equilíbrio para que seja aberta uma posição; e
- (iii) quando a posição deve ser encerrada.

Da perspectiva do gerenciamento de risco é importante especificar um período de tempo máximo que se pode manter a posição em aberto. Também deve-se definir o nível máximo de valor em risco permitido (VaR), e se possível, definir medidas de redução de risco, como limites de perda.

Apesar de serem amplamente utilizadas pelos participantes do mercado, surpreendentemente, as estratégias de valor relativo têm recebido pouca atenção da comunidade acadêmica. Uma das principais referências sobre o assunto é o artigo de Gatev et alii (2006), em que os autores usaram dados diários do mercado de ações norte-americano. Foram empregadas regras de *trading* simples e alcançaram retornos de 11% ao ano, com baixa exposição às fontes de risco sistemático. Gatev et alii (2006) usaram uma abordagem baseada na distância, em que o co-movimento de um par é medido pela *distância* relativa, ou, a soma do quadrado das diferenças entre duas séries de preços normalizadas.<sup>2</sup> Elliott et alii (2005) propuseram uma abordagem estocástica para o *spread* no contexto de *pairs trading*, a qual modela o comportamento de reversão à média do *spread* em tempo contínuo, sendo o *spread* definido como a diferença entre os preços das duas ações. Do e Hamza (2006) propuseram uma estratégia *pairs trading* baseada na modelagem

<sup>1</sup> O mercado de aluguel de ações no Brasil era praticamente insignificante na década de 1990. Entretanto, acompanhando o crescimento da indústria de fundos, o seu crescimento tem sido muito expressivo nos últimos anos.

<sup>2</sup> Aplicações similares são encontradas em Nath (2006) e Do e Hamza (2006).

dos desvios dos retornos, ao invés de trabalhar diretamente com os preços em nível. Avellaneda e Leey (2010) empregaram análise de componentes principais (PCA) e ETFs (Exchange Trade Funds) setoriais no contexto de estratégias de arbitragem estatística, onde os componentes de retornos residuais ou idiossincráticos são tratados como processos com reversão à média.

A abordagem de cointegração apresentada inicialmente em Vidyamurthy (2004) busca parametrizar estratégias *pairs trading* explorando a possibilidade de cointegração entre pares de ações (ver Engle e Granger 1987; Engle e Yoo 1987). Cointegração possibilita incorporar reversão à média em uma estrutura de *pairs trading*, a qual é a mais importante relação estatística requerida para o sucesso da estratégia (Avellaneda e Leey 2010). Outros exemplos de aplicações de cointegração no contexto de estratégias *pairs trading* são Enders e Granger (1998), Alexander e Dimitriu (2005b), Herlemont (2000) e Lin et alii (2006). Outros autores aplicaram a metodologia de cointegração para otimização de carteiras no contexto de estratégias *index tracking* e estratégias *long-short* com carteiras de ações (por exemplo Caldeira e Portugal 2010; Alexander e Dimitriu 2005a; Dunis e Ho 2005). Aplicações envolvendo dados de alta frequência são encontradas em Burges (2003), Lin et alii (2006), Galenko e Popova (2012).

Estratégias neutras ao mercado são geralmente identificadas como uma modalidade de investimento que apresentam propriedades atrativas, tais como baixa exposição ao fator de risco de mercado e relativamente baixa volatilidade (Capocci 2006). Recentemente o desempenho destas estratégias apresentou considerável deterioração, conforme relatado em Gatev et alii (2006) e Khandani e Lo (2007). Na mesma linha o *Hedge Fund Equity Market Neutral Index* (Índice HFRXEMN na Bloomberg) não mostra as supostas qualidades pelas quais as estratégias neutras ao mercado são atrativas, isto é, crescimento estável e baixa volatilidade.

Este artigo implementa uma estratégia de arbitragem estatística com pares de ações e estuda sua lucratividade usando dados do mercado de ações brasileiro. O objetivo é aplicar o conceito de cointegração para derivar relações dinâmicas de equilíbrio de longo prazo entre pares ações negociadas no mercado brasileiro, que inerentemente implique em reversão à média nas séries de preços componentes. Inicialmente seleciona-se os pares de ações a serem usados na estratégia *pairs trading* através dos testes de cointegração de Johansen e Engle-Granger, ver Juselius (2006). A série dos resíduos da equação de cointegração, ou *spread*, que descreve a dinâmica do portfolio com reversão à média é usada para modelar a meia-vida do processo que é utilizada nas regras de *trading*. Posteriormente é empregado um indicador de rentabilidade em simulações dentro da amostra no processo de seleção dos pares. Por fim, calcula-se o *spread* padronizado (*score* padronizado), que é usado como critério para definir os momentos de abertura e encerramento das posições.

Os dados utilizados consistem dos preços diários de fechamento das principais ações constituintes do índice Ibovespa durante o período de janeiro de 2005 a dezembro de 2009, perfazendo um total de 1250 observações diárias. Empregando

parâmetros definidos em simulações dentro da amostra, a rentabilidade da estratégia é avaliada em simulações fora da amostra. Alcançando rentabilidade média de 17.49% ao ano, índice de Sharpe de 1.29, baixa correlação com o mercado e volatilidade em níveis relativamente baixos, os resultados obtidos reforçam o uso do conceito de cointegração no desenvolvimento de estratégias quantitativas que podem ser usadas tanto por investidores individuais, gestores institucionais e fundos multimercados..

Além desta introdução, o artigo está organizado em mais 4 seções. Na Seção 2 são discutidos os aspectos teóricos que dão sustentação à estratégia de arbitragem estatística, precificação relativa e eficiência de mercado. Na Seção 3 são apresentadas de forma mais detalhada os conceitos de arbitragem estatística e estratégias *pairs trading*, além de contextualizar o uso da cointegração nesta classe de estratégias. Na Seção 4 descreve-se o modelo proposto para implementação da estratégia, a base de dados e são analisados os resultados obtidos. Na Seção 5 é apresentada uma conclusão dos resultados empíricos e indicação de passos que podem potencializar os resultados da estratégia.

## 2. Arbitragem Estatística, Precificação Relativa e Eficiência de Mercado

O uso do termo “estatística” no contexto de estratégias de investimento é um indicativo do caráter especulativo da estratégia de investimento em questão. Baseia-se no pressuposto de que os padrões observados no passado vão ser repetir no futuro. Isto se opõe às estratégias de investimentos fundamentalistas que buscam explorar e prever o comportamento das forças econômicas que influenciam os preços das ações. Assim, arbitragem estatística é uma abordagem puramente estatística que tem por objetivo explorar ineficiências dos mercados de ações definidas como desvios do equilíbrio de longo-prazo entre preços de ações observados no passado. Uma das principais motivações para implementação de arbitragem estatística com técnicas econométricas avançadas é que oferece uma forma arbitrária de computar o *spread*, cuja cotação é baseada na normalização dos preços das ações. No caso da análise de cointegração a normalização é baseada na estimação do vetor de cointegração (a esse respeito ver, por exemplo, Bessler e Yang 2003; Yang et alii 2004). A inacurácia das estimativas pode impactar negativamente a rentabilidade de uma estratégia de trading.

Conforme Gatev et alii (2006), a precificação de ativos pode ser vista em termos absolutos ou em termos relativos. Precificação absoluta envolve atribuir valor aos ativos a partir de fundamentos, como o fluxo de caixa descontado, o que é um processo difícil e com considerável margem para erros. Já precificação em termos relativos significa que dois ativos que são considerados substitutos próximos deveriam ser vendidos pelo mesmo preço – não define qual deveria ser o preço de cada ativo. Assim, precificação relativa admite a possibilidade da ocorrência de bolhas na economia, mas não necessariamente arbitragem ou especulação rentável.

A Lei do Preço Único [LOP] e uma “quase” LOP são aplicáveis à precificação relativa – mesmo se os preços estiverem errados. Ingersoll (apud Gatev et alii 2006) define a LOP como uma proposição: dois ativos com o mesmo *payoff* em todos os estados de natureza deveriam ter o mesmo valor corrente. Em outras palavras, dois ativos com os mesmos preços em todos os estados possíveis deveriam ser vendidos pelo mesmo valor.

