



UNIVERSIDADE CATÓLICA PORTUGUESA

Pairs Trading

Aplicação da Distância Mínima

Trabalho Final na modalidade de Dissertação
apresentado à Universidade Católica Portuguesa
para obtenção do grau de mestre em Finanças

por

Luís Filipe Carreira de Campos Figueiredo

sob orientação de

Prof. Doutor Ricardo Cunha

Prof. Doutor Paulo Alves

Católica Porto Business School,
Universidade Católica Portuguesa

Janeiro de 2017

Agradecimentos

Gostaria de agradecer à minha mãe, ao meu pai, à minha irmã e à Cristina por todo o apoio que me deram neste meu percurso académico, pois sem eles não teria chegado aqui.

Queria agradecer também ao Professor Ricardo Cunha e ao Professor Paulo Alves por me terem apoiado no desenvolvimento deste trabalho e me terem guiado na direção certa.

Agradeço especialmente ao Professor Ricardo Ribeiro, cujo apoio no Excel e no Matlab foi essencial e sem o qual não teria sido possível realizar este trabalho.

Também agradeço ao Doutor Nuno Henriques por todo o apoio que me deu na utilização do Datastream e por me ter facilitado o acesso à plataforma.

Um último, mas não menos importante, agradecimento ao meu amigo Stéphane Oliveira Leite por todo o apoio dado a nível de formatação e apresentação.

Resumo

Neste trabalho analisamos uma estratégia de arbitragem estatística conhecida como *Pairs Trading*, em que fazemos inicialmente um enquadramento teórico e uma revisão das principais estratégias e metodologias desenvolvidas. De seguida, replicamos a estratégia utilizada em Gatev, et al. (2006), com algumas diferenças significativas, utilizando as ações americanas cotadas no New York Stock Exchange, para um período compreendido entre janeiro de 2007 e junho de 2016.

O nosso trabalho destaca-se pelos seguintes motivos: em primeiro lugar, utilizamos restrições de liquidez, em cada período de simulação, com o objetivo de que os pares de ações encontrados apresentassem preços de fecho, volume de transações e dias de transações semelhantes.

Em segundo lugar, utilizamos cotações *bid/ask*, em vez dos preços diários de fecho, de forma a que a simulação fosse o mais aproximada possível de uma aplicação real da estratégia.

Em último lugar, definimos um valor fixo de 20 pares de ações, em cada período de simulação, de modo a que os nossos resultados e conclusões fossem os mais robustos e sólidos possíveis.

Obtivemos um retorno excessivo médio de 12,89% para períodos de simulação 6 meses, um retorno excessivo total de 244,93% em 9,5 anos e um retorno excessivo médio de 0,92% por par.

Palavras-chave: *Pairs Trading*; Distância mínima; Restrições de liquidez

Abstract

In this thesis, we study an arbitrage strategy known as Pairs Trading, in which we initially present a theoretical background and a revision of the main used strategies. Thereafter we replicated the strategy used in Gatev et al. (2006), with a few but significant differences, using American stocks listed at New York Stock Exchange, for a period between January 2007 and June 2016.

Our work stands out for the following reasons: in the first place, we used liquidity restrictions, in each simulation period, so that the selected pairs of stocks had similar close prices, trading volume and transaction days.

Secondly, we used bid/ask quotes instead of the daily close prices, so that the strategy simulation could be as similar as possible to a real application of the strategy.

Lastly, we chose to select 20 pairs of stocks, per simulation period, so that our results and conclusions could be as sound and solid as possible.

We obtained an average excess return of 12.89% for 6 months' simulation periods, a total excess return of 244.93% in 9.5 years and an average excess return of 0.92% per pair.

Keywords: Pairs Trading, Minimum distance; Liquidity restrictions

Índice

Agradecimentos	iii
Resumo	v
Abstract	vii
Introdução.....	15
Capítulo 1	
Revisão Bibliográfica	19
1.1 Eficiência de Mercado	19
1.2 Risco de Mercado.....	22
1.3 Revisão de estratégias desenvolvidas.....	27
1.3.1 Outras ferramentas: A questão da cointegração	30
1.3.2 Outros mercados: A questão dos setores económicos.....	31
Capítulo 2	
Implementação da Estratégia.....	33
2.1 Dados.....	33
2.2 Formação dos pares.....	34
2.3 Cálculo da distância mínima.....	36
2.4 Sinais de entrada e saída de mercado.....	38
2.5 Cálculo dos excessos de retorno	40
Capítulo 3	
Resultados.....	41
3.1 Resultados.....	41
3.2 Resultados detalhados da estratégia.....	43
3.3 Neutralidade da estratégia face ao mercado	47
Capítulo 4	
Conclusões	49
4.1 Conclusões	49
Bibliografia.....	51

Índice de Figuras

Figura 1: Exemplo da estrutura dos períodos de formação e simulação	34
Figura 2: Evolução dos retornos normalizados do par de ações WRI-PLD.....	38
Figura 3: Exemplo do período de simulação do par de ações WRI-PLD	39
Figura 4: Evolução da rentabilidade da estratégia e do NYSE	48

Índice de Tabelas

Tabela 1: Base de dados das empresas por setor económico.....	34
Tabela 2: Resultados das estratégias de Gatev et al. (2006) e da nossa	42
Tabela 3: Resultados detalhados da estratégia	44
Tabela 4: Resultados detalhados dos pares consoante o dia de fecho	46

Introdução

As estratégias de arbitragem estatística, como a metodologia *Pairs Trading*, são estratégias de investimento quantitativas desenhadas para gerar retornos económicos quando ocorrem ineficiências de mercado. São estratégias que requerem uma análise prévia da evolução dos preços, de forma a poderem definir-se estratégias de negociação que serão, posteriormente, automatizadas no mercado através de sistemas informáticos avançados.

A metodologia *Pairs Trading* foi introduzida nos mercados financeiros americanos, nos anos 80, por Nunzio Tartaglia e pela sua equipa de físicos, matemáticos e informáticos, do grupo Morgan Stanley, que tinham como objetivo descobrir oportunidades de arbitragem nos mercados financeiros. Para tal, a equipa utilizou métodos estatísticos sofisticados para desenvolver programas de *trading* avançados, que permitissem automatizar as entradas e saídas de mercado, de forma a usufruírem disciplinadamente das ineficiências de mercado que ocorriam.

Neste processo, a equipa de Nunzio Tartaglia identificou pares de ações cujos preços se moviam tendencialmente em conjunto, ou seja, apesar de por vezes os preços das ações se distanciarem, tendencialmente, eles voltam a aproximar-se.

A equipa de Nunzio Tartaglia definiu então uma estratégia de negociação, que viria a ficar conhecida como *Pairs Trading*, com o objetivo de obter retornos através destes períodos de divergência e posterior convergência entre os preços das ações que, tendencialmente, se moviam em conjunto.

A estratégia desenvolve-se da seguinte forma: quando a divergência entre os preços das duas ações for superior a um determinado valor, previamente definido, abre-se uma posição longa na ação que apresentar o preço relativamente mais baixo e uma posição curta na ação que apresentar o preço relativamente mais elevado. Desta forma, quando os preços das duas ações convergirem, fecham-se as posições abertas, obtendo-se um retorno ou com a subida do preço da ação na qual se abriu uma posição longa, ou com a descida do preço da ação na qual se abriu uma posição curta, ou com ambas.

Aplicando esta estratégia, a equipa de Nunzio Tartaglia gerou um lucro de \$50 milhões de dólares americanos para o grupo Morgan Stanley. Mais tarde, quando a equipa se desmembrou, a estratégia difundiu-se e ganhou popularidade, tornando-se uma estratégia comum de investimento.

Although the Morgan Stanley group disbanded in 1989 after a couple of bad years of performance, pairs trading has since become an increasingly popular “market neutral” investment strategy used by individual and institutional traders as well as hedge funds. - (Gatev et al. 2006, p. 799)

Neste trabalho, numa fase inicial, iremos explorar a forma como esta estratégia se enquadra nas hipóteses de eficiência do mercado e na Lei do Preço Único, iremos aprofundar uma das suas principais características no âmbito do *Capital Asset Pricing Model* (Markowitz, 1952; Sharpe, 1964; Lintner, 1965; Mossin, 1966), a neutralidade relativamente ao mercado, e iremos analisar as limitações e os resultados obtidos em alguns trabalhos de investigação desenvolvidos em torno desta estratégia.

Numa segunda fase, iremos replicar uma estratégia *Pairs Trading* desenhada por Gatev et al. (2006) em que, através da distância entre os preços normalizados e através das restrições de liquidez aplicadas, definimos o top 20 de pares de ações para 19 períodos de simulação, com a duração de 6 meses cada.

