

FUNDAÇÃO GETULIO VARGAS  
ESCOLA DE ECONOMIA DE SÃO PAULO

PAULA ANDREA SOTO

**ARBITRAGEM ESTATÍSTICA NO MERCADO  
BRASILEIRO DE AÇÕES: UMA ABORDAGEM POR  
VECM**

**SÃO PAULO**

**2016**

PAULA ANDREA SOTO

**ARBITRAGEM ESTATÍSTICA NO MERCADO  
BRASILEIRO DE AÇÕES: UMA ABORDAGEM POR  
VECM**

Dissertação apresentada ao Programa de  
Mestrado Profissional da Escola de Economia  
de São Paulo da Fundação Getúlio Vargas,  
como requisito para a obtenção do título de  
Mestre em Economia.

Área de concentração:  
Finanças Quantitativas.

Orientador:  
Prof. Dr. Juan Carlos Ruilova Terán

SÃO PAULO

2016

Soto, Paula Andrea.

Arbitragem estatística no mercado brasileiro de ações: uma abordagem por VECM / Paula Andrea Soto. - 2016.  
63 f.

Orientador: Juan Carlos Ruilova Terán

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Ações (Finanças). 2. Mercado de opções. 3. Modelos econométricos. 4. Cointegração. 5. Volatilidade (Finanças). I. Ruilova Terán, Juan Carlos. II. Dissertação (MPFE) - Escola de Economia de São Paulo. III. Título.

CDU 336.764.2(81)

PAULA ANDREA SOTO

# **ARBITRAGEM ESTATÍSTICA NO MERCADO BRASILEIRO DE AÇÕES: UMA ABORDAGEM POR VECM**

Dissertação apresentada ao Programa de Mestrado Profissional da Escola de Economia de São Paulo da Fundação Getúlio Vargas, como requisito para a obtenção do título de Mestre em Economia.

Área de concentração:  
Finanças Quantitativas.

Data da Aprovação: 11 / 08 / 2016

Banca Examinadora:

---

**Prof. Dr. Juan Carlos Ruilova Terán**  
(Orientador)  
FGV - EESP

---

**Prof. Dr. Afonso de Campos Pinto**  
FGV - EESP

---

**Prof. Dr. Alessandro Marques**  
FGV - EESP

# RESUMO

Ao modelar séries de preços de ativos financeiros, a prática usual é tomar a primeira diferença das séries, e trabalhar assim com retornos ou logretornos. Utilizando VECM (*Vector Error Correction Models*, em inglês), torna-se possível trabalhar diretamente com as séries sem diferenciar, o que possibilita o estudo de tendências comuns e cointegração. Este trabalho utiliza VECM para gerar estratégias de arbitragem estatística no mercado brasileiro de ações. Tendências comuns são identificadas por PCA (*Principal Components Analysis*, em inglês, ou análise de componentes principais, em português) e os resultados foram utilizados para definir portfólios cointegrados. Foram propostos dois métodos de geração de sinais para estratégias de *trading* do tipo *longshort*. Um total de cinco diferentes estratégias de *trading* foram simuladas e a existência de arbitragem estatística em cada caso foi testada pelo teste proposto em (JARROW et al., 2012). Conclui-se que, ao considerar séries de preços não diferenciadas, a metodologia abordada permite identificar e modelar candidatos de portfólios cointegrados. Quando bem calibradas, as estratégias testadas geram ganhos significativos em todos os portfólios.

**Palavras-chave:** Arbitragem estatística. Cointegração. PCA. VECM.

# ABSTRACT

*Common practice for modelling stock prices is to use their differences in form of returns or logreturns. Using VECM (Vector Error Correction Models), it is possible to work with the series of prices without differentiation, which allows looking into common trends and cointegration. This work uses VECM to create trading strategies for the Brazilian stock market. Common trends are obtained using PCA (Principal Components Analysis) and prices are modelled using VECM. Five longshort-type trading strategies are simulated in diversified portfolios, and tested for statistical arbitrage using the test proposed by (JARROW et al., 2012). The methodology for identifying common trends and modelling prices allows for trading strategies with good results for all portfolios.*