Um grande número de estudos empíricos conclui que o mercado de ações parece contradizer a hipótese de mercados eficientes. Por exemplo, Jegadeesh e Titman (1993) avaliam uma estratégia de investimento que consiste na compra de ações que apresentam bom desempenho passado e na venda de ações que apresentaram fraco desempenho. Os autores documentam excessos de retornos médios de 12% ao ano, onde os retornos em excesso são definidos em relação a um padrão CAPM. Lakonishok et alii (1994) chegam a conclusões similares através da compra de ações de *valor* e venda de ações que chamam de *glamour*, que são identificadas com variáveis como índice de preço/lucro, dividendos, valor de mercado, fluxo de caixa e crescimento de vendas. Chan et alii (1996) também confirmam os resultados através de portfólios formados com base em anúncio de resultados. No entanto, todos os estudos acima mencionados sofrem do problema hipótese conjunta. Fama (1998) adverte contra a rejeição de eficiência do mercado, porque os testes anteriores dependem de um modelo específico para os retornos de equilíbrio. De acordo com Fama (1998) a maioria das anomalias de longo prazo também são sensíveis à metodologia estatística utilizada e, portanto, inferências usando retornos de longo prazo (por exemplo, mais de 5 anos) devem ser vistos com ressalvas.

Em vista deste dilema fundamental, Hogan et alii (2004) propõem testar a eficiência do mercado através da verificação se as anomalias constituem ou oportunidades de arbitragem estatística. Os autores argumentam que a arbitragem estatística é independente de qualquer modelo de equilíbrio ou formulação para os retornos esperados e sua existência contradiz a eficiência de mercado. A estratégia de investimento proposta neste artigo pode ser considerada como um teste da eficiência de mercado com base no conceito de arbitragem estatística. De maneira simples, arbitragem estatística envolve uma oportunidade de negócio que gera lucro livre de risco. Como tal, arbitragem estatística é uma extensão natural das estratégias de investimento utilizadas na literatura existente sobre anomalias (Jarrow et alii 2012). Mais importante, arbitragem estatística é definido sem referência a qualquer modelo de equilíbrio, portanto, sua existência é incompatível com o equilíbrio do mercado e, conseqüentemente, a eficiência do mercado (ver Hogan et alii 2004). Assim, arbitragem estatística permite a rejeição da eficiência de mercado sem evocar ao conjunto de hipóteses de um modelo de equilíbrio. Ao invés disso, o conjunto de hipóteses de um modelo de equilíbrio é substituído por um processo estocástico assumido para os lucros da estratégia.

Neste artigo é implementa uma estratégia de arbitragem estatística através de um algoritmo para seleção dos pares baseado na análise de cointegração, com isso o critério consiste em identificar pares de ações que apresentem uma relação estável historicamente. Assim, é analisado se as anomalias identificadas nos preços relativos

através da análise do *spread* constitui oportunidade de arbitragem estatística. São formados pares de ações de cujos preços historicamente andam juntos, por que a Lei do Preço Único, neste caso, sugere que em um mercado eficiente seus preços devem se mover juntos. Neste contexto, o presente estudo pode ser visto com um teste da LOP, quase-LOP (Gatev et alii 2006) e um teste de eficiência (ver Hogan et alii 2004) no mercado de ações brasileiro, sob certas condições de estacionariedade. A aplicação do conceito de cointegração para lidar com ineficiência dos mercados não é nova e há uma extensa discussão a esse respeito, por exemplo, Carcano et alii (2005); Gatev et alii (2006); Gatarek et alii (2011); Serban (2010); Triantafyllopoulos e Montana (2011), entre outros.

### 3. A Abordagem da Cointegração: Pairs Trading e Arbitragem Estatística

#### 3.1. *Contextualização histórica*

O pioneirismo na utilização de estratégias *pairs trading* estritamente estatísticas é atribuído ao *quant* de Wall Street Nunzio Tartaglia, que trabalhou no banco Morgan Stanley em meados da década de 1980. Tartaglia montou um grupo de matemáticos, físicos e cientistas da computação com o objetivo de desenvolver estratégias de arbitragem estatística usando o estado da arte das técnicas estatísticas. As estratégias desenvolvidas pelo grupo foram automatizadas a tal ponto que podiam gerar operações de forma mecanizada, se necessário, executá-las perfeitamente através de sistemas de negociação automatizados.

Uma das estratégias de investimento empregadas pelo grupo envolvia a negociação com pares de ações. O processo consistia em identificar pares de ações cujos os preços se moviam juntos. Quando era identificada uma anomalia na relação, o par era negociado com a idéia de que a anomalia se corrigiria. Tal estratégia ficou conhecida como *pairs trading*. Tartaglia e seu grupo empregaram a estratégia *pairs trading* com grande sucesso ao longo de 1987. Apesar do grupo ter sido desmontado em 1989, após dois anos de resultados ruins, a estratégia *pair trading* se tornou cada vez mais popular entre operadores individuais, investidores institucionais e gestores de *hedge-funds*.

Recentemente, em decorrência da crise dos mercados financeiros, foi amplamente divulgado na mídia especializada que o ano de 2007 foi especialmente desafiador para os *hedge-funds* quantitativos, (ver Khandani e Lo 2007; Avellaneda e Leey 2010), sobretudo para as estratégias de arbitragem estatística. A estratégia proposta aqui é analisada no período em questão e os resultados encontrados corroboram com os encontrados por outros autores.

### 3.2. *Pairs Trading e Arbitragem Estatística*

A arbitragem, no seu sentido puro, envolve a obtenção de lucro econômico sem risco e sem o desembolso de caixa. A Hipótese de Mercados Eficientes (HME), segundo a qual os preços dos ativos refletem os valores de seus fundamentos, pressupõe que os investidores são racionais e rapidamente incorporam nos preços dos ativos todas as informações disponíveis, levando os preços para os seus valores intrínsecos. Quando os preços dos ativos se desviam dos seus valores intrínsecos, as operações de arbitragem promovem a sua convergência, Aldrighi e Milanez (2005).

A existência de oportunidades de arbitragem estatística é determinada pelos lucros obtidos com uma estratégia de investimento. Formalmente, Hogan et alii (2004) definem uma estratégia de arbitragem estatística como tendo que satisfazer a quatro condições:

- (i) envolve custo inicial zero e é uma estratégia que se auto-financia;
- (ii) no limite tem lucro esperado descontado positivo;
- (iii) apresenta probabilidade de perda convergindo para zero; e
- (iv) variância convergindo para zero se a probabilidade de perda não vai para zero em tempo finito.

Isto é, a quarta condição aplica-se apenas quando existe sempre uma probabilidade positiva de perder de dinheiro.

Jacobs et alii (1993) definem estratégias *long-short* com ações como sendo neutras ao mercado. Estratégias neutras ao mercado envolvem posições compradas e vendidas com equivalentes exposições ao risco de mercado em todo o tempo. Isto é alcançado igualando os betas de mercado ponderados das posições comprada e vendida dentro do portfólio. Esta abordagem elimina exposição ao risco direcional de mercado, tal que o retorno obtido não deve apresentar correlação com o índice de referência do mercado, o que é equivalente a um portfólio beta-zero. Os retornos são gerados pelo isolamento do alfa, ajustado pelo risco. Os recursos obtidos com a posição vendida geralmente são integralmente usados para financiar a parte comprada. Segundo Fund e Hsieh (1999), uma estratégia é dita neutra ao mercado se seu retorno é independente do retorno do mercado relevante. Fundos neutros ao mercado buscam ativamente evitar exposição a fatores de risco, apostando em movimentos dos preços relativos. Um portfólio neutro ao mercado exhibe risco sistemático igual a zero e é praticamente interpretado como tendo beta de mercado igual a zero. Já Alexander e Dimitriu (2005a), afirmam que estratégias *long-short* com ações, apesar de serem naturalmente vistas como neutras ao mercado, se não forem construídas especificamente para ter beta de mercado igual a zero, não necessariamente são neutras ao mercado.

*Pairs trading* envolve a venda de uma ação relativamente sobrevalorizada e a compra de outra ação subvalorizada com a idéia de que a distorção nos preços irá se corrigir no futuro. A distorção mútua entre os preços das duas ações é capturada pelo *spread*. Quanto maior o *spread* maior a magnitude da distorção nos preços e maior o lucro potencial. Assim, uma posição *long-short* com duas

ações é construída de forma a ter um beta mínimo, e, portanto, mínima exposição ao mercado. Consequentemente, os retornos da operação são não correlacionados com os retornos do mercado, uma característica típica de estratégias neutras ao mercado. Uma extrapolação natural de *pair trading* consiste em operar um grupo de ações contra outro grupo de ações, ou *pair trading* generalizado, (ver, por exemplo, Alexander e Dimitriou 2005a; Dunis e Ho 2005; Avellaneda e Leey 2010; Caldeira e Portugal 2010).