Na terceira e última fase do trabalho, iremos analisar detalhadamente os resultados obtidos na nossa estratégia, fazendo as devidas comparações com os

resultados de Gatev et al. (2006), bem como apresentar as limitações e conclusões encontradas.

Capítulo 1

Revisão Bibliográfica

1.1 Eficiência de Mercado

Um mercado eficiente caracteriza-se por apresentar ativos cujos preços refletem com exatidão os sinais corretos à alocação de recursos, ou seja, um mercado onde quer as empresas quer os investidores têm acesso a informação verdadeira e transparente que lhes possibilite tomar decisões de produção/investimento e de distribuição de poupanças, respetivamente. “A market in which prices always ‘fully reflect’ available information is called ‘efficient’.” (Fama, 1970, p. 383)

Esta definição, no entanto, é muito ampla e houve necessidade de testar o conceito de informação disponível em vários níveis, na sua versão fraca, semi-forte e forte, de modo a conseguir-se testar empiricamente a eficiência do mercado.

Durante o século 20, foi primeiramente desenvolvida a teoria *Random Walk* (Bachelier, 1900) que defendia duas hipóteses: as sucessivas alterações nos preços dos ativos financeiros são independentes e estas apresentam distribuições idênticas. Isto é, a teoria *Random Walk* defende que não existe dependência linear significativa numa série temporal de preços de um ativo, o que significa que não é possível utilizar os dados históricos para prever os preços futuros de um ativo.

Esta hipótese (*Random Walk*) apenas pretende explicar a eficiência de mercado na sua versão fraca, ou seja, apenas quer saber se a informação histórica disponível se encontra refletida nos preços. Mais tarde, os economistas viriam a

preocupar-se com as restantes versões, a semi-forte que testa a velocidade de ajustamento dos preços e a versão forte que testa se os investidores têm acesso monopolístico a informações.

O laureado Nobel Eugene F. Fama, na década de 60, considerava que a hipótese *Random Walk* representava a realidade de forma adequada (Fama, 1965, p. 35). Mais tarde, devido aos contributos de Samuelson (1965) e de Mandelbrot (1966), Fama viria a admitir que estava errado e que os preços seguem na realidade um processo *sub-martingale*, que os modelos *Fair Game* dos retornos esperados desempenham um papel importante na eficiência de mercado e que a teoria *Random Walk* apenas se verifica nalgumas situações (Fama, 1970, p. 391).

Fama (1970) apresenta então a sua Hipótese da Eficiência de Mercado, desta vez fundamentada nos modelos *Fair Game* dos retornos esperados, mas com uma característica comum à hipótese *Random Walk*: não existe dependência linear estatisticamente significativa entre os retornos e os seus desfasamentos (Fama, 1970, p. 394).

No entanto, por esta altura, Fama já tinha observado situações de dependência linear estatisticamente significativa em preços diários quando desenvolveu um trabalho em parceria com Marshall Blume, em 1966, em que utilizaram testes de filtros muito pequenos que indicaram que era possível criar sistemas de trading, de muito curto-prazo, que superam em média o tradicional *buy-and-hold* por uma grande margem (Fama, 1970, p. 395).

Não obstante, Fama (1970) considera que tal estratégia de filtros não é lucrativa quando se tem em consideração os custos de transação e que a dependência linear encontrada nos movimentos de muito curto-prazo não é suficiente para rejeitar os modelos *Fair Game*.

Em suma, os modelos *Fair Game* dos retornos esperados parecem representar a realidade de forma adequada, pelo menos na sua forma fraca e semi-forte, como

mostrou Fama, no seu trabalho de 1970, onde a sua Hipótese de Eficiência de Mercado foi reconhecida com o Prémio Nobel de Economia, em 2013.

No entanto, a possibilidade de o mercado financeiro não ser eficiente nunca deixou de existir, ou não tivesse o Prémio Nobel de Economia de 2013 sido partilhado por três economistas, Eugene F. Fama, Robert Shiller e Lars Peter Hansen, sendo que Fama e Shiller têm visões contrárias precisamente sobre este tema.

A atribuição do Prémio Nobel de Economia de 2013 aos economistas deveu-se às suas análises empíricas sobre os preços dos ativos.

Eugene Fama demonstrou que os movimentos dos preços das ações são impossíveis de prever no curto-prazo e que a nova informação afeta os preços quase imediatamente, o que implica que o mercado é eficiente. Os principais trabalhos, de Fama, são, “The Behavior of Stock Market Prices” (1965), “The Adjustment of Stock Prices to New Information” (1969), com os coautores, Lawrence Fisher, Michael C. Jensen e Richard Roll, e o “Efficient Capital Markets: A Review of Theory and Empirical Work” (1970).

Por outro lado, Robert Shiller descobriu que os preços das ações podem ser previstos no longo-prazo, argumentando, portanto, que o mercado é ineficiente. Os seus trabalhos de referência são, “Do stock prices move too much to be justified by subsequent changes in dividends?” (1981) em que desafia a hipótese da eficiência de mercado e o livro *Irrational Exuberance* (2000).

Lars Peter Hansen desenvolveu um método estatístico que permitiu demonstrar que os resultados de Shiller não podem ser inteiramente explicados apenas pelos modelos existentes na altura. Atualmente o método de Hansen é utilizado em toda a investigação económica. Contribuiu também na coedição dos livros *Advances in Economics and Econometrics* (2003) e *Handbook of Financial Econometrics* (2009).

A par da hipótese da eficiência de mercado há que analisar um outro conceito económico importante, a Lei do Preço Único, que de forma geral nos diz que, “two

investments with the same payoff in every state of nature must have the same current value”, Ingersol (1987, p. 33). Aplicando este conceito aos mercados financeiros, isto implica que uma ação deve ter o mesmo preço em diferentes mercados, tendo já em consideração as taxas de câmbio e os custos de transação, porque, de outra forma, surgiriam oportunidades de arbitragem e o mercado não seria eficiente.

Gatev et al. (2006, p. 800-801) argumentam que, na ótica da Lei do Preço Único, se dois instrumentos financeiros têm o mesmo valor e os mesmos fatores de risco, então o preço relativo dos dois instrumentos deve resultar numa relação estável, caso contrário surgem oportunidades de arbitragem. De forma sucinta, isto significa que se dois ativos são substitutos, então devem ser vendidos pelo mesmo preço.

Os investigadores estendem este raciocínio argumentando que, num mercado eficiente, se dois ativos apresentam preços históricos semelhantes, então os seus preços atuais devem ser também semelhantes. O que isto implica é que, quando os preços se distanciam um do outro, estamos perante uma ineficiência de mercado ou uma violação à Lei do Preço Único, o que se traduz numa oportunidade para o arbitragista de aplicar a estratégia *Pairs Trading*.

Desta forma, Gatev et al. (2006) consideram que os lucros obtidos através do *Pairs Trading* são uma compensação que os arbitragistas obtêm por fazer cumprir a Lei do Preço Único e “zelarem” pela eficiência do mercado. “If the U.S. equity market were efficient at all times, risk-adjusted returns from pairs trading should not be positive” (Gatev et al. 2006, p. 797).

1.2 Risco de Mercado

As estratégias *Pairs Trading* são consideradas estratégias neutras face ao mercado. Para que possamos entender este conceito há que explicar brevemente o

modelo *Capital Asset Pricing Model* (CAPM), um modelo que nos permite prever a relação que devemos observar entre o risco de um ativo e o seu retorno esperado.

O modelo CAPM, considerado um dos principais modelos da teoria financeira moderna, foi desenvolvido em artigos, por Sharpe (1964), Lintner (1965) e Mossin (1966) dando continuação ao trabalho de Markowitz (1952).

A ideia geral do modelo CAPM é a de que os investidores devem ser compensados pelo valor temporal do investimento e pelo risco que incorrem. Enquanto que o valor temporal do investimento é compensado pela taxa de juro sem risco, que normalmente é definida por obrigações do tesouro, a compensação pelo risco do investimento é calculada de acordo com o *beta* de cada ativo, uma medida que reflete o risco de um determinado ativo, tendo em conta o retorno e a volatilidade esperada do mesmo face ao mercado.

Então, de forma a perceber qual a rentabilidade e a volatilidade esperada do mercado como um todo, utiliza-se um portefólio de mercado, isto é, um portefólio que funciona como um *proxy* uma vez que inclui todos os ativos disponíveis no mercado, de acordo com a sua capitalização.