**Keywords:** *Statistical arbitrage. Cointegration. PCA. VECM.*

# Lista de ilustrações

|   |    |
|---|----|
| Figura 1 – Valores diários dos índices IBOVESPA e IBX100 . . . . .  | 22 |
| Figura 2 – Fatores latentes extraídos por PCA, utilizando todo o período da amostra (2007-2015). . . . .                                | 23 |
| Figura 3 – Conjunto de séries $X_t$ no período 2007-2015 para diferentes tipos de seleção de portfólio . . . . .                        | 27 |
| Figura 4 – $S_t$ estimadas na preamostra, considerando o VECM com 1 <i>lag</i> e 27 vetores de cointegração. . . . .                    | 28 |
| Figura 5 – Ganhos para os portfólios de um fator, detalhados pelo número de ativos que compõem os portfólios. . . . .                   | 37 |
| Figura 6 – Ganhos para os portfólios de um fator com $N = 10$ , detalhados pelo número de vetores de cointegração considerados. . . . . | 38 |
| Figura 7 – Ganhos para o portfólio de três fatores . . . . .  | 39 |
| Figura 8 – Ganhos para o portfólio com ativos da mesma empresa, PETR3 e PETR4   | 40 |
| Figura 9 – Ganhos para o portfólio do setor bancário . . . . .  | 41 |
| Figura 10 – Resultado do teste JTTW para todas as estratégias no portfólio de três fatores . . . . .                                    | 42 |
| Figura 11 – <i>Backtest</i> : Preços diários dos ativos do portfólio de três fatores . . . . .  | 44 |
| Figura 12 – <i>Backtest</i> : Resultado das estratégias no portfólio de três fatores . . . . .  | 45 |
| Figura 13 – Detalhamento do resultado para a estratégia (1) no portfólio de um fator com $N = 10$ ativos . . . . .                      | 51 |
| Figura 14 – Detalhamento do resultado para a estratégia (2) no portfólio de um fator com $N = 10$ ativos . . . . .                      | 51 |
| Figura 15 – Detalhamento do resultado para a estratégia (3) no portfólio de um fator com $N = 10$ ativos . . . . .                      | 52 |
| Figura 16 – Detalhamento do resultado para a estratégia (4) no portfólio de um fator com $N = 10$ ativos . . . . .                      | 53 |
| Figura 17 – Detalhamento do resultado para a estratégia (5) no portfólio de um fator com $N = 10$ ativos . . . . .                      | 54 |
| Figura 18 – Detalhamento do resultado para a estratégia (1) no portfólio de três fatores . . . . .                                      | 54 |
| Figura 19 – Detalhamento do resultado para a estratégia (2) no portfólio de três fatores . . . . .                                      | 55 |
| Figura 20 – Detalhamento do resultado para a estratégia (3) no portfólio de três fatores . . . . .                                      | 55 |
| Figura 21 – Detalhamento do resultado para a estratégia (4) no portfólio de três fatores . . . . .                                      | 56 |

|   |    |
|---|----|
| Figura 22 – Detalhamento do resultado para a estratégia (5) no portfólio de três<br>fatores . . . . . | 56 |
|---|----|