O processo de precificação de ativos pode ser visto em termos absolutos ou relativos. Em termos absolutos a precificação de ativos se dá através de fundamentos, como fluxo de caixa futuro descontado, por exemplo. Precificação relativa significa que dois ativos que são substitutos próximos um do outro deveriam ser negociados a preços parecidos – não diz qual deve ser o preço. Assim, precificação em termos relativos admite a ocorrência de bolhas na economia. Em mercados eficientes, a precificação com base no modelo CAPM (*Capital Asset Pricing Model*) e a lei do preço único requer preços próximos para ativos equivalentes ao longo do tempo. O *spread* de preços entre ativos substitutos próximos deve obedecer a um equilíbrio estável de longo prazo ao longo do tempo. Hendry e Juselius (2001) usam este princípio para mostrar que os desvios de curto prazo da condição de preços equivalentes pode criar oportunidades de arbitragem dependendo da duração do desvio dos preços.

Quando é identificado um desvio suficientemente grande da relação de equilíbrio de longo prazo entre os preços (medido pelo *spread*), abre-se uma posição comprando a ação relativamente subvalorizada e vendendo a ação relativamente sobrevalorizada, simultaneamente. A posição é encerrada quando o *spread* converge para o nível de equilíbrio de longo prazo. O lucro líquido da operação é composto pela soma dos resultados das partes comprada e vendida, calculado como a diferença entre os preços de abertura e preços de fechamento (descontados os custos envolvidos na operação).

Nath (2006) e Lin et alii (2006) argumentam que a implementação da estratégia na prática é mais complicado e os retornos são menores do que em teoria devido à presença de fricções e imperfeições de mercado, como os saltos dos preços das ações no dia em que passam a ser negociadas ex-dividendos, custos de transação, custos de financiamento e impostos. Além disso, Nath (2006) alerta para a presença dos riscos de microestrutura de mercado, mudanças nos fundamentos econômicos e a presença de riscos provenientes dos movimentos dos investidores irracionais (*noise traders*). Esse último fator faz com que, mesmo que os fundamentos econômicos permaneçam inalterados, os desvios de apreçamento aumentem ainda mais em relação ao momento em que se abriu a posição.

Com base na discussão acima, nota-se que parte decisiva para obter sucesso com estratégias *pairs trading* consiste na identificação dos pares de ações. Gatev et alii (2006) empregam uma abordagem puramente empírica para atingir esse objetivo. A metodologia de escolha dos pares se baseia inteiramente nos movimentos históricos dos preços das ações. Os resultados são comparados com aqueles obtidos em um experimento de bootstrap. Além do conjunto de pares definidos com base

nos preços históricos, um outro conjunto de pares é definindo aleatoriamente. Os resultados das operações com pares executados com base na abordagem empírica foram, então, comparados com os resultados das operações com os pares definidos aleatoriamente. A diferença entre os retornos obtidos com os dois grupos foi estatisticamente significativa, e o retorno alcançado com a estratégia que empregou pares definidos metodologicamente foi superior à situação onde os pares foram definidos aleatoriamente.

### 3.3. A abordagem da cointegração

A partir dos trabalhos seminais de Markowitz (1959), Sharpe (1964), Lintner (1965) e Black (1972), a ferramenta estatística fundamental para otimização de portfólios é a análise da correlação entre os retornos dos ativos. Os modelos de otimização para construção de portfólio, em sua maioria, focam em minimizar a variância de uma carteira de ativos para um dado nível de retorno, com restrições adicionais a respeito dos investimentos permitidos, restrições quanto a venda a descoberto e os custos de transação associados ao rebalanceamento da carteira.

Na última década o conceito de cointegração foi amplamente aplicado em econometria financeira, em conexão com análise de séries temporais e macroeconomia. Tem se colocado, portanto, como uma técnica estatística extremamente útil, dada a possibilidade de se aplicar métodos de estimação simples (tais como mínimos quadrados e máxima verossimilhança) a séries não-estacionárias. A observação fundamental que justifica a aplicação do conceito de cointegração na análise de preços de ações é que um sistema envolvendo preços de ações não-estacionários em nível, pode ter uma tendência estocástica comum, (Stock e Watson 1988).

Quando comparado ao conceito de correlação, a principal vantagem da cointegração está na possibilidade de se utilizar integralmente o conjunto de informações contidas nas séries financeiras em nível. Alexander e Dimitriu (2005a) demonstram que o lucro potencial da arbitragem entre duas ações depende necessariamente da presença de um *spread* de equilíbrio de longo prazo entre os preços das ações, da ocorrência de desvios de curto prazo e da convergência para o equilíbrio. Nesta situação, a técnica estatística usada para implementar uma estratégia de *pair trading* deve ser capaz de fornecer um efetivo modelo para o comportamento do preço relativo ao longo do tempo e detectar os valores da relação de equilíbrio, além de uma medida dos desvios de curto prazo da relação de equilíbrio. (Alexander e Dimitriu 2005a,b; Gatev et alii 2006; Caldeira e Portugal 2010), sugerem que a metodologia de cointegração oferece uma estrutura mais adequada para estratégias de arbitragem estatística que as demais.

A abordagem de cointegração, descrita em Vidyamurthy (2004), é uma tentativa de parametrizar estratégias *pairs trading* explorando a possibilidade de cointegração. Cointegração é uma relação estatística onde duas séries de tempo que são integradas de mesma ordem,  $d$ , podem ser combinadas linearmente para

produzir uma única série integrada de ordem  $d-b$ , onde ( $d > b > 0$ ). Nas aplicações envolvendo *pairs trading*, geralmente são considerados casos onde séries de preços  $I(1)$  são combinadas para produzir séries temporais estacionárias de um portfólio, ou  $I(0)$ .

DEFINIÇÃO 2. 1. Uma série de tempo  $X_t$  é dita  $I(1)$  se a primeira diferença da série é estacionária, denotada por  $I(0)$ .

Muitas séries de preços de ações são  $I(1)$ . Portanto, a seguinte definição de cointegração é baseada em séries  $I(1)$ .

DEFINIÇÃO 2. 2. Seja  $X_{1,t}, X_{2,t}, \dots, X_{k,t}$  uma sequência de séries de tempo  $I(1)$ . Se existem números reais  $\beta_1, \beta_2, \dots, \beta_k$ , diferentes de zero, tal que:

$$\beta_1 X_{1,t} + \beta_2 X_{2,t} + \dots + \beta_k X_{k,t} \quad (1)$$

forme uma série  $I(0)$ , então  $X_{1,t}, X_{2,t}, \dots, X_{k,t}$  são ditas cointegradas.

Em outras palavras, um conjunto de séries  $I(1)$  são ditas cointegradas se existe uma combinação linear dessas séries que é estacionária. Séries de preços de ações cointegradas possuem uma relação estável de equilíbrio de longo prazo, com a propriedade de reversão à média. Por definição, alguma combinação linear de séries de preços cointegradas é estacionária e sempre reverte para a média de longo prazo. Para mais detalhes a respeito da análise de cointegração, ver Hamilton (1994), Enders (2002), Hendry e Juselius (2000) e Hendry e Juselius (2001).

Séries de tempo cointegradas podem ser representadas através de um vetor de correção de erros (*VECM* – Teorema da Representação de Granger), no qual a dinâmica das séries de tempo é modelada como função de seus valores defasados, dos valores defasados de seu par cointegrado e de um componente de correção de erro, o qual corrige os desvios da relação de equilíbrio nos períodos anteriores. Neste artigo, o modelo *VEC* é empregado para calcular a meia vida dos pares utilizados na estratégia *pairs trading*, que é utilizada para definir o período em que a operação deve ser mantida em aberto.

A idéia por traz da correção de erros é que sistemas cointegrados tem um equilíbrio de longo-prazo, isto é, a média de longo prazo da combinação linear das duas séries. Se ocorre um desvio em relação ao equilíbrio de longo prazo, uma ou as duas séries ajustam para restaurar o equilíbrio. Seja  $\varepsilon_t$  o erro correspondente à equação de cointegração de  $X_{1,t}$  em  $X_{2,t}$ . A representação de correção de erros é:

$$\Delta X_{1,t} = \alpha (X_{1,t-1} - \beta X_{2,t-1}) + \varepsilon_t \quad (2)$$

O termo  $\alpha (X_{1,t-1} - \beta X_{2,t-1})$  é a parte da correção de erros, onde  $(X_{1,t-1} - \beta X_{2,t-1})$  representa o desvio do equilíbrio de longo prazo e  $\beta$  é o coeficiente de cointegração.  $\alpha$  é a taxa de correção de erro, indicando a velocidade à qual as séries se ajustam para manter o equilíbrio, que é usado para determinar o período ótimo para se manter aberta uma estratégia baseada em reversão à média.