De forma a relacionar a rentabilidade esperada do mercado com a rentabilidade esperada de um ativo (ativo p), definiu-se no modelo CAPM uma condição ótima, a *Securities Market Line* (SML), representada pela seguinte equação:

$$E[r_p] = r_f + \beta_p(E[r_M] - r_f)$$

Em que r_f representa a taxa de juro sem risco, $E[r_M]$ representa o retorno esperado do portefólio de mercado (*proxy*) e β_p representa o *beta*, ou seja, o risco do ativo em questão. Este último pode ser calculado da seguinte forma:

$$\beta_p = \frac{Cov(r_p, r_M)}{\sigma^2(r_M)}$$

Observa-se então que o beta do ativo p , (β_p), é o rácio entre a volatilidade do retorno do mesmo face ao risco de mercado e a volatilidade do mercado como um todo, calculadas respetivamente pela covariância entre o retorno do ativo e o retorno do portefólio de mercado, e pela variância do retorno do portefólio de mercado.

Introduzindo agora o conceito de risco, de acordo com Bodie, Kane & Marcus (2010, p. 280-281), existem dois tipos de risco, o risco de mercado que está relacionado com as condições gerais da economia e o risco específico de cada empresa.

O risco de mercado engloba variáveis que não podem ser previstas com um grau de confiança aceitável, como a inflação, taxas de juro, taxas de câmbio e ciclos económicos. Este tipo de risco, também denominado de risco sistémico ou não diversificável, não pode ser eliminado.

O risco específico de cada empresa, também conhecido por risco idiossincrático ou risco diversificável, pode ser eliminado através da diversificação eficiente de ativos num portefólio. Isto significa que um investidor pode eliminar o risco específico de cada empresa ao combinar, de forma ótima, vários ativos no mesmo portefólio. Independentemente de realizar esta operação ou não, o investidor fica sujeito ao risco de mercado.

Vidyamurthy (2004) considera que, na ótica do modelo CAPM, o retorno do ativo em questão pode ser separado em duas componentes, na componente de mercado ou sistémica e na componente residual ou não sistémica. Desta forma, Vidyamurthy (2004, p. 3) representa a SML da seguinte forma:

$$r_p = \beta r_M + \theta_p$$

Em que r_p representa o retorno de um ativo, βr_M representa a componente de mercado ou sistémica do retorno do ativo e θ_p representa a componente residual

ou não sistémica do retorno do ativo. Dos pressupostos do modelo CAPM, Vidyamurthy (2004) sublinha que as duas componentes de retorno são não correlacionadas e que o valor esperado da componente residual é zero.

Desenvolvendo o seu raciocínio, Vidyamurthy (2004) afirma o seguinte:

We remind ourselves of an oft-quoted adage about the markets, 'A rising tide raises all boats'. The statement indicates that when the market goes up, we can typically expect the price of all securities to go up with it. Thus, a positive return for the market usually implies a positive return for the asset, that is, the sum of the market component and the residual component is positive. – (Vidyamurthy, 2004, p. 3)

Vidyamurthy (2004) entende que, se a componente residual do retorno for nula (ou diminuta), como é esperado que seja, então o retorno positivo de um ativo é explicado (quase) inteiramente pela componente de mercado. Como tal, um retorno positivo no portefólio de mercado e, conseqüentemente, no ativo em questão, implica um valor positivo da componente de mercado, o que, por sua vez, se traduz num valor positivo para *beta*.

De seguida, Vidyamurthy (2004) introduz o conceito de estratégias neutras face ao mercado:

They are strategies that are neutral to market returns, that is, the return from the strategy is uncorrelated with the market return. Regardless of whether the market goes up or down, in good times and bad the market neutral strategy performs in a steady manner, and results are typically achieved with a lower volatility. This desired outcome is achieved by trading market neutral portfolios. – (Vidyamurthy, 2004, p. 5)

No contexto do modelo CAPM, entende-se que portefólios neutros face ao mercado são portefólios que apresentam um *beta* igual a zero. Isto porque, ao considerarmos na equação SML que o *beta* é zero, o retorno do portefólio passa a ser determinado apenas pela componente residual, $r_p = \theta_p$, o que significa que deixa de estar relacionado com o retorno do mercado, dado que as componentes residual e de mercado são não correlacionadas.

Vidyamurthy (2004) explica então que, ao construir um portefólio com apenas posições longas em ativos, é expectável que o *beta* do portefólio seja positivo, isto porque, retornos positivos no mercado resultam em retornos positivos para os ativos, logo um retorno positivo para o portefólio. De igual forma se entende que um portefólio, com apenas posições curtas, deverá ter um *beta* negativo. Então para se construir um portefólio neutro face ao mercado, ou seja, com *beta* igual a zero, há que combinar posições longas e curtas no mesmo portefólio.

Vidyamurthy (2004) vem demonstrar matematicamente este raciocínio. Considerando os retornos de dois portefólios, A e B, representados por r_A e r_B :

$$\begin{aligned} r_A &= \beta_A * r_M + \theta_A \\ r_B &= \beta_B * r_M + \theta_B \end{aligned}$$

Em que β_A e β_B representam *betas* positivos, θ_A e θ_B representam as componentes residuais dos retornos dos portefólios e r_M representa o retorno do portefólio de mercado.

Então, ao construir um portefólio AB em que se abre uma posição curtas em r unidades do portefólio A e uma posição longa numa unidade do portefólio B, o retorno de AB é dado por:

$$\begin{aligned} r_{AB} &= -r * r_A + r_B (=) \\ (=) r_{AB} &= (-r * \beta_A + \beta_B) * r_M + (-r * \theta_A + \theta_B) \end{aligned}$$

De onde se retira que o *beta* do portefólio AB é:

$$\beta_{AB} = -r * \beta_A + \beta_B$$

Se considerarmos que abrimos posições curtas em $r = \frac{\beta_B}{\beta_A}$ unidades do portefólio A, então o *beta* do portefólio AB será zero, o que se traduz num portefólio neutro face ao mercado.

Desta forma se compreende porque é que as estratégias *Pairs Trading* são estratégias neutras face ao mercado, uma vez que utilizam sempre uma posição longa e uma posição curta, simultaneamente, em cada par.

Na última fase do trabalho iremos tentar retirar conclusões sobre o desempenho da nossa estratégia *Pairs Trading* relativamente ao mercado.

1.3 Revisão de estratégias desenvolvidas

Os investigadores Gatev et al. (1999 e 2006) publicaram dois trabalhos de investigação sobre uma estratégia *Pairs Trading*, em que o trabalho Gatev et al. (1999) serviu para estudar o mercado e definir os parâmetros que viriam a ser utilizados na estratégia do trabalho Gatev et al. (2006).

Então, utilizando a estratégia original de 1999, os investigadores testaram a rentabilidade de uma estratégia *Pairs Trading* para o período de 1999 a 2002, obtendo em média um retorno excessivo anual de 10,4%, com um desvio-padrão anual de 3,8%, tendo obtido para os pares top 5 e top 20, em média, um retorno excessivo mensal de 1,31% e 1,44%, respetivamente.

Utilizando preços diários das ações americanas, retirados do CRSP, Gatev et al. (2006) separaram o estudo em períodos de análise de 12 meses (períodos de formação), em que selecionavam os pares e definiam parâmetros, e em períodos de teste de 6 meses (períodos de simulação) em que testavam os pares selecionados, imediatamente a seguir ao respetivo período de formação.

Outros autores optaram por diferentes horizontes temporais, como por exemplo Perlin (2009), que analisou o comportamento das ações durante 2 anos, para cada período de formação.

Utilizando os retornos (incluindo dividendos reinvestidos) das ações, normalizados, Gatev et al. (2006) calcularam a soma do quadrado dos desvios para todas as combinações possíveis de pares de ações, selecionando posteriormente,

para cada período, os pares top 5 e top 20 que apresentassem a mínima distância entre os retornos.

Gatev et al. (2006) argumentam que a ferramenta da distância mínima explora as relações estatísticas existentes entre os pares de ações, ao nível dos preços, tendo a vantagem de não ser necessário um modelo económico para a utilizar o que elimina o risco de erros de especificações e de estimações, mas, por outro lado, como nos diz Perlin (2009), a distância mínima não permite prever quando irá ocorrer a suposta convergência dos preços, dado que é uma ferramenta paramétrica.

Em Gatev et al. (2006) as entradas e saídas do mercado, em cada período de simulação, foram definidas da seguinte forma: é aberta uma posição longa-curta quando a distância entre os retornos normalizados dos pares ultrapassa duas vezes o desvio-padrão (dos retornos), calculado durante o respetivo período de formação, e fecha-se a posição (curta-longa) quando os retornos normalizados das ações se cruzam ou, em última instância, no último dia de cada período de simulação.