# Lista de tabelas

|   |    |
|---|----|
| Tabela 1 – Valor-p do teste KPSS para fatores estimados por PCA, por ano. . . .               | 24 |
| Tabela 2 – Valor-p do teste de raiz unitária para fatores estimados por PCA, por ano. . . . . | 24 |
| Tabela 3 – Valor-p do teste KPSS para fatores estimados por ICA, por ano. . . .               | 24 |
| Tabela 4 – Valor-p do teste de raiz unitária para fatores estimados por ICA, por ano.         | 25 |
| Tabela 5 – Variância explicada acumulada dos fatores estimados por PCA, por ano.              | 25 |
| Tabela 6 – Variância explicada acumulada dos fatores estimados por ICA, por ano.              | 25 |
| Tabela 7 – Combinação de parâmetros para cada estratégia . . . . .                            | 32 |
| Tabela 8 – Resumo da estratégia (1) no portfólio de um fator . . . . .                        | 57 |
| Tabela 9 – Resumo da estratégia (2) no portfólio de um fator . . . . .                        | 57 |
| Tabela 10 – Resumo da estratégia (3) no portfólio de um fator . . . . .                       | 57 |
| Tabela 11 – Resumo da estratégia (4) no portfólio de um fator . . . . .                       | 58 |
| Tabela 12 – Resumo da estratégia (5) no portfólio de um fator . . . . .                       | 58 |
| Tabela 13 – Resumo da estratégia (1) no portfólio de três fatores . . . . .                   | 58 |
| Tabela 14 – Resumo da estratégia (2) no portfólio de três fatores . . . . .                   | 59 |
| Tabela 15 – Resumo da estratégia (3) no portfólio de três fatores . . . . .                   | 59 |
| Tabela 16 – Resumo da estratégia (4) no portfólio de três fatores . . . . .                   | 60 |
| Tabela 17 – Resumo da estratégia (5) no portfólio de três fatores . . . . .                   | 60 |
| Tabela 18 – Resumo da estratégia (1) no portfólio composto por PETR3 e PETR4                  | 60 |
| Tabela 19 – Resumo da estratégia (2) no portfólio composto por PETR3 e PETR4                  | 61 |
| Tabela 20 – Resumo da estratégia (3) no portfólio composto por PETR3 e PETR4                  | 61 |
| Tabela 21 – Resumo da estratégia (4) no portfólio composto por PETR3 e PETR4                  | 61 |
| Tabela 22 – Resumo da estratégia (5) no portfólio composto por PETR3 e PETR4                  | 61 |
| Tabela 23 – Resumo da estratégia (1) no portfólio do setor bancário . . . . .                 | 61 |
| Tabela 24 – Resumo da estratégia (2) no portfólio do setor bancário . . . . .                 | 62 |
| Tabela 25 – Resumo da estratégia (3) no portfólio do setor bancário . . . . .                 | 62 |
| Tabela 26 – Resumo da estratégia (4) no portfólio do setor bancário . . . . .                 | 63 |
| Tabela 27 – Resumo da estratégia (5) no portfólio do setor bancário . . . . .                 | 63 |

# Sumário

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Introdução</b>   | <b>11</b> |
| <b>2</b> | <b>Revisão Bibliográfica</b>  | <b>13</b> |
| <b>3</b> | <b>Objetivo</b>   | <b>15</b> |
| <b>4</b> | <b>Conceitos teóricos</b>   | <b>16</b> |
| 4.1      | Séries de preços  | 16        |
| 4.2      | Tendências comuns   | 16        |
| 4.3      | PCA e ICA   | 17        |
| 4.4      | Integração e cointegração   | 17        |
| 4.5      | Testes de estacionariedade, raiz unitária e cointegração                  | 18        |
| 4.6      | VECM  | 18        |
| 4.7      | AIC e BIC   | 19        |
| 4.8      | Teste de arbitragem estatística   | 19        |
| <b>5</b> | <b>Seleção de portfolio</b>   | <b>21</b> |
| 5.1      | Os dados  | 21        |
| 5.2      | Determinando a existência de tendências comuns                            | 22        |
| 5.3      | Seleção de portfólios candidatos  | 25        |
| <b>6</b> | <b>As estratégias de <i>trading</i></b>                                   | <b>28</b> |
| 6.1      | Proposição de estratégias definidas em função dos vetores de cointegração | 28        |
| 6.2      | Estratégias definidas em função de <i>forecasts</i>                       | 30        |
| 6.3      | Implementação   | 31        |
| 6.3.1    | Custos de operação e a taxa livre de risco                                | 32        |
| <b>7</b> | <b>Resultados</b>   | <b>34</b> |
| 7.1      | Portfólio de um fator latente   | 34        |
| 7.2      | Portfólio de três fatores latentes  | 35        |
| 7.3      | Portfólios de ativos da mesma empresa                                     | 36        |
| 7.4      | Portfólios de ativos do mesmo setor                                       | 36        |
| 7.5      | Teste de arbitragem estatística para o portfólio de três fatores          | 40        |
| 7.6      | <i>Backtest</i>   | 43        |
| <b>8</b> | <b>Considerações finais</b>   | <b>46</b> |
|          | <b>Referências</b>  | <b>48</b> |

|   |           |
|---|-----------|
| <b>Apêndices</b>  | <b>50</b> |
| <b>APÊNDICE A Gráficos dos ganhos acumulados das simulações</b> | <b>51</b> |
| A.1 Portfólio de um fator                                       | 51        |
| A.2 Portfólio de três fatores                                   | 54        |
| <b>APÊNDICE B Medidas resumo das simulações</b>                 | <b>57</b> |
| B.1 Portfólio de um fator com $N = 10$ ativos                   | 57        |
| B.2 Portfólio de três fatores                                   | 58        |
| B.3 Portfólio composto por PETR3 e PETR4                        | 60        |
| B.4 Portfólio com ativos do setor bancário                      | 61        |