Note que a reversão à média de uma série de tempo pode ser modelada também pela equação chamada fórmula de Ornstein-Uhlenbeck (ver Uhlembeck e Ornstein 1930; Gillespie 1996; Mudchanatongsuk et alii 2008). Seja  $z_t$  o *spread* entre um par de ações, com propriedade de reversão à média. Então, podemos escrever:

$$dz_t = \theta(\mu - z_t) dt + \sigma \Delta W_t + \varepsilon_t \quad (3)$$

onde  $\mu$  é a média de longo prazo do *spread* entre os preços,  $z_t$  é o valor do *spread* em um dado ponto do tempo,  $\theta$  a força de reversão à média,  $\sigma$  o desvio padrão,  $W_t$  o processo de Wiener e  $\varepsilon_t$  o resíduo da regressão. Matematicamente se diz que o valor médio de  $z_t$  segue um decaimento exponencial para sua média  $\mu$ , e a meia-vida do decaimento exponencial é igual a  $\ln(2)/\theta$ , que é tempo esperado que o *spread* leve para reverter a metade do desvio inicial em relação à média. Esta meia-vida pode ser usada para determinar o período ótimo de manutenção de uma posição baseada em reversão à média.

#### 4. Modelo de *Pairs Trading*, Dados e Análise dos Resultados

A estratégia de investimento analisada neste artigo pode ser considerada neutra ao mercado. Neste sentido, busca-se identificar pares de ações que apresentem uma relação de equilíbrio de longo prazo entre os preços. Assumindo, simultaneamente, posição comprada em uma das ações e posição vendida na outra, o beta de mercado do par deve ser próximo a zero e o resultado obtido é igual ao alfa. O desafio nesta estratégia é identificar ações que tendem a andar juntas, e, portanto, são pares potenciais.

##### 4.1. *Estratégia Pairs Trading Baseada em Cointegração*

É claro que se a série de preços de uma ação é estacionária, ela é uma grande candidata para estratégias baseadas em reversão a média. Porém, a maioria das séries de preços de ações são não estacionárias. Portanto, é possível encontrar pares de ações tais que, assumindo posição comprada em uma delas e vendida na outra, o valor de mercado do par é estacionário. Neste caso, as duas séries individuais são ditas *cointegradas*. Se as séries de preços (de uma ação, um par de ações, ou um portfolio de ações) é estacionária, então uma estratégia baseada em reversão à média é seguramente rentável, ao menos enquanto a estacionariedade persistir no futuro (o que não é garantido). Geralmente, duas ações que formam um par cointegrado pertencem a um mesmo setor, porém, neste artigo não restringimos atenção a pares de ações que pertençam a um mesmo setor.

A primeira parte do desenvolvimento da estratégia *pairs trading* consiste em identificar potenciais pares de ações. Definido o conjunto de ações considerado (subseção 4.2), o primeiro passo é checar se todas as séries são integradas de mesma ordem,  $I(1)$ . Isto é feito através do teste ADF (*Augmented Dickey Fuller Test*). Passado o teste ADF, são realizados testes de cointegração em todas as combinações possíveis da amostra. Os métodos de cointegração utilizados são

o teste de cointegração de Dickey-Fuller aumentado (*Cointegrating augmented Dickey-Fuller test*) e teste de Johansen.<sup>3</sup>

Definidos os pares a serem empregados na estratégia é estimada a relação linear entre as duas séries de preços das ações através da regressão por mínimos quadrados, de onde se obtém o *hedge-ratio* (o beta da regressão):

$$P_t^l - \beta P_t^s = z_t \quad (4)$$

onde o termo do lado esquerdo representa a combinação linear das duas séries<sup>4</sup> e  $\beta$  é o coeficiente de cointegração.  $z_t$  é o valor do *spread* no tempo  $t$ ,  $P_t^l$  e  $P_t^s$  são os preços das ações  $l$  e  $s$  em  $t$ . Assim, dado que as séries são cointegradas, o *spread* tem a propriedade de reversão à média. Esta característica é de fundamental importância, pois *pairs trading* é uma aposta em que a série dos resíduos irá convergir para sua média, ou, equilíbrio de longo prazo. Nesta etapa calcula-se também a meia-vida, que basicamente nos diz a metade do tempo necessário para o *spread* convergir para a média de longo prazo.

Primeiro é estimado o *spread* da série com base na equação (4). O *spread* é então normalizado subtraindo-se sua média e dividindo pelo desvio padrão. A média e o desvio padrão são calculados com base no período dentro da amostra e são usados para normalizar *spread* tanto no período dentro da amostra quanto fora da amostra. Nesta etapa uma alternativa é usar algum critério de suavização exponencial para a média e desvio padrão, de tal forma que as duas medidas possam variar ao longo do tempo. O critério para abrir uma posição consiste em vender (comprar) o *spread* quando está a dois desvios padrões acima (abaixo) da média, a posição é liquidada quando o *spread* está a menos de 0.5 desvio da sua média, ou, decorrido o período de meia vida do *spread*.

Como critério para selecionar os 20 pares a serem utilizados na composição da carteira, usando os mesmos critérios empregados no período fora da amostra, é calculado o índice de Sharpe (*SR*) para simulações dentro da amostra e os pares são ranqueados com base no *SR*. Os 20 pares que apresentam maior *SR* nas simulações dentro da amostra são utilizados para compor uma carteira de pares a ser empregada nos testes da estratégia fora da amostra.

A estratégia implementada aqui busca ser beta-neutra, assim os valores financeiros alocados nas partes *long* e *short* podem não ser iguais.<sup>5</sup> Uma vez que o *spread* se distancia do equilíbrio de longo prazo, aposta-se que ele retornará para sua média, porém não se sabe se a convergência será proveniente da posição *long*

<sup>3</sup> São empregadas as funções *cadf* e *johansen*, disponíveis em [www.spatial-econometrics.com](http://www.spatial-econometrics.com). Uma descrição detalhada do método pode ser encontrada no manual, também disponível no mesmo site.

<sup>4</sup> O modelo é estimado sem a inclusão de termo constante. Intuitivamente falando, quando o preço de uma ação aproxima de zero, porque deveria existir um preço mínimo a partir do qual o preço da segunda ação não poderia cair? Além disso, não incluindo constante temos um modelo com menos parâmetros para serem estimados.

<sup>5</sup> Uma alternativa é definir o investimento como financeiro-neutro, assumindo volumes financeiros iguais nas partes *long* e *short*. Entretanto, na prática isto nem sempre é possível, uma vez que o investidor não pode comprar frações de uma ação.

ou *short*.<sup>6</sup> Uma aberta uma posição, o portfolio não é rebalanceado. Portanto, depois que a posição é aberta, mesmo quando os preços se movimentam e a posição pode deixar de ser neutra, o portfolio não é rebalanceado. Apenas dois tipos de transações são admitidas pela metodologia da estratégia, entrar numa nova posição, ou a liquidação total de uma posição aberta previamente.

#### 4.2. Base de dados

Os dados utilizados neste estudo consistem dos preços diários de fechamento das 50 ações com maior participação na composição do índice Ibovespa no último quadrimestre do período empregado para formação dos pares. Todas as ações utilizadas são listadas no Ibovespa, o que significa que estão entre as maiores e mais ativamente negociadas no mercado brasileiro. Esta característica é importante para uma estratégia *pairs trading*, desde que a falta de liquidez nas partes *long* e *short* constitui um risco fundamental quando se implementa esta estratégia de investimento. Os dados foram obtidos da Bloomberg, compreendendo o período de janeiro de 2005 a dezembro de 2009. Os dados são ajustados para pagamentos de dividendos e *splits* nos preços das ações.<sup>7</sup>

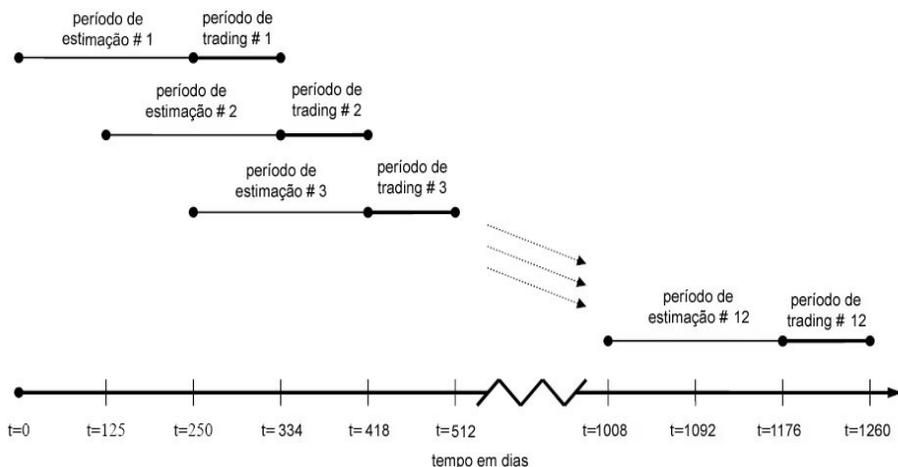
O objetivo ao usar apenas as 50 ações com maior participação no Ibovespa é trabalhar com ações que apresentam elevada liquidez e que são negociadas em todos os dias do período amostral, assegurando que a estratégia seja “tomadora de preços”. Utilizar ações com baixa liquidez pode implicar em maiores custos operacionais (*bid & ask spread*) e dificuldade para se obter o aluguel de alguma ação quando a posição for mantida em aberto por período superior a três dias úteis, quando ocorrem as liquidações das operações na CBLC – Câmara Brasileira de Liquidação e Custódia.