Apesar de Gatev et al. (2006) terem utilizado os retornos normalizados para parametrizar a estratégia em cada período de formação, quando testaram a rentabilidade dos pares, em períodos de simulação, fizeram-no utilizando preços diários de fecho. Ou seja, simularam entradas e saídas de mercado utilizando o mesmo preço, o preço diário de fecho, o que não é muito realista tendo em conta que se está a utilizar um preço aproximado à cotação *bid* quando se quer comprar e um preço aproximado à cotação *ask* quando se quer vender.

Numa primeira fase, Gatev et al. (2006) admitiram que os retornos da sua estratégia estão enviesados devido a este fator, afirmando mesmo que parte das divergências de preço observadas se devem potencialmente a movimentos entre cotações *bid* e *ask*. Então, de forma a resolverem este problema, os investigadores decidiram simular as entradas e saídas de mercado, respetivamente, um dia

depois de os pares começarem a divergir e um dia depois de se voltarem a cruzar. Esta alteração resultou numa diminuição, já esperada, dos retornos excessivos médios, de 20-55 pontos-base.

No nosso trabalho, em vez de utilizarmos os preços diários de fecho desfasados no tempo, optamos por utilizar, em cada período de simulação, as reais cotações *bid/ask*, garantindo desta forma que os resultados empíricos obtidos sejam os mais aproximados possíveis de uma aplicação real da estratégia.

No entanto, as cotações *bid/ask* que utilizamos são, na mesma, cotações diárias de fecho, o que na prática significa que não poderíamos utilizar essas cotações para abrir ou fechar posições porque essas são as últimas cotações que o mercado apresenta para cada dia específico. Admitindo esta limitação, assumimos que a abertura e fecho de posições foram feitas no dia a seguir ao que seria suposto, mas com as cotações *bid/ask* de fecho do dia anterior.

Uma forma de resolver esta última limitação seria utilizar, em vez de utilizar cotações *bid/ask* de fecho diárias, utilizar cotações de hora em hora ou de períodos de tempo ainda mais curtos, dado que assim poderíamos simular as transações no mercado no momento exato em que o deveríamos fazer caso estivéssemos a aplicar a estratégia na realidade. Infelizmente, tal não nos foi possível dado que a base de dados Datastream da Thomson Reuters só permite o acesso a preços intradiários para os últimos 3 meses.

Ainda relativamente à periodicidade dos preços, há que sublinhar que Miao (2014) estudou a viabilidade de uma estratégia *Pairs Trading* em que observou preços de 15 em 15 minutos, aplicando *high frequency trading*¹ tendo obtido retornos excessivos de 34,35% num ano. No entanto, não nos é possível comparar os resultados da nossa estratégia com os de Miao (2014), uma vez que o autor

¹“High Frequency Trading (HFT) is a type of algorithmic and quantitative trading which is characterized by short holding periods, specifically the use of sophisticated and powerful computing methods to trade securities rapidly. HFT has positions in equities, options, currencies and all other financial instruments that possess electronic trading capability, aiming to capture small profits and/or fractions of a cent of profit on every short-term trade.” - (Miao, 2014, p. 97)

utilizou outras ferramentas para selecionar os pares de ações, nomeadamente a correlação e a cointegração.

1.3.1 Outras ferramentas: A questão da cointegração

A cointegração é um modelo matemático inovador nas áreas de estatística e econometria, desenvolvido pelos laureados Prémio Nobel, Robert F. Engle e Clive W.J. Granger (1987).

De forma a conseguirmos compreender melhor a utilidade desta ferramenta numa estratégia *Pairs Trading* há que introduzir o conceito de série temporal estacionária, trata-se de um processo estocástico com média e variância, constantes e finitas, e uma autocorrelação independente no tempo, dependendo apenas do número de desfasamentos.

Este tipo de séries temporais, também conhecidas como séries integradas de ordem zero, denominadas de $I(0)$, evoluem ao longo do tempo em torno de uma média constante de longo-prazo apresentando, portanto, características que permitem a sua modelação e previsão.

Por sua vez, as séries temporais dos preços das ações não são estacionárias, mas são sim integradas de ordem “ d ”, denominadas $I(d)$, ou seja, precisam de ser diferenciadas um número mínimo de “ d ” vezes para serem consideradas estacionárias. Isto significa que é inútil modelar e tentar prever o comportamento futuro destas séries temporais.

No entanto, um modelo multivariado em que as séries dos preços das ações sejam cointegradas poderá revelar informações sobre o equilíbrio de longo-prazo no sistema.

If a spread is found to be mean reverting we know that, wherever one series is in several years' time, the other series will be right along there with it. Cointegrated asset prices have a common stochastic trend; they are tied together in the long run because spreads are mean reverting, even though they might drift apart in the short run. – (Alexander, 1999, p. 2042)

Caldeira & Moura (2013) utilizaram a cointegração para identificar os pares de ações que viriam a utilizar na sua estratégia *Pairs Trading*, justificando o uso desta ferramenta da seguinte forma:

Cointegration enable us to combine the two stocks in a certain linear combination so that the combined portfolio is a stationary process. If two stocks share a long-run equilibrium relationship, then deviations from this equilibrium are only short-term and are expected to die out in future periods. – (Caldeira & Moura, 2013, p. 51)

Caldeira & Moura (2013) obtiveram, com a sua estratégia, retornos excessivos anuais médios de 16,38% com baixa volatilidade e fraca correlação com o mercado. Daí que os autores tenham concluído da seguinte forma:

The results show the pairs trading strategy based on cointegration is persistently profitable even in the period of global crisis, reinforcing the usefulness of cointegration in quantitative strategies. – (Caldeira & Moura, 2013, p. 52)

Serve então este subcapítulo para demonstrar que esta ferramenta pode ser utilizada numa estratégia *Pairs Trading*, com potencial sucesso, em substituição do cálculo da distância mínima, apesar de não haver trabalhos de investigação que comparem as duas ferramentas quanto aos resultados, neste tipo de estratégias.²

1.3.2 Outros mercados: A questão dos setores económicos

Gatev et al. (2006) analisaram a sua estratégia em duas bases de dados diferentes, numa incluíram todas as ações disponíveis no mercado, na outra separaram a seleção dos pares por setor económico.

² No nosso trabalho, ponderámos fazer uma análise comparativa entre a aplicação da distância mínima e da cointegração, de forma a verificarmos qual a ferramenta mais adequada para uma estratégia *Pairs Trading*. No entanto, apesar de termos os conhecimentos para testar a cointegração de um determinado par de ações, utilizando o programa Stata por exemplo, faltam-nos os conhecimentos de programação necessários para podermos testar todas as combinações possíveis inerentes a uma base de dados de 2676 ações. De qualquer das formas, fica desde já, uma proposta de trabalho para o futuro.

Os autores observaram que, na estratégia que abrangeu todas as ações do mercado, a percentagem dos pares selecionados que pertencem a setores diferentes foi bastante inferior a 50%, o que indica que é mais provável encontrar potenciais pares de ações quando pertencem ao mesmo setor do que quando pertencem a setores diferentes.

Gatev et al. (2006) consideraram este fenómeno normal uma vez que empresas concorrentes, logo do mesmo setor económico, devem gerar pares de ações naturais dado que estão sujeitas a riscos de mercado semelhantes.

No entanto, este fenómeno não parece ter influência na rentabilidade da estratégia, como demonstraram Gatev et al. (2006), ao obter retornos excessivos mensais médios semelhantes entre os setores: 1,08%, 0,58%, 0,78% e 0,61% nos setores das Utilidades, Transportes, Financeiras e Industriais, respetivamente.

Apesar das conclusões de Gatev et al. (2006), Miao (2014) simulou a sua estratégia apenas no setor energético e, apesar de ter obtido resultados bastante elevados, não é possível perceber se tal se deveu apenas a ter utilizado ações do mesmo setor económico, ou se outros fatores possam ter tido influência, como por exemplo, a já referida utilização de preços intradiários.

Tendo em consideração estes resultados, decidimos aplicar a nossa estratégia utilizando o mercado como um todo, não fazendo a separação entre setores nem escolhendo um setor em particular.

Capítulo 2

Implementação da Estratégia

2.1 Dados

Na nossa estratégia utilizamos os dados referentes às ações americanas, cotadas no New York Stock Exchange (NYSE). Escolhemos o mercado americano porque foi o mercado utilizado em Gatev et al. (2006), o que nos irá permitir fazer comparações de forma correta.