# 1 Introdução

Técnicas que procuram se beneficiar de mercados inefficientes já existem há décadas. Com origem nos anos 1980, o *pairs trading* é considerado o pai das estratégias de arbitragem estatística. A ideia consistia em identificar um par de ativos cujos preços seguissem uma tendência comum, de modo que o *spread* entre eles apresentasse uma reversão à média. Ativos com essa característica são chamados de ativos cointegrados, e sua identificação e modelagem vem sendo aprofundada ao longo dos anos a fim de identificar oportunidades de arbitragem, especialmente com o uso de estratégias *longshort*. Este tipo de estratégia tornou-se muito popular por ser teoricamente independente de tendências do mercado, pois ela depende apenas a relação entre os ativos.

Com o tempo, a técnica de *pairs trading* foi sendo generalizada para técnicas mais abrangentes de arbitragem estatística, sempre com o objetivo de gerar retornos com risco zero, procurando falhas na precificação de ativos e instrumentos financeiros. Estas falhas tornaram-se mais raras e difíceis de detectar a medida que os mercados foram ganhando liquidez e os preços se ajustaram mais rapidamente ao valor justo de mercado. Deste modo, os modelos ganharam complexidade para poder identificar tais oportunidades. Para construir um portfólio, em vez de procurar por dois ativos e uma tendência comum, passou-se a considerar grandes grupos de ativos cointegrados com múltiplas tendências comuns influenciando seus preços.

Modelos de séries temporais destacaram-se como ferramenta principal de análise. Durante muito tempo, os modelos disponíveis supunham que as séries eram estacionárias e consequentemente, para trabalhar com séries financeiras, a prática usual sempre foi diferenciá-las. Para poder estudar a relação entre as tendências das séries, os modelos precisaram ser adaptados para lidar com não estacionariedade e, paralelamente, métodos para identificar a existência de tendências comuns ficaram cada vez mais relevantes.

Como será mostrado nos próximos capítulos, é possível modelar um conjunto de séries de preços de ações não estacionárias utilizando VECM. O objetivo é estudar a cointegração das séries de preços sem diferenciar, o que é uma grande vantagem para criação de estratégias *longshort*. Aplicando o modelo em séries de preços do mercado brasileiro, é possível criar estratégias de *trading* que se aproveitam de oportunidades de arbitragem estatística. As estratégias propostas são definidas pelos vetores de cointegração estimados pelo VECM, e os portfólios de ativos são determinados pelos fatores latentes estimados através de PCA. Os resultados das duas estratégias de *trading* propostas são comparados àqueles de estratégias usuais onde os sinais de compra e venda são definidos pelos *forecasts*, na mesma linha dos modelos de (FOCARDI; FABOZZI; MITOV, 2016).

O trabalho é estruturado da seguinte forma: no capítulo 2 faz-se uma revisão da literatura referente à arbitragem estatística e à modelagem de séries cointegradas. O capítulo 3 explica o objetivo deste trabalho. No capítulo 4 faz-se uma breve revisão dos conceitos teóricos chaves de séries temporais, a formulação de VECM e o tipo de parametrização que será utilizada. No capítulo 5 é feita uma análise da amostra e são explicados os critérios de seleção de portfólios. O capítulo 6 descreve as metodologias propostas para gerar os sinais das estratégias *longshort*. Os resultados das simulações das estratégias de *trading* são apresentados no capítulo 7. Conclusões finais são apresentadas no capítulo 8.

## 2 Revisão Bibliográfica

A arbitragem estatística teve sua origem no pairs trading, com os primeiros trabalhos datados de 1980, incluindo (BOSSAERTS, 1988), que mostrou evidências de cointegração e (LEHMANN, 1990) identificando oportunidades de arbitragem a curto prazo.