A Figura 1 ilustra o período de formação e período de *trading*. Tem-se um total de doze períodos de *trading*, para um total de 1250 dias de negociação na amostra, sendo que as primeiras 250 observações são usadas para período inicial de formação dos pares. A partir disso, a cada quadrimestre os parâmetros são reestimados, para coincidir com o período de recomposição do Ibovespa, os pares escolhidos são testados fora da amostra nos quatro meses seguintes, conforme Figura 1.

É comum em estratégias *pairs trading* restringir que as ações que compõem um par pertençam ao mesmo setor, por exemplo em Chan (2009) e Dunis et alii (2010), aqui esta restrição não é considerada. Portanto, pares de ações de empresas pertencentes a diferentes setores podem ser negociados, desde que satisfaçam os critérios estabelecidos.

<sup>6</sup> Não sabemos qual dos casos ocorrerá primeiro: se as ações retornam para seu equilíbrio de longo prazo porque a ação sobrevalorizada cai mais ou porque o preço da ação subvalorizada sobe mais, ou ambos apresentam a mesma performance.

<sup>7</sup> Dados diários são automaticamente ajustados pela Bloomberg.

Fig. 1. Período de Formação e Período de *Trading*

Nota: A figura mostra o período de estimação e período de trading. Nós definimos um ano como 250 dias de negociação e um período de 4 meses para *trading*. O período amostral vai de jan-2005 a dez-2009. Assim, temos 12 períodos de *trading* não sobrepostos, perdendo 250 dias para o primeiro período de estimação.

#### 4.3. Resultados fora da amostra

O primeiro ano<sup>8</sup> da amostra é usado como período de formação, com esta base de dados são definidos os primeiros pares. Uma vez que são empregadas as séries de preços das 50 ações com maior liquidez, existem 1225 pares possíveis. São realizados testes de cointegração em todos os pares, posteriormente calcula-se o *SR* em simulações dentro da amostra e meia-vida de todos os pares cointegrados. Dos 1250 pares possíveis, em média foram obtidos 91 pares cointegrados em cada período. Em seguida os pares que passaram nos testes de cointegração são ranqueados com base no *SR* obtido nas simulações dentro da amostra. Seguindo Gatev et alii (2006) e Andrade et alii (2005), a estratégia é analisada com base numa carteira composta por 20 pares, sendo estes pares definidos como aqueles que apresentaram maior *SR* para o período dentro da amostra. Após o período de formação os mesmos são usados por quatro meses nas operações fora da amostra. No final de cada período de *trading* as operações que estiverem abertas são encerradas, todos os parâmetros são reestimados e pares podem ser substituídos.

Abrindo uma posição no par um dia após a divergência, seja a coordenada (*l*) referente a ação comprada e (*s*) referente a ação vendida. Seja  $P_t^l$  preço da ação *l* no dia *t* e  $P_t^s$  o preço das ação *s* em *t*. Assim, o retorno do par *i* no dia *t* é definido como:

<sup>8</sup> Ao longo deste artigo, definimos um ano como 250 dias de negociação e quatro meses como 84 dias de negociação.

$$r_{i,t}^{raw} = \ln \left[ \frac{P_t^l}{P_{t-1}^l} \right] - \ln \left[ \frac{P_t^s}{P_{t-1}^s} \right] \tag{5}$$

O retorno líquido diário do portfólio com  $N$  pares no dia  $t$  é:

$$R_t^{net} = \sum_{i=1}^N w_{i,t} r_{i,t} + 2n \ln \left( \frac{1-C}{1+C} \right) \tag{6}$$

onde  $w_{j,t}$  é o peso do par  $j$  no portfólio no tempo  $t$ ,  $N$  é o número de pares considerados e  $C$  refere-se aos custos de transação. A segunda parte da equação (6) leva em conta os custos de transação.<sup>9</sup>

As carteiras de pares são avaliadas em termos de rentabilidade acumulada, variância dos retornos ( $\hat{\sigma}^2$ ), índice de Sharpe ( $SR$ ) e *Maximum drawdown* ( $MDD$ ). O drawdown é a medida do declínio de um pico histórico em alguma variável (geralmente o lucro acumulado ou rentabilidade de uma estratégia de investimento). Neste artigo,  $MDD$  é definido como queda máxima percentual incorrida desde um pico até um vale em determinado período de tempo. Estas estatísticas são calculadas como segue:

$$R^A = 252 * \frac{1}{T} \sum_{t=1}^T R_t$$

$$\hat{\sigma}^A = \sqrt{252} * \frac{1}{T} \sum_{t=1}^T (R_t - \hat{\mu})^2$$

$$SR = \frac{\hat{\mu}}{\hat{\sigma}}, \text{ onde } \hat{\mu} = \frac{1}{T} \sum_{t=1}^T w_{i,t} R_{i,t}$$

$$MDD = \sup_{t \in [0, T]} \left[ \sup_{s \in [0, t]} R_s - R_t \right],$$

onde  $w_{j,t}$  é o peso do par  $j$  na carteira no tempo  $t$ . Por construção, o *maximum drawdown* é maior, em valor absoluto, que a perda máxima. Realmente, trata-se do cenário mais pessimista.

<sup>9</sup> Esta fórmula que leva em conta diretamente os custos de transação pode ser explicada de uma forma bastante intuitiva. Suponha que se compre uma ação  $\xi$  ao preço  $P_{t-1}^\xi$  no tempo  $t-1$  e venda-a um dia depois, no tempo  $t$ , por  $P_t^\xi$  unidades monetárias (preço de mercado), de fato os custos de comprar a ação são  $P_{t-1}^\xi(1+C)$  e o lucro obtido com a venda  $P_t^\xi(1-C)$ . Isto corresponde ao retorno líquido decomposto:

$$\ln \left[ \frac{P_t^\xi(1-C)}{P_{t-1}^\xi(1+C)} \right] = \ln \left[ \frac{P_t^\xi}{P_{t-1}^\xi} \right] + \ln \left[ \frac{(1-C)}{(1+C)} \right] = r_t^\xi + \ln \left[ \frac{(1-C)}{(1+C)} \right]$$

onde o segundo termo leva em conta os custos de transação da compra e da venda.

A Tabela 1 apresenta os pares empregados na estratégia no período referente ao último quadrimestre da amostra, setembro a dezembro de 2009. De 1250 pares possíveis para o período, 97 passaram nos testes de cointegração de Johansen e Engle-Granger, dos quais foram selecionados os 20 pares que apresentaram maior *SR* nas simulações dentro da amostra para serem empregados no período fora da amostra. Apesar de não haver restrição a que os ativos de um par se refiram a empresas de um mesmo setor, nota-se que a maioria dos pares são formados por ações de empresas com algum tipo de relação, sejam do mesmo ramo de atividade ou participantes de um mesmo segmento do mercado. Nota-se também que muitos dos pares apresentam meia vida inferior a 10 dias, reforçando a característica de reversão à média, o que é desejável para a estratégia. Todos os pares passam nos testes de cointegração de Johansen ( $\lambda_{\text{traço}}$ ) e Engle-Granger (*ADF*) a 95% (os valores críticos a 95% de confiança são 13.43 e 3.38, respectivamente). Nota-se que todos os 20 pares apresentam *SR* positivo nas simulações dentro da amostra, mesmo assim, nem todos obtiveram rentabilidade positiva no período fora da amostra. No período em questão, 6 dos 20 pares que compunham a carteira apresentaram resultados negativos, em média foi obtida rentabilidade líquida de 3.82% por par no período.

A Figura 2 ilustra o comportamento de um par<sup>10</sup> específico e do respectivo *spread*. O par é formado pelas ações *VALE5* e *BRAP4*, que formam um dos pares utilizados no período de setembro a dezembro de 2009. Também é mostrada a série do *spread* padronizado, que define os momentos de abrir e encerrar as operações. Conformepode ser observado na Tabela 1, este par apresenta meia vida de aproximadamente 12 dias e acumulou retorno líquido de 7.96% no período fora da amostra analisado.