A escolha pela praça financeira NYSE deveu-se ao facto de um dos requisitos para uma empresa aqui poder ser cotada, é o ter mais de 100 000 ações transacionadas por mês, o que é uma maior garantia de liquidez das ações escolhidas para a nossa estratégia.

A partir do Datastream obtivemos os retornos diários (incluindo dividendos), os preços *bid/ask* e o volume transacionado por dia, referentes às ações em questão, para um período de 10,5 anos compreendido entre janeiro de 2006 e junho de 2016.

A nossa amostra consiste em 2 676 empresas, incluindo 752 empresas que já não se encontram ativas e que foram, naturalmente, utilizadas nos períodos de formação em que estiveram ativas, evitando assim a introdução de um efeito de *survivalship bias*.

Separando as empresas por setor, utilizando a *General Industry Classification* da Thomson Reuters, a distribuição é a seguinte:

Setor	Nº Empresas	%
<i>Industrial</i>	1 818	67,94%
<i>Utility</i>	186	6,95%
<i>Transportation</i>	76	2,84%
<i>Bank/Saving & Loans</i>	94	3,51%
<i>Insurance</i>	111	4,15%
<i>Other financial</i>	387	14,46%
NA	4	0,15%
Total	2 676	100%

Tabela 1: Base de dados das empresas por setor económico

2.2 Formação dos pares

Utilizando a metodologia de Gatev et al. (2006), definimos então 19 períodos de formação e 18 períodos de simulação, com a duração de 12 e 6 meses, respetivamente. Cada período de simulação inicia-se imediatamente a seguir ao respetivo período de formação e, por sua vez, cada período de formação inicia-se 6 meses após o início do anterior. A figura abaixo exemplifica o processo de formação de pares:

Jan - Jun 2006	Jul - Dez 2006	Jan - Jun 2007	Jul - Dez 2007	Jan - Jun 2008	Jul - Dez 2008
Formação 1		Simulação 1			
	Formação 2		Simulação 2		
		Formação 3		Simulação 3	
			Formação 4		Simulação 4
				Formação 5	

Figura 1: Exemplo da estrutura dos períodos de formação e simulação

Esta metodologia permite-nos analisar 19 períodos de formação em vez dos 10 que obteríamos caso utilizássemos apenas um período de formação por cada ano civil.

De forma a garantirmos a liquidez das ações que integraram a nossa estratégia, para além de termos escolhido uma praça financeira que apenas permite a entrada de empresas relativamente líquidas (NYSE), aplicámos duas outras metodologias.

Em primeiro lugar, para cada período de simulação, retirámos todas as ações que apresentassem menos do que 200 dias de transações no respetivo período de formação. Para além de garantir a liquidez das ações escolhidas, de forma geral, esta metodologia permite-nos excluir ações de empresas que tenham dado entrada no mercado há pouco tempo e, perante as quais, o mercado ainda se esteja a adaptar.

A segunda metodologia foi adaptada do trabalho de dissertação de Ribeiro (2015), sobre estratégias *Pairs Trading*, onde nos explica que a liquidez de uma determinada ação deve ser, pelo menos, semelhante à liquidez do seu par, porque, caso estejamos perante um par de ações em que uma delas é bastante transacionada e a outra não, poderemos não conseguir abrir posições longas-curtas simultaneamente. Para resolver este problema, Ribeiro (2015) utilizou uma metodologia denominada de *Specific Liquidity* em que tem em conta os preços de fecho diários e o volume transacionado de cada ação, para determinar se a liquidez de cada uma das empresas envolvidas é semelhante ou não.

A fórmula utilizada foi a seguinte:

$$Specific\ Liquidity_i = \bar{X}(Preço\ Diário\ de\ Fecho_{i,t} * Volume\ Transacionado_{i,t})$$

Estabeleceu-se então a seguinte regra: caso a *Specific Liquidity* de uma ação, calculada em cada período de formação, seja superior ao dobro da *Specific Liquidity* do respetivo par (ou vice-versa), esse par será excluído da lista de pares que irá integrar o respetivo período de simulação, independentemente do valor de distância mínima que apresente.

Utilizámos a metodologia de Ribeiro (2015), mas de forma ligeiramente diferente. Ribeiro (2015) selecionou os 20 pares que apresentassem a menor distância entre os seus retornos (normalizados), posteriormente filtrou esses pares relativamente ao critério da *Specific Liquidity* obtendo, em média, 6,74 pares de ações por período de simulação. No nosso trabalho, o que fizemos foi testar as combinações possíveis até obtermos 20 pares que apresentassem a menor distância possível, tendo em conta que tinham de cumprir o requisito da *Specific Liquidity*.

Esta ligeira alteração deve-se a dois motivos. Em primeiro lugar permite-nos obter uma maior significância estatística nos resultados, dada a maior quantidade de pares analisados, apesar de, em contrapartida, não se utilizar pares cuja distância seria mais reduzida caso apenas filtrássemos o top 20. Em segundo lugar, Gatev et al. (2006) observaram que existem benefícios em combinar vários pares no mesmo portefólio dado que, à medida que o número de pares num portefólio aumenta, o retorno (da estratégia) mínimo observado aumenta, o retorno máximo observado mantém-se estável e o desvio-padrão do portefólio diminui.

2.3 Cálculo da distância mínima

No cálculo da distância mínima, seguimos mais uma vez a metodologia de Gatev et al. (2006), em que, inicialmente, normalizámos os retornos (incluindo dividendos reinvestidos) diários de cada ação, para o dia 1 de cada período de formação, utilizando a seguinte fórmula, retirada de Papadakis e Wysocki, (2007, p. 6).

$$P_t^A = \prod_{\tau=1,t} (1 + r_t^A)$$

onde r_t^A é o retorno (incluindo dividendos reinvestidos) da ação A no final do dia t , e P_t^A é o respectivo retorno normalizado.

Após os retornos diários estarem normalizados pudemos calcular a distância mínima entre cada par possível de ações. Para tal, calculamos a soma do quadrado das diferenças dos retornos normalizados de cada par, para cada combinação de par de ações possível, de cada dia do período de formação. Utilizou-se a seguinte fórmula,

$$\text{Distância Mínima}_{A,B} = \sum_{i=1}^n \left(\frac{Pa_{i+1}}{Pa_i} - \frac{Pb_{i+1}}{Pb_i} \right)^2$$

onde $\frac{Pa_{i+1}}{Pa_i}$ e $\frac{Pb_{i+1}}{Pb_i}$ representam os retornos normalizados, para cada dia, das ações A e B , respetivamente.

Utilizando o programa Matlab obtivemos então uma lista das combinações possíveis de pares, sendo que o primeiro par da lista é o que apresenta a menor distância entre os retornos, logo uma maior relação de proximidade entre os preços.

A figura seguinte representa a evolução dos retornos normalizados, de um dos pares de ações que foram selecionados no período de formação 1.

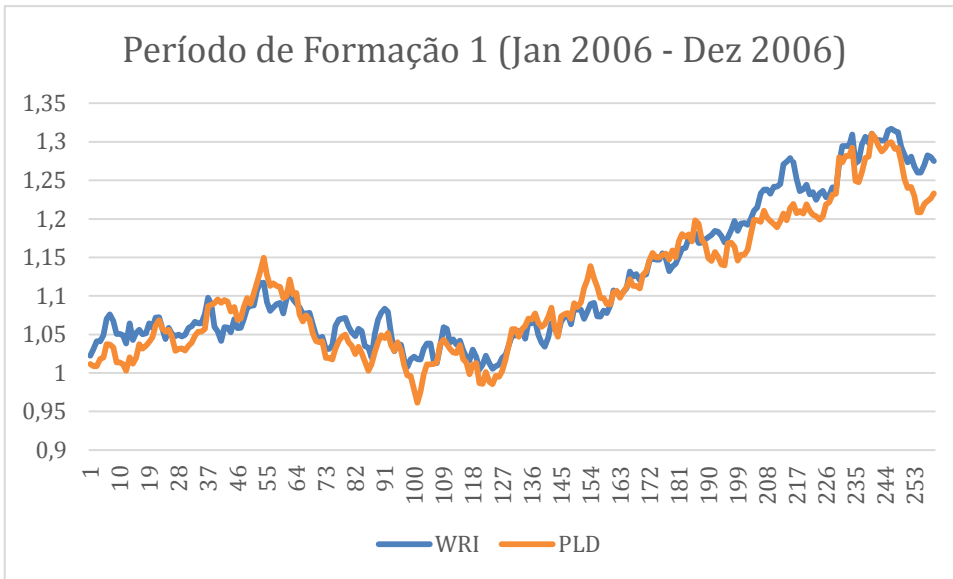


Figura 2: Evolução dos retornos normalizados do par de ações WRI-PLD

Como se pode observar na figura 2, as ações apresentam alterações muito semelhantes entre os preços.