Desde então, houve muitos aportes à literatura sobre técnicas para identificação de pares. Um exemplo é (ZHANG; DAI; DJARBIROV, 2014), que utiliza PCA e *clustering* para encontrar candidatos de portfólios cointegrados, incluindo nas análises outras variáveis além do preço do ativo. Em relação à modelagem do *spread* entre pares, trabalhos relevantes incluem (ELLIOTT; HOEK; MALCOLM, 2005) que sugere um processo Ornstein-Uhlenbeck, depois estendido por (TRIANTAFYLLOPOULOS; MONTANA, 2011), no qual se introduzem coeficientes dependentes do tempo. Ou (BERTRAM, 2009), onde as séries seguem processos de difusão de Itô.

No contexto brasileiro, (CALDEIRA; MOURA, 2013) procura através de testes de cointegração candidatos de pares de ações de entre todas as possíveis combinações de pares disponíveis no mercado local. Já em (PUCCIARELLI, 2014), a escolha de pares foi feita selecionando ativos de mesma empresa ou índices e ETFs (*Exchange Traded Funds*, em inglês). Os pares são validados utilizando o modelo dinâmico de reversão à média de (TRIANTAFYLLOPOULOS; MONTANA, 2011) e uma alteração no modelo é proposta para incluir novos parâmetros de cointegração entre os preços dos ativos.

Alguns trabalhos recentes estudam as propriedades de um número grande de séries com tendências comuns, como (BAI; NG, 2002), (BAI; NG, 2013) e (AMENGUAL; WATSON, 2006). Nesses trabalhos, as tendências em um conjunto de séries são estudadas por DFM (*Dynamic Factor Models*, em inglês), em que os fatores latentes ganham dinâmica no tempo ao serem expressados em termos de seus próprios *lags*. Em particular, (BAI; NG, 2002) explica que os fatores latentes podem ser estimados por PCA e o número ótimo de fatores pode ser determinado por meio de critérios de informação no caso de grandes amostras.

A limitação de DFM é que seu comportamento não está claro quando se trabalha com séries claramente não estacionárias. A literatura citada, por exemplo, assume que tendências podem ser removidas através de componentes estáticos. (PEÑA; PONCELA, 2006) faz uma abordagem formal da metodologia de fatores dinâmicos para séries não estacionárias, introduzindo fatores integrados, e propõe métodos para estimar o número total de fatores. (BAI, 2004) também aborda séries de fatores não estacionárias no contexto de amostras grandes. No entanto, eles não tratam o assunto de cointegração.

Apesar de suas limitações, há muitas aplicações de DFM para séries financeiras. Por

exemplo, (DIEBOLD; LI, 2006) usa a técnica para prever o comportamento da taxa de juros de títulos do governo americano. (CARVALHO, 2013) faz uma adaptação para o mercado brasileiro de taxas de juros dos títulos públicos NTN-Bs. Recentemente, (FOCARDI; FABOZZI; MITOV, 2016) mostra que modelos de DFM nas bases de (PEÑA; PONCELA, 2006) proporcionam boas previsões para as séries de logpreços dos ativos pertencentes ao índice S&P, pois existe uma tendência comum entre eles. Os autores alegam que o comportamento desta tendência comum é bem capturado ao se trabalhar com séries de logpreços e não logretornos, permitindo gerar boas previsões com um número reduzido de parâmetros a serem estimados. Porém, neste contexto, a análise da cointegração dos logpreços é apenas uma ferramenta para a geração *forecasts*. As estratégias de *trading* apresentadas no trabalho não estão se aproveitando diretamente do movimento de reversão dos preços ao equilíbrio de cointegração.

Para obter estimativas do número de tendências estocásticas em um conjunto de ativos, importantes avanços foram feitos por (CHANG; MILLER; PARK, 2005) e (CHANG; JIANG; PARK, 2008). Os autores propõem uma reformulação de VECM com *lags* infinitos dentro de um modelo de espaço-estado, cuja estimação é feita através do filtro de Kalman. O principal motivo do uso do filtro é para poder estimar os preços futuros e sua variância condicional. Para criar as estratégias de *trading* propostas neste trabalho, no entanto, não é essencial a interpretação de tendências comuns, e sim obter através delas um conjunto de vetores de cointegração que viabilizem a formulação de estratégias eficientes.