A Tabela 3 resume algumas estatísticas relativas aos resultados da estratégia *pairs trading* implementada. Os resultados apresentados referem-se a análise fora da amostra para o período de janeiro de 2006 a dezembro de 2009. As rentabilidades apresentadas já estão descontadas dos custos de transação.<sup>11</sup> Em média foram obtidos 91 pares cointegrados por período, porém os resultados apresentados referem-se a uma carteira composta com pesos iguais, pelos 20 pares selecionados. A rentabilidade acumulada no período é de 89.78%, com *SR* anualizado de 1.29. Nota-se também que a estratégia apresenta volatilidade relativamente baixa, 12.49% em termos anualizados, e coeficiente de correlação com o mercado de 0.079, indicando que a estratégia pode ser considerada neutra ao mercado. Optou-se por empregar o coeficiente de correlação de Spearman<sup>12</sup> levando-se em conta que é uma medida não-paramétrica que não pressupõe que as séries de dados provenham de uma distribuição normal bivariada. Isso se justifica uma vez que estratégias com

<sup>10</sup> O par *VALE5* e *BRAP4* foi escolhido apenas para ilustrar a forma de implementação da estratégia.

<sup>11</sup> Os custos considerados são de 0.5% na entrada e 0.5% no encerramento da posição, totalizando 0.10% por operação.

<sup>12</sup> O coeficiente de correlação de Spearman é uma medida não paramétrica, não sensível a assimetrias nas distribuições nem à presença de *outliers*, não exigindo que os dados provenham de duas populações normais. O coeficiente de correlação de Spearman é definido por:

Tabela 1

Estatísticas dos pares utilizados no último quadrimestre (setembro a dezembro de 2009)

Ativo A	Ativo B	EG (ADF)	JH ( $\lambda_{trao}$ )	IR ( <i>in-sample</i> )	Meia Vida	Rent. Liq.
ITUB4	ITSA4	-3.67	18.32	4.33	4.58	6.60
USIM5	USIM3	-3.47	18.00	3.29	16.82	9.03
VALE3	BRAP4	-4.25	24.23	3.13	12.19	7.96
AMBV4	NATU3	-4.27	19.42	2.69	6.00	10.19
AMBV4	JBSS3	-4.35	20.63	2.57	9.01	4.00
CSNA3	BRAP4	-3.78	19.27	2.39	11.62	10.10
BBAS3	LREN3	-3.47	16.86	1.90	6.31	-19.68
CYRE3	GFSA3	-4.17	18.20	1.70	4.81	24.19
VALE3	CCRO3	-3.39	14.24	1.61	10.87	-4.00
BBAS3	USIM3	-3.69	19.58	1.60	8.61	-10.69
BRFS3	PCAR5	-4.15	19.20	1.59	6.41	-6.59
NETC4	JBSS3	-3.87	17.23	1.58	8.39	-8.28
CPLE6	PCAR5	-4.05	17.26	1.57	12.22	13.78
CPLE6	CCRO3	-3.58	15.97	1.56	13.25	5.37
LAME4	AMBV4	-3.85	17.06	1.46	14.85	8.91
ITUB4	CCRO3	-3.92	17.60	1.45	6.18	13.45
USIM5	BBAS3	-4.39	22.13	1.40	6.51	0.40
AMBV4	CCRO3	-3.82	18.76	1.24	6.98	22.09
BRFS3	CPLE6	-3.55	19.38	1.15	7.04	-14.38
BVMF3	NETC4	-3.98	19.39	1.10	17.01	4.02

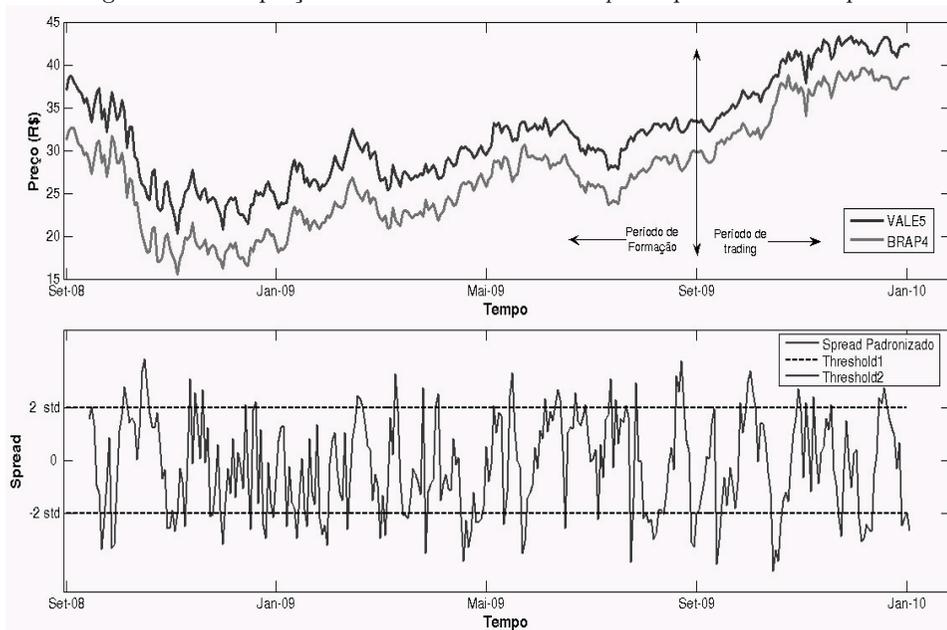
Nota: Rentabilidade líquida apurada para o período fora da amostra, expressa em %. Meia vida expressa em dias úteis.

ações que assumem posições compradas e vendidas geralmente apresentam retornos leptocúrticos.

É apresentado também o *maximum drawdown* da estratégia no período analisado, que foi de 24.49%. *Maximum drawdown* é a diferença entre o máximo global e o

$$\rho = \frac{\left(1 - 6 \sum_{i=1}^n d_i^2\right)}{n^3 - n}$$

onde  $d_i$  é a diferença entre os *ranks* dos correspondentes valores de  $X_i$  e  $Y_i$ , e  $n$  é o número de valores em cada conjunto.

Fig. 2. Séries de preços de VALE5 e BRAP4 e *spread* padronizado do par

Nota: Par empregado na estratégia no período setembro a dezembro de 2009.

mínimo global da curva de rentabilidade acumulada, (neste caso o mínimo global deve ocorrer depois do máximo global). Em outras palavras, é simplesmente uma medida da queda, em termos percentuais, em relação ao pico do retorno acumulado, pode ser usada como uma medida de quão agressivamente pode-se aumentar a alavancagem na estratégia. Pode-se observar na Figura 3, que traz a rentabilidade acumulada da estratégia e a volatilidade, que o *maximum drawdown* ocorreu no primeiro semestre de 2008, período em que o mercado financeiro nacional foi mais afetado pela crise financeira internacional, com os principais ativos negociados apresentado expressivas perdas e níveis de volatilidade nunca antes observados.

Nota-se na Figura 3, segundo painel, que a estratégia *pairs trading* implementada apresentou padrão de volatilidade relativamente baixo ao longo de praticamente todo o período analisado, ficando em patamares inferiores a 15%, em termos anualizados, em quase todo o período analisado. Mesmo no período mais agudo da crise financeira internacional, quando a volatilidade no mercado de ações doméstico ultrapassou 120%, a volatilidade da estratégia não chegou a 40%.

Parte relevante na avaliação de uma estratégia *pairs trading* é análise da correlação com o principal índice de referência do mercado, já que um dos objetivos da estratégia é a neutralidade ao mercado. Nesse sentido pode-se dizer que o objetivo foi alcançado, uma vez que a estratégia exibiu correlação com o Ibovespa inferior a 0.30 em todo o período analisado, sendo que na maior parte do tempo foi inferior a 0.10, como pode ser observado na Figura 4.