2.4 Sinais de entrada e saída de mercado

Seguindo, mais uma vez, a metodologia de Gatev et al. (2006), abre-se uma posição longa-curta sempre que os retornos normalizados das ações de cada par, selecionado no período de simulação (lado esquerdo da equação), divergirem em mais de dois desvios-padrão históricos, parametrizados no período de formação (lado direito da equação).

$$\left| \frac{Pa_{i+1}}{Pa_i} - \frac{Pb_{i+1}}{Pb_i} \right| > \pm 2 * \text{Desvio - padrão} \left(\frac{Pa_{i+1}}{Pa_i} - \frac{Pb_{i+1}}{Pb_i} \right)$$

Quando isto ocorre,

- Se $\frac{Pa_{i+1}}{Pa_i} > \frac{Pb_{i+1}}{Pb_i}$, abre-se uma posição longa em b e uma posição curta em a ;
- Se $\frac{Pa_{i+1}}{Pa_i} < \frac{Pb_{i+1}}{Pb_i}$, abre-se uma posição longa em a e uma posição curta em b .

Na figura seguinte está representado o período de simulação do par de ações da figura anterior, onde se pode observar que abrimos uma posição no dia 25 e fechámo-la no dia 74 quando os preços se cruzaram, voltando a abrir no dia 97 e fechando novamente no dia 112, quando os preços se voltaram a cruzar, a poucos dias do final do período de simulação. Neste período em concreto, este par de ações obteve um retorno de 10,65% (6 meses).

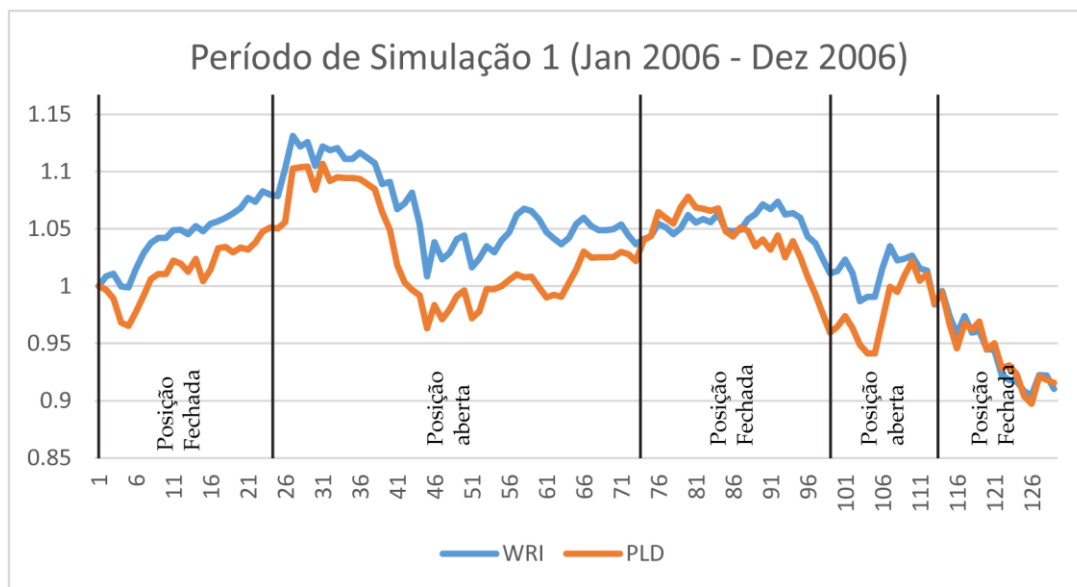


Figura 3: Exemplo do período de simulação do par de ações WRI-PLD

2.5 Cálculo dos excessos de retorno

Relativamente ao cálculo dos retornos em si utilizámos a seguinte fórmula,

$$\left(\frac{P.abertura_{curta} - P.fecho_{curta}}{P.abertura_{curta}} + \frac{P.fecho_{longa} - P.abertura_{longa}}{P.abertura_{longa}} \right)$$

Nas situações em que os retornos (normalizados) de um par de ações, após terem sido abertas posições, não se cruzem até ao fim do período de simulação, fecham-se as posições no último dia do período, registando-se os lucros/prejuízos.

Nos casos em que se abram posições e que, posteriormente, uma das empresas do par deixe de estar ativa no mercado, fecham-se as posições utilizando os preços do último dia em que a empresa esteve ativa.

Capítulo 3

Resultados

3.1 Resultados

Utilizando 19 períodos de simulação de 6 meses cada, simulámos a estratégia para um período contínuo de 9,5 anos, compreendido entre janeiro de 2007 e junho de 2016.

Na tabela seguinte, apresentamos os nossos resultados principais, juntamente com os resultados de Gatev et al. (2006), considerando a estratégia em que utilizaram os preços de fecho diários do próprio dia (Gatev A) e a estratégia em que utilizaram os preços de fecho do dia a seguir (Gatev B), para colmatar o facto de não terem utilizado as cotações *bid/ask*.

Comparando inicialmente as duas estratégias de Gatev et al. (2006), pode-se observar que os retornos de Gatev A são superiores ao dobro dos retornos de Gatev B, evidenciando claramente que os retornos de Gatev A estavam enviesados, devido a terem ignorado a situação das cotações *bid/ask*. Naturalmente, os retornos mais baixos da estratégia Gatev B fazem-se acompanhar também de uma diminuição da volatilidade, observável através dos valores do desvio-padrão, valor mínimo e valor máximo, e de uma diminuição da distorção.

Resultados	Gatev A (5/20)	Gatev B (5/20)	Estratégia
Retorno (6 meses)	40,90%	16,39%	12,89%
Retorno Total	777,10%	311,41%	244,93%
Retorno por par (média)	1,31% / 1,44%	0,78% / 0,81%	0,92%
Retorno máximo	350,51%	175,63%	149,58%
Retorno mínimo	-42,24%	-27,62%	-79%
Mediana	23,90%	18,04%	8,73%
Curtose	4,72	7,44	2,79
Distorção	4,02	2,29	1,04
Desvio-padrão	0,95	0,44	0,47
Nº de Pares	25	25	20
Nº de transações (média)	32,31	10,94	5,74
Nº dias posição (média)	73,81 (2,46 meses)	75,52 (2,52 meses)	78,92 (2,63 meses)

Tabela 2: Resultados das estratégias de Gatev et al. (2006) e da nossa

Observando agora a nossa estratégia, é possível verificar-se que os resultados são, em muito, semelhantes aos da estratégia Gatev B, se bem que ligeiramente inferiores a nível dos retornos, o que significa que mesmo utilizando a política de abrir/fechar posições um dia a seguir, a estratégia Gatev B continua ligeiramente enviesada.

Pode-se observar que os valores da mediana, da curtose e da distorção, da nossa estratégia, são mais baixos que os da estratégia Gatev B, o que significa que os valores dos retornos dos pares da nossa estratégia aproximam-se mais do seu valor médio, comparativamente. Estes valores permitem-nos argumentar que, ao seleccionarmos ações com características de liquidez mais semelhantes, poderemos estar a acrescentar alguma consistência à estratégia sem que as rentabilidades média e total se alterem muito, isto apesar de termos obtido um retorno mínimo bastante elevado, em termos absolutos.

Uma outra informação importante que se retira da tabela 2 é a grande diferença que existe entre o número médio de transações das duas estratégias, 10,94 em Gatev B e 5,74 na nossa. Na prática, isto significa que o número de períodos de

divergência/convergência em cada par foi, em média, bastante inferior na nossa estratégia.

Tendo em consideração que as grandes diferenças entre a nossa estratégia e a de Gatev B foram as restrições de liquidez impostas, podemos argumentar que uma maior semelhança entre as ações leva a um menor número de períodos de divergência/convergência, sem que, contudo, a rentabilidade total e média tenham diminuído, pelo contrário, uma vez que, como um menor número de transações se traduz em menores custos de transação, a rentabilidade real da nossa estratégia poderá ser superior.

Relativamente ao número de dias que cada posição está aberta, verifica-se que não existe grande diferença entre as estratégias, o que é normal dado que também a nossa estratégia utiliza apenas ações americanas.

3.2 Resultados detalhados da estratégia

Iremos agora fazer uma análise detalhada dos resultados obtidos na nossa estratégia, onde apresentamos os resultados das transações positivas e negativas, os das transações que fecharam antes do último dia do respetivo período de simulação, e os das transações que foram forçados a fechar.