Para definir se há arbitragem estatística, um teste muito utilizado é aquele proposto por (JARROW et al., 2004) e depois aprimorado por (JARROW et al., 2012). A definição de arbitragem estatística apresentada nestes trabalhos é muito utilizada porque não depende de suposições de equilíbrio no mercado. O teste depende apenas dos ganhos acumulados e da taxa livre de risco adotada. O uso do teste é comum na literatura: o já citado (ZHANG; DAI; DJARBIROV, 2014) utiliza o teste para testar suas estratégias. (MAYORDOMO; PEÑA; ROMO, 2011) aplica o teste em estratégias baseadas em séries de preços de derivativos de crédito. (GÖNCÜ, 2014) utiliza a definição de arbitragem estatística de (JARROW et al., 2004) dentro do *framework* de Black & Scholes para *Buy & Hold* de um ativo só, e depois generaliza as conclusões para todo um portfólio em (GÖNCÜ, 2015).

Com a ideia de trabalhar com séries não diferenciadas surge a proposta criar estratégias de *trading* com o uso de VECM. Este trabalho contribui na literatura da arbitragem estatística ao mostrar que este tipo de modelos pode ser utilizado no mercado brasileiro de ações para criar estratégias com bons resultados financeiros. Em uma generalização do *pairs trading*, simulações das estratégias são feitas em portfólios com mais do que dois ativos e mais de uma tendência comum.

### 3 Objetivo

O objetivo principal deste trabalho é utilizar VECM para gerar estratégias de arbitragem estatística no mercado brasileiro de ações. Quando se tem um portfólio com tendências comuns, a reversão dos preços a estas tendências pode ser modelada por um conjunto de parâmetros referidos aqui como vetores de cointegração. Estes parâmetros estimados podem ser utilizados para descobrir desvios de precificação e assim criar regras para estratégias de *trading*. As estratégias são simuladas e os ganhos são estudados para determinar a sua viabilidade e a existência de arbitragem estatística.

Para poder realizar as simulações, primeiramente precisa-se de um portfólio que tenha pelo menos uma tendência comum tendências entre todos os ativos. Considerando primeiramente todo o universo de ações mais líquidas do mercado local BMF&BOVESPA, as tendências comuns são estimadas por técnicas de PCA e ICA (explicadas na seção seguinte). Os portfólios são então selecionados considerando grupos de ações que mais são influenciados por essas tendências.

Para cada portfólio selecionado, um total de cinco estratégias de *trading* são simuladas ao longo do tempo. Duas das estratégias são propostas neste trabalho, onde os sinais de compra e venda e os pesos de alocação são definidos pelos vetores de cointegração estimados. Outras três estratégias também simuladas têm os sinais de compra e venda definidos por *forecasts* e são comuns nos mercados de ações. Os resultados são analisados em função da performance e da existência de arbitragem estatística.



## 4 Conceitos teóricos

Nesta seção apresentados os conceitos teóricos básicos relacionados à metodologia proposta e a notação que será utilizada ao longo do texto.

### 4.1 Séries de preços

Um conjunto de séries de preços denominado por  $X$  tem a forma:

$$X = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,N} \\ X_{2,1} & \ddots & & \\ \vdots & & & \vdots \\ X_{T,1} & X_{T,2} & \cdots & X_{T,N} \end{bmatrix}. \quad (4.1)$$

Cada observação deste conjunto é denominada por  $X_{t,i}$ , onde  $i$  identifica o ativo e  $t$  é o ponto no tempo.

Séries de preços são comumente transformadas em séries de logretornos, dados por

$$R_{t,i} = \log \frac{X_{t,i}}{X_{t-1,i}} = \log X_{t,i} - \log X_{t-1,i}. \quad (4.2)$$

Por definição, séries de logretornos são séries diferenciadas. Elas são em geral estacionárias e ergódicas, duas propriedades estatísticas que facilitam o uso de modelos usuais de séries temporais. Eles também não têm escala, o que é uma vantagem ao estudar grupos de séries de empresas diversas ou mesmo tipos de séries diferentes. A estacionariedade dos logretornos, porém, torna impossível analisar e comparar tendências.

### 4.2 Tendências comuns

Séries de preços de um mesmo mercado ou setor podem ter seu comportamento associado a um número reduzido de séries influentes chamadas de fatores latentes. Fatores latentes são muitas vezes associados a séries exógenas às séries de preços, como por exemplo variáveis macroeconômicas, variáveis de mercado ou fatores específicos ao setor de operação das empresas. Quando os fatores são não estacionários são criadas tendências comuns entre os ativos.