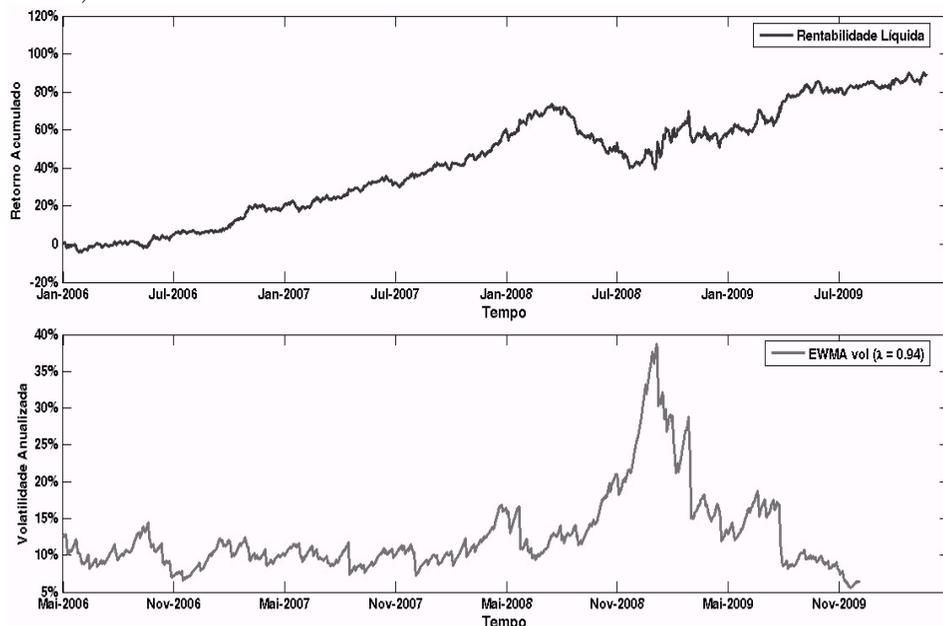
Tabela 2

Estatísticas da Estratégia *Pairs Trading* (Jan-2006 a Dez-2009)

<b>Descrição dos Dados</b>	
Total de dias na amostra	1250
Dias em cada período de formação	250
Dias em cada período de trading	84
Número de períodos de trading na amostra	12
Número de pares em cada período de trading	20
<b>Resumo da Rentabilidade da Estratégia <i>Pairs Trading</i></b>	
Retorno médio diário da carteira	0.069%
Maior rentabilidade em um dia	6.25%
Menor rentabilidade em um dia	-4.80%
Desvio padrão do retorno diário	0.85%
Índice de Sharpe dos retornos diários	0.081
Rentabilidade acumulada no período	89.78%
Rentabilidade média anual da carteira	17.37%
Desvio padrão do retorno (anualizado)	12.49%
Índice de Sharpe dos retornos (anualizado)	1.29
Coefficiente de correlação (Spearman)	0.079
Assimetria	0.269
Curtose	4.546
Máximo <i>Drawdown</i>	24.49%

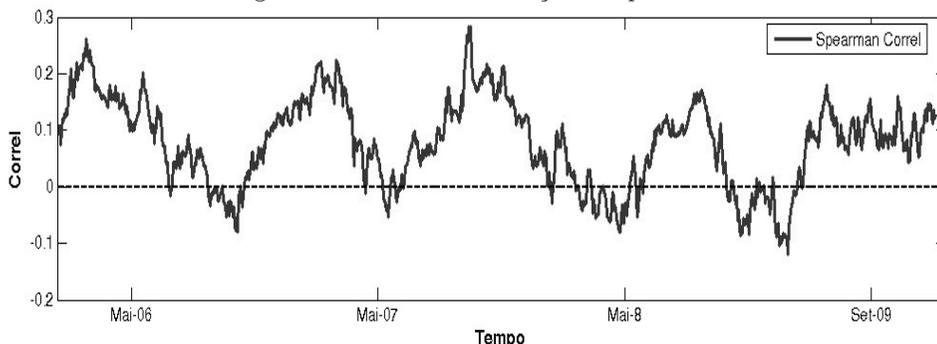
Nota: Esta tabela apresenta estatísticas descritivas da estratégia de arbitragem estatística, *pairs trading*, que implementamos. O período amostral vai de janeiro de 2005 a dezembro de 2009.

A Tabela 3 resume estatísticas anuais da estratégia *pairs trading*. Pode-se observar que o ano de 2008 foi quando a estratégia apresentou pior performance, acumulando rentabilidade líquida de 2.85% e volatilidade levemente superior aos demais anos. Conforme destacado em Khandani e Lo (2007) e Avellaneda e Liley (2010), o segundo semestre de 2007 e primeiro semestre de 2008 foram bastante complicados para fundos de investimentos quantitativos. Particularmente para estratégias de arbitragem estatística, que experimentaram expressivas perdas no

Fig. 3. Resultados Acumulados da Estratégia *Pairs Trading* e Volatilidade (Jan-06 a Dez-09)

Nota: No primeiro painel rentabilidade acumulada da estratégia *pairs trading* e no segundo volatilidade anualizada (EWMA Vol com  $\lambda = 0.94$ ).

Fig. 4. Coeficiente de correlação de Spearman



Nota: Coeficiente de correlação de Spearman calculado com base numa janela móvel de 84 observações.

período, com posterior recuperação em alguns casos. Muitos gestores sofreram perdas de tal forma que tiveram que desalavancar suas carteiras, não se beneficiando da subsequente recuperação. Os resultados encontrados neste estudo corroboram com aqueles encontrados por Avellaneda e Leey (2010) e Khandani e Lo (2007).

Nota-se na Figura 3 que a estratégia *pairs trading* proposta apresentou perda

expressiva no primeiro semestre de 2008, vindo a recuperar-se no segundo semestre. Khandani e Lo (2007) e Avellaneda e Leey (2010) sugerem que os eventos de 2007-2008 poderiam ser consequência de falta de liquidez, causada por fundos que tiveram que desfazer suas posições.

Tabela 3

**Estatísticas Descritivas Anuais da Estratégia Pairs Trading – 2006 a 2009**

Ano	Máx	Min	# Pares	Média	Desv. Pad.	Curtose	Assimetria	Rent. Liq.	Sharpe	MDD
2006	7.41	-3.11	48	1.44	0.12	0.04	5.57	17.37	1.72	5.57
2007	4.32	-1.48	97	2.13	0.12	0.23	4.71	24.94	2.62	4.71
2008	12.41	-6.19	59	0.29	0.14	0.68	6.98	2.85	0.15	24.98
2009	3.89	-1.76	155	1.77	0.04	0.34	5.24	21.25	1.86	5.24
Total	12.41%	-6.19%	359	1.19%	0.124	0.51	6.21	89,78%	1.29	24.98

Nota: Estatísticas descritivas para excesso de retorno percentual anual (líquido) do portfólio com os 20 pares selecionados para o período entre 2006 e 2009.

Estratégias desta natureza resultam em carteiras alavancadas, que assumem centenas de posições compradas e vendidas em ações. Mesmo cada posição sendo relativamente pequena em relação ao mercado como um todo, o efeito agregado das várias posições podem potencializar as perdas. Deve-se considerar também que estratégias de arbitragem estatística, como *pairs trading*, baseadas em cointegração podem sofrer em momentos de quebras estruturais nos mercados, como pode ser o caso do período em questão. Nestes casos a alternativa seria o uso de metodologias que permitam identificar eventuais quebras, por exemplo, estimar os parâmetros através do filtro de Kalman.

## 5. Conclusão

Neste artigo propõe-se a implementação de uma estratégia de arbitragem estatística conhecida como *pairs trading* aplicada às ações constituintes do índice Ibovespa. A estratégia é desenvolvida com base em cointegração, explorando a característica de reversão à média dos pares de ações cointegrados. Assim, o primeiro objetivo é desenvolver um método para identificar pares que possam ser usados na estratégia *pairs trading*. Após a identificação dos pares é conduzida uma análise dos co-movimentos entre as ações buscando identificar a força de reversão à média através da estimação da meia-vida, a qual é relevante para aplicação da estratégia.

Empregando as séries de preços das ações constituintes do Ibovespa no último quadrimestre de cada período de estimação, primeiramente são realizados testes de cointegração de Johansen e Engle-Granger em todas as combinações de pares possíveis com o objetivo de identificar pares de ações que compartilham uma relação de equilíbrio de longo prazo. Posteriormente, para os pares que passam nos testes de

cointegração, é calculada a meia-vida, através da equação de Ornstein-Uhlenbeck. Subsequentemente, calcula-se o *spread* padronizado entre as ações e simula-se *tradings* no período dentro da amostra com base em três regras simples. Uma posição é aberta sempre que o *spread* se distancia por mais que dois desvios da sua média de longo prazo. Todas as posições são liquidadas quando o *spread* retorna para sua média de longo prazo,<sup>13</sup> ou a posição está aberta por um período igual à meia vida do *spread*. A partir disso, é montada uma carteira composta pelos 20 pares que apresentaram maior *SR* para o período simulado dentro da amostra.