De notar que, na tabela 2, uma transação corresponde a uma abertura ou fecho de posição numa ação, ou seja, se abirmos uma posição longa em A e uma posição curta em B, e depois fecharmos a posição longa de A e a posição curta de B, temos 4 transações. No entanto, na tabela seguinte, uma transação corresponde à abertura e fecho de uma posição o que, no exemplo anterior, corresponde apenas a 2 transações.

Períodos		Transações		Transações positivas		Transações negativas		Fecharam		Não fecharam	
Ano		Nº	Retorno (%)	Nº	Retorno (%)	Nº	Retorno (%)	Nº	Retorno (%)	Nº	Retorno (%)
2007	1ª metade	56	42,02%	30	188,99%	26	-146,97%	26	82,59%	30	-40,56%
	2ª metade	78	-19,91%	43	257,23%	35	-277,14%	44	169,66%	34	-189,59%
2008	1ª metade	70	51,20%	35	265,42%	35	-214,21%	32	140,67%	38	-89,46%
	2ª metade	92	-11,30%	52	544,92%	40	-556,23%	52	256,58%	40	-267,87%
2009	1ª metade	65	149,58%	40	462,80%	26	-313,21%	28	134,42%	38	15,14%
	2ª metade	66	-79,00%	36	427,29%	32	-506,28%	30	155,95%	38	-234,94%
2010	1ª metade	34	23,13%	21	98,77%	13	-75,68%	10	41,12%	24	-18,02%
	2ª metade	44	27,32%	24	138,52%	18	-111,21%	12	45,00%	30	-17,67%
2011	1ª metade	50	-24,22%	24	131,02%	26	-155,26%	16	56,63%	34	-80,82%
	2ª metade	55	8,73%	27	231,12%	27	-222,40%	22	69,11%	32	-60,38%
2012	1ª metade	48	-25,10%	26	118,64%	22	-143,74%	12	27,19%	36	-52,28%
	2ª metade	50	29,57%	32	104,50%	18	-74,92%	22	64,97%	28	-35,39%
2013	1ª metade	50	79,42%	29	177,42%	19	-98,03%	24	109,81%	24	-30,38%
	2ª metade	55	-27,99%	23	108,72%	25	-136,71%	10	39,17%	38	-67,17%
2014	1ª metade	50	19,38%	26	173,18%	24	-153,81%	14	39,84%	36	-20,46%
	2ª metade	58	22,99%	29	181,92%	29	-158,91%	20	70,80%	38	-47,80%
2015	1ª metade	48	-19,80%	26	154,05%	22	-173,88%	16	44,73%	32	-64,52%
	2ª metade	64	4,69%	35	163,17%	27	-158,47%	26	53,20%	36	-48,51%
2016	1ª metade	58	-5,79%	34	254,93%	24	-260,72%	18	48,98%	40	-54,76%
Média		57	12,89%	31	220,14%	26	-207,25%	23	86,86%	34	-73,97%

Tabela 3: Resultados detalhados da estratégia

Observando a tabela 3 é possível verificar-se que o número médio de transações positivas (31), por período, é superior ao número médio de transações negativas (26), no entanto, tal não precisaria de acontecer para que a estratégia tivesse sucesso.

Quanto às transações que fecharam antes do último dia do período (23) é expectável que todas apresentem resultados positivos uma vez que, se fecharam, foi porque os preços convergiram e se cruzaram, após as posições terem sido abertas, o que resulta em situações de lucro, com exceção de pares cujas ações tenham saído do mercado.

Relativamente às transações que não fecharam antes do último dia do período (34), ou seja, aqueles pares que foram forçados a fechar no último dia, é também expectável que os seus retornos médios sejam negativos, uma vez que representam pares de ações que após terem divergido não voltaram a convergir, pelo menos naquele período, o que significa que, quando forçámos o fecho no último dia, a distância entre os pares, na maioria dos casos, era maior do que quando abrimos as posições, resultando, portanto em prejuízos.

Podemos afirmar que estes pares de ações não conseguiram fechar antes do último dia do período porque, ou não tiveram tempo para o fazer, pois as posições podem ter sido abertas numa data próxima do final do período, ou porque a identificação da relação de proximidade entre os pares de ações foi mal conseguida, pelos menos nalguns casos.

As soluções possíveis para este problema, caso se trate realmente de uma questão de tempo para os pares convergirem, podem ser ou estabelecer uma data limite a partir da qual não se abrem posições, tendo em conta o número de dias médio que cada posição se mantém aberta (79 dias aproximadamente), ou então alargar o período de simulação/*trading*.

Iremos agora analisar em detalhe os retornos das transações consoante o dia em que tenham sido fechados, ou seja, a coluna “Fecharam” representa os pares que

fecharam antes do último dia do período, e a coluna “Não fecharam” que representa os pares que foram forçados a fechar no último dia.

Períodos		Fecharam				Não fecharam			
		Positivos		Negativos		Positivos		Negativos	
Ano		Nº	Retorno (%)	Nº	Retorno (%)	Nº	Retorno (%)	Nº	Retorno (%)
2007	1ª metade	26	82,59%	0	0,00%	10	46,32%	20	-86,88%
	2ª metade	44	169,66%	0	0,00%	8	7,97%	26	-197,56%
2008	1ª metade	32	140,67%	0	0,00%	8	5,01%	30	-94,47%
	2ª metade	52	256,58%	0	0,00%	16	35,59%	24	-303,46%
2009	1ª metade	28	134,42%	0	0,00%	18	89,54%	20	-74,40%
	2ª metade	30	155,95%	0	0,00%	2	1,97%	36	-236,91%
2010	1ª metade	10	41,12%	0	0,00%	18	28,01%	6	-46,03%
	2ª metade	12	45,00%	0	0,00%	12	19,70%	18	-37,37%
2011	1ª metade	16	56,63%	0	0,00%	6	10,22%	28	-91,04%
	2ª metade	22	69,11%	0	0,00%	14	28,22%	18	-88,60%
2012	1ª metade	12	27,19%	0	0,00%	8	13,00%	28	-65,28%
	2ª metade	22	64,97%	0	0,00%	10	7,87%	18	-43,26%
2013	1ª metade	24	109,81%	0	0,00%	10	5,36%	14	-35,74%
	2ª metade	10	39,17%	0	0,00%	14	15,67%	24	-82,84%
2014	1ª metade	14	39,84%	0	0,00%	12	24,32%	24	-44,78%
	2ª metade	20	70,80%	0	0,00%	20	15,39%	18	-63,19%
2015	1ª metade	16	44,73%	0	0,00%	12	11,80%	22	-76,33%
	2ª metade	26	53,20%	0	0,00%	18	20,06%	16	-68,57%
2016	1ª metade	18	48,98%	0	0,00%	14	11,59%	26	-66,35%
Média		23	86,86%	0	0,00%	12	20,93%	22	-94,90%

Tabela 4: Resultados detalhados dos pares consoante o dia de fecho

Como se pode observar na tabela anterior, todos as transações que fecharam até ao último dia do respetivo período apresentaram retornos positivos, o que é normal, como já referimos. Houve, no entanto, um par de ações que teve que ser fechado antes do último dia porque uma das ações saiu do mercado, o que

poderia ter dado prejuízo, mas tal não aconteceu porque a diferença entre os seus preços, no dia de fecho, foi inferior à diferença no dia de abertura.

Como já tínhamos visto, em média, por período, 23 transações fecharam antes do último dia do período, com um retorno médio por período de 86,86%.

Relativamente às transações que foram forçados a fechar no último dia do período, em média, 12 por período apresentaram retornos positivos, o que equivale a um valor médio de 35,29% de transações positivas, no conjunto dos que não fecharam antes do último dia. Estes valores levantam a questão se o prazo de cada período de *trading* (ou simulação no caso de trabalho de investigação) não deveria ser alargado, tendo em conta que, em média, 35,29% das transações poderiam já estar a convergir, podendo vir, eventualmente, a cruzarem-se o que levaria a retornos superiores.

3.3 Neutralidade da estratégia face ao mercado

Considerando então que uma das principais características das estratégias *Pairs Trading* é apresentar neutralidade de resultados face ao mercado, resolvemos comparar a rentabilidade da nossa estratégia face à do NYSE.

A figura seguinte representa a evolução da rentabilidade da nossa estratégia e do NYSE para os 19 períodos de simulação da nossa estratégia, o que corresponde a um período compreendido entre janeiro 2007 e junho 2016.

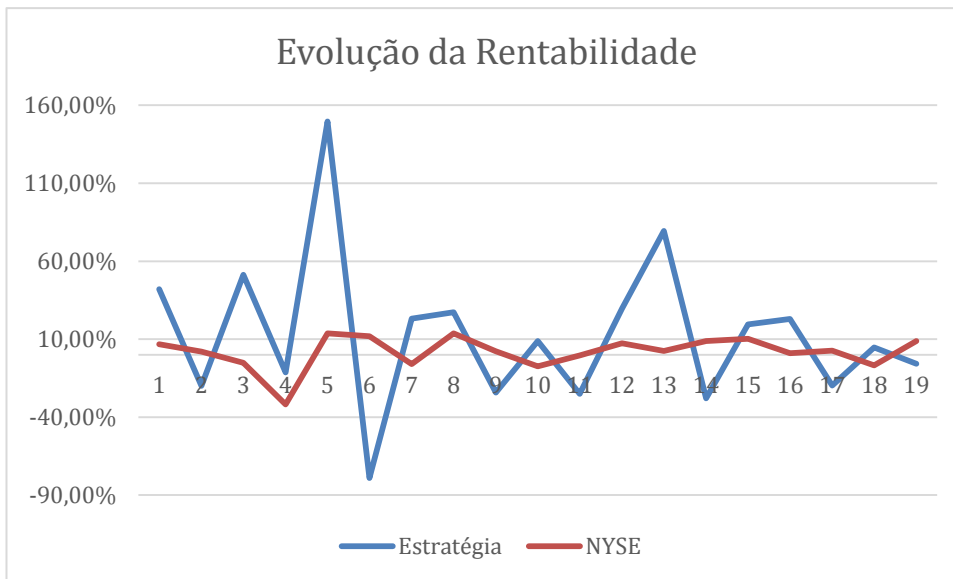


Figura 4: Evolução da rentabilidade da estratégia e do NYSE

Observando a figura 4 é possível perceber que a nossa estratégia *Pairs Trading* apresenta uma volatilidade bastante superior relativamente ao NYSE.

A correlação entre as evoluções de rentabilidade é de 15,13% o que indica que, apesar de forma reduzida, a estratégia está positivamente correlacionada com o mercado, o que nos parece indicar que a estratégia não apresenta neutralidade face ao mercado, como seria de esperar.

Capítulo 4

Conclusões

4.1 Conclusões

No nosso trabalho replicamos então a estratégia *Pairs Trading* de Gatev et al. (2006) com duas grandes diferenças. Em primeiro lugar, aplicamos restrições de liquidez na seleção dos pares tais como, a utilização da praça financeira NYSE que é constituída por ações de elevada liquidez, o requisito de um mínimo de 200 dias de transações, por cada ação, em cada período de formação, e aplicámos o critério *Specific Liquidity*, que garante que cada par de ações apresenta semelhanças aos níveis dos preços de fecho e volume de transações. Em segundo lugar, utilizámos cotações *bid/ask* em vez dos preços diários de fecho, com o objetivo de aproximar a estratégia o mais possível a uma aplicação real.

Simularam-se transações para 20 pares de ações, para cada um dos 19 períodos de simulação, em que o retorno excessivo médio foi de 12,89% (6 meses), o retorno excessivo total foi de 244,93% (9,5 anos), o retorno excessivo médio por par foi de 0,92% e os retornos excessivos mínimos e máximos foram de -79% e 149,58%, respetivamente. Estes resultados provam então que, utilizando esta metodologia, as estratégias *Pairs Trading* continuam a apresentar retornos excessivos elevados.

Relativamente ao trabalho de Gatev et al. (2006), os retornos da nossa estratégia foram ligeiramente inferiores devido à utilização de cotações *bid/ask* em vez dos preços diários de fecho, o que prova que a estratégia de Gatev et al. (2006) está

ligeiramente enviesada, como os próprios admitiram, mesmo tendo utilizado os preços do dia a seguir.

Observamos também que na nossa estratégia o número de transações necessário foi bastante inferior ao necessário na estratégia de Gatev et al. (2006). Argumentámos que este fator se deveu às restrições de liquidez impostas e que, na prática, podem traduzir-se num aumento da rentabilidade real, tendo em consideração os custos de transação.

Quanto à neutralidade da estratégia face ao mercado, observámos que existe uma correlação bastante baixa, mas positiva, de 15,13%, o que parece indicar que a estratégia não é neutra face ao mercado. No entanto, não nos devemos precipitar na conclusão de que as estratégias *Pairs Trading* não são neutras face ao mercado, pois tal resultado pode dever-se a não termos utilizado uma amostra suficientemente grande.

Como sugestão para trabalhos futuros, é nossa opinião que é importante testar a cointegração como ferramenta de seleção de pares, testar períodos mais alargados de simulação e, por último, simular os resultados tendo em conta os custos de transação em vigor no mercado.

Bibliografia

Ait-Sahalia, Y. & Hansen, L. P. 2009. **Handbook of Financial Econometrics**. Oxford/Amsterdam: Elsevier.

Alexander, C. 1999. Optimal Hedging Using Cointegration. **Journal of Philosophical Transactions of the Royal Society A**, 357: 2039-2058.

Bachelier, L. 1900. **Théorie de la Speculation**. Paris: Gauthier-Villars, and also published in English in Cootner, P. **The Random Character of Stock Market Prices**. Cambridge: M.I.T., 1964, pp. 17-78.

Bodie, Z.; Kane, A. & Marcus, A. 2010. **Investments**. New York: McGraw-Hill/Irwin.

Caldeira, J. F. & Moura, G. V. 2013. Selection of a Portfolio of Pairs Based on Cointegration: A Statistical Arbitrage Strategy. **Revista Brasileira de Finanças**, Vol. 11, No. 1: 49-80.

Dewatripont, M; Hansen, L. P. & Turnovsky, S. J. 2003. **Advances in Economics and Econometrics**. Cambridge: Cambridge University Press.

Engle, R. F. & Granger, C. W. J. 1987. Co-Integration and Error Correction: Representation, Estimation and Testing. **Journal of the Econometric Society**, 55: 251-276.

Fama, E. F. 1965. The Behavior of Stock-Market Prices. **Journal of Business**, 38: 34-105.

Fama, E. F. & Blume, M. 1966. Filter Rules and Stock Market Trading Profits. **Journal of Business**, 39: 226-41.

Fama, E. F. 1969. The Adjustment of Stock Prices to New Information. **International Economic Review**, Vol. 10, No. 1: 1-21.

Fama, E. F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. **Journal of Finance**, 25: 383-417.

Gatev, E.; Goetzmann, W. & Rouwenhorst, K. 2006. Pairs Trading: Performance of a Relative-Value Arbitrage Rule. **Review of Financial Studies**, 19: 797-827.

Ingersol, J. E. Jr. 1987. **Theory of Financial Decision Making**. Lanham: Rowman & Littlefield Publishers, Inc.

Lintner, J. 1965. Security Prices, Risk, and Maximal Gains from Diversification. **Journal of Finance**, 20: 587-615.

Mandelbrot, B. 1966. Forecasts of Future Prices, Unbiased Markets, and Martingale Models. **Journal of Business**, 39: 242-55.

Markowitz, H. 1952. Portfolio Selection. **Journal of Finance**, 7: 1-119.

Mossin, J. 1966. Equilibrium in a Capital Asset Market. **Econometrica: Journal of the econometric society**, Vol. 34, No 4: 768-783.

Miao, G. 2014. High Frequency and Dynamic Pairs Trading Based on Statistical Arbitrage Using a Two-Stage Correlation and Cointegration Approach. **International Journal of Economics and Finance**, 6: 96-110.

Papadakis, G. & Wysocki, P. 2007. **Pairs Trading and Accounting Information**. Working Paper

Perlin, M. 2009. Evaluation of Pairs-Trading Strategy at the Brazilian Financial Market. **Journal of Derivatives and Hedge Funds**, 15: 122-136.

Ribeiro, R. 2015. **Pairs Trading: Strategy Refinements**. Porto: Universidade Católica do Porto.

Samuelson, P. A. 1965. Proof That Properly Anticipated Prices Fluctuate Randomly. **Industrial Management Review**, 6: 41-9.

Sharpe, W. F. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. **Journal of Finance**, Vol. 19, No.3: 425-442.

Shiller, R. J. 1980. Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? **The American Economic Review**, Vol. 71, No. 3: 421-436.

Shiller, R. J. 2000. **Irrational Exuberance**. Princeton: Princeton University Press.

Vidyamurthy, G. 2004. **Pairs Trading: Quantitative Methods and Analysis.**
Hoboken, New Jersey: John Wiley & Sons, Inc.