### 4.3 PCA e ICA

Para identificar fatores latentes, além da opção de procurar séries exógenas que influenciam os dados, uma alternativa é tentar extraí-los das próprias séries estudadas. PCA (*Principal Component Analysis*) e ICA (*Independent Component Analysis*) são duas técnicas muito utilizadas para identificar fatores latentes através de combinações lineares das variáveis originais. Os dois métodos usam uma representação dos dados da forma

$$X^s = F\Lambda' \quad (4.3)$$

onde  $X^s$  é a matriz  $T \times N$  de séries padronizadas (isto é, cada série original menos sua média, dividida pelo seu desvio padrão).  $F$  é a matriz  $T \times r$  de vetores de fatores latentes e  $\Lambda$  é a *loading matrix* de dimensão  $N \times r$  que expressa a relação das séries originais com os fatores. Os componentes extraídos por PCA maximizam a variância da amostra em um número menor de variáveis ortogonais por construção. Já os componentes extraídos por ICA presumem que as variáveis não são processos gaussianos e visam maximizar momentos de ordens mais altas. Uma grande vantagem de ambos os métodos é poder trabalhar com uma amostra de tamanho reduzido sem perda substancial de informação. Isto é, é possível explicar a maior parte da variância da amostra através de um número reduzido de variáveis.

### 4.4 Integração e cointegração

Uma série não estacionária é integrada se puder ser transformada em uma série estacionária através de diferenciação. Séries de preços, por exemplo, são em geral integrados de ordem um, pois precisa-se tomar apenas uma diferença para obter uma série de retornos estacionária.

Um conjunto  $X$  de séries integradas de ordem um é cointegrado se existe pelo menos uma combinação linear  $S = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_N X_N$  com todo  $\beta_i \neq 0$  que seja estacionária. O vetor  $\beta_1, \beta_2, \dots, \beta_N$  é chamado de vetor de cointegração e ele representa a reversão das séries ao equilíbrio em torno das suas tendências comuns. O número total de vetores de cointegração está diretamente ligado ao de tendências comuns: se em um conjunto de  $N$  séries existem  $r < N$  tendências comuns, então existem  $N - r$  vetores de cointegração. Por isso, uma abordagem equivalente a procurar o número de vetores de cointegração é procurar pelo número de tendências comuns.

## 4.5 Testes de estacionariedade, raiz unitária e cointegração

O teste mais comum de estacionariedade é o teste Dickey-Fuller, usualmente chamado apenas de teste de raiz unitária. Ele testa a hipótese nula de que uma série é não estacionária, contra a hipótese alternativa da série ser estacionária. Um teste alternativo é o teste KPSS, que testa a hipótese nula de que a série é não integrada (e, portanto, estacionária), versus a hipótese alternativa da série ser integrada de ordem um.

Uma série pode ser não estacionária por causa de componentes determinísticas. Neste caso, apesar de não estacionária, a série não é integrada. Uma série integrada, por outro lado, sempre será não estacionária. Por causa disto, uma boa prática é usar os testes KPSS e de raiz unitária em conjunto.

Para testar a cointegração em um conjunto grande de séries, uma opção para conjuntos com um grande número de séries é o teste Nyblom-Harvey proposto em (NYBLOM; HARVEY, 2000), que é uma generalização do já citado teste KPSS. Dado um número  $K$  de tendências comuns, as hipóteses testadas são:  $H_0$ : Há  $K_0 < K$  tendências comuns nos dados, contra  $H_1$ : Há  $K$  tendências comuns nos dados.

## 4.6 VECM

Os modelos *Vector Error Correction* (VECM) são uma adaptação dos conhecidos modelos VAR para séries cointegradas, que incluem na formulação um termo de correção  $B'X_{t-1}$ . Este termo mede a distância das séries em relação a um equilíbrio estacionário definido pelas tendências comuns. O VECM é da forma

$$\Delta X_t = AB'X_{t-1} + \sum_{j=1}^b B_j \Delta X_{t-j} + \epsilon_t, \quad (4.4)$$

onde  $X_t$  é uma matriz  $N \times 1$  que contém os valores das  $N$  séries de preços no tempo  $t$ ,  $\Delta X_{t-1}$  é uma matriz  $N \times 1$  contendo a primeira diferença das séries no tempo  $t - 1$ ,  $b$  é a quantidade de *lags* considerados