De 1250 pares possíveis, em média foram obtidos 91 pares cointegrados em cada período de formação. A carteira diversificada, composta pelos 20 pares que apresentaram os melhores indicadores dentro da amostra (com pesos iguais), alcançou índice de Sharpe de 1.29 no período de *trading* fora da amostra. A rentabilidade líquida acumulada no período de testes fora da amostra de quatro anos foi de 89.79%, média de 17.37% anuais. Além disso, a estratégia *pairs trading* aqui implementada apresentou níveis de volatilidade relativamente baixos e correlação com o índice Ibovespa inferior a 0.10, em termos absolutos, na maior parte do período analisado, corroborando com o objetivo de neutralidade ao mercado da estratégia. Os resultados são atrativos quando comparados a outras estratégias empregadas por *hedge funds*, principalmente se for levado em conta que a estratégia praticamente não envolve desembolso de recursos, resumindo-se basicamente às margens de garantia, permitindo maior nível de alavancagem e a rentabilização dos recursos envolvidos em ativos livres de risco. Existem algumas direções a serem seguidas em trabalhos futuros com o objetivo de potencializar a rentabilidade e mitigar riscos, entre elas, o emprego de algum método para identificar e monitorar a estabilidade dos parâmetros, por exemplo, estimação através do filtro de Kalman. Outra alternativa é aplicar a metodologia proposta a dados de alta frequência. Portanto, os resultados encontrados reforçam o uso do conceito de cointegração como importante ferramenta na gestão quantitativa de fundos.

Além dos resultados da estratégia serem atrativos no contexto de estratégias de investimentos, a estratégia de arbitragem estatística implementada pode ser vista como um teste de eficiência de mercado para o caso brasileiro, seguindo o argumento apresentado por Hogan et alii (2004), que introduz o conceito de arbitragem estatística como forma de testar eficiência de mercado sem a necessidade de especificar um modelo de equilíbrio. Segundo os autores, no limite, arbitragem estatística converge para arbitragem no sentido amplo, conseqüentemente, o dilema do conjunto de hipóteses é evitado, apelando para uma estratégia de investimento de longo prazo.

---

<sup>13</sup> Define-se retornar para a média de longo prazo quando a distância do *spread* em relação à média de longo prazo é menor que 0.5 desvios.

## Referências bibliográficas

- Aldrichi, D. M. & Milanez, D. Y. (2005). Finança comportamental e a hipótese dos mercados eficientes. *Revista de Economia Contemporânea*, 9:41–72.
- Alexander, C. & Dimitriu, A. (2005a). Indexing and statistical arbitrage: Tracking error or cointegration? *Journal of Portfolio Management*, 31:50–63.
- Alexander, C. & Dimitriu, A. (2005b). Indexing, cointegration and equity market regimes. *International Journal of Finance and Economics*, 10:213–231.
- Andrade, S. C., Pietro, V., & Seasholes, M. S. (2005). Understanding the profitability of pairs trading. UC Berkeley Working Papers, UC Berkeley.
- Avellaneda, M. & Leey, J.-H. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, 10:1–22.
- Bessler, D. A. & Yang, J. (2003). The structure of interdependence in international stock markets. *Journal of International Money and Finance*, 22(2):261–287.
- Black, F. (1972). Capital market equilibrium and restricted borrowing. *Journal of Business*, 45:444–454.
- Burges, A. N. (2003). Using cointegration to hedge and trade international equities. In *Applied Quantitative Methods for Trading and Investment*, pages 41–69. John Wiley & Sons.
- Caldeira, J. F. & Portugal, M. S. (2010). Estratégia long-short, neutra ao mercado, e index tracking baseadas em portfólios cointegrados. *Revista Brasileira de Finanças*, 8(4):469–504.
- Capocci, D. P. (2006). The neutrality of market neutral funds. *Global Finance Journal*, 17:309–333.
- Carcano, G., Falbo, P., & Stefani, S. (2005). Speculative trading in mean reverting markets. *European Journal of Operational Research*, 163(1):132–144.
- Chan, E. P. (2009). *Quantitative Trading: How to Build Your Own Algorithmic Trading Business*. Wiley trading series, New Jersey.
- Chan, L. K. C., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *Journal of Finance*, 51(6):1681–1713.
- Do, B., R. F. & Hamza, K. (2006). A new approach to modeling and estimation for pairs trading. Working paper series, Monash University.
- Dunis, C. L., Giorgioni, G., Laws, J., & Rudy, J. (2010). Statistical Arbitrage and High-Frequency Data with an Application to Eurostoxx 50 Equities. Cibef working papers, CIBEF.
- Dunis, C. L. & Ho, R. (2005). Cointegration portfolios of European equities for index tracking and market neutral strategies. *Journal of Asset Management*, 1:33–52.
- Elliott, R., der Hoe, J. V., & Malcolm, W. (2005). Pairs trading. *Quantitative Finance*, 5:271–276.
- Enders, W. (2002). *Applied Econometric Times Series*. John Wiley & Sons, New York.
- Enders, W. & Granger, C. W. J. (1998). Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. *Journal of Business & Economic Statistics*, 16:304–311.
- Engle, R. & Granger, C. (1987). Co-integration and error correction: Representation, estimation and testing. *Econometrica*, 2:251–76.
- Engle, R. & Yoo, B. (1987). Forecasting and testing in co-integrated systems. *Journal of Econometrics*, 35:143–159.

- Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, 49(3):283–306.
- Fund, W. & Hsieh, D. (1999). A primer of hedge funds. *Journal of Empirical Finance*, 6:309–331.
- Galenko, A. & Popova, E. (2012). Trading in the presence of cointegration. *The Journal of Alternative Investments*, 15(1):85–97.
- Gatarek, L. T., Hoogerheide, L., van Dijk, H. K., & Verbeek, M. (2011). A simulation-based Bayes' procedure for robust prediction of pairs trading strategies. Tinbergen Institute Discussion Papers 09-061/4, Tinbergen Institute.
- Gatev, E., Goetzmann, G. W., & Rouwenhorst, K. (2006). Pairs trading: Performance of a relative value arbitrage rule. *The Review of Financial Studies*, 19:797–827.
- Gillespie, D. T. (1996). Exact Numerical Simulation of the Ornstein-Uhlenbeck Process and Its Integral. *Physical Review E*, 54:2084–91.
- Hamilton, J. D. (1994). *Times Series Analysis*. IE-Princeton, New Jersey.
- Hendry, D. F. & Juselius, K. (2000). Explaining cointegration analysis: Part 1. *The Energy Journal*, 0(Number 1).
- Hendry, D. F. & Juselius, K. (2001). Explaining cointegration analysis: Part ii. *The Energy Journal*, 0(Number 1).
- Herlemont, D. (2000). Pairs trading, convergence trading, cointegration. Yats working papers, YATS Finances & Technologies.
- Hogan, S., Jarrow, R., Teo, M., & Warachka, M. (2004). Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial Economics*, 73(3):525–565.
- Jacobs, B., Levy, K., & Starer, D. (1993). Long-short equity investing. *Journal of Portfolio Management*, 1:52–64.
- Jarrow, R., Teo, M., Tse, Y. K., & Warachka, M. (2012). An improved test for statistical arbitrage. *Journal of Financial Markets*, 15(1):47–80.
- Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, 48:65–91.
- Juselius, K. (2006). *The cointegrated VAR model: Methodology and Applications*. Oxford University Press, Oxford.
- Khandani, A. E. & Lo, A. W. (2007). What happened to the quants in August 2007? *Journal of Investment Management*, 5:5–54.
- Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance*, 49(5):1541–78.
- Lin, Y., McRae, M., & Gulati, C. (2006). Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Journal of Applied Mathematics and Decision Sciences*, 6:1–14.
- Lintner, J. (1965). The valuation of risk assets and the selection of risk investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 4:13–37.
- Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. John Wiley & Sons, New York.
- Mudchanatongsuk, S., Primbs, J. A., & Wong, W. (2008). Optimal pairs trading: A stochastic control approach. In *Proceedings of American Control Conference*. American Control Conference.
- Nath, P. (2006). High frequency pairs trading with u.s treasury securities: Risks and rewards for hedge funds. Working paper series, London Business School.

- Serban, A. F. (2010). Combining mean reversion and momentum trading strategies in foreign exchange markets. *Journal of Banking & Finance*, 34(11):2720–2727.
- Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442.
- Stock, J. & Watson, M. (1988). Testing for common trends. *Journal of the American Statistical Association*, 83:1097–1107.
- Triantafyllopoulos, K. & Montana, G. (2011). Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science*, 8(1):23–49.
- Uhlenbeck, G. & Ornstein, L. (1930). On the theory of Brownian motion. *Physical Review*, 36:823–841.
- Vidyamurthy, G. (2004). *Pairs Trading, Quantitative Methods and Analysis*. John Wiley & Sons, Canada.
- Yang, J., Kolari, J. W., & Sutanto, P. W. (2004). On the stability of long-run relationships between emerging and us stock markets. *Journal of Multinational Financial Management*, 14(3):233–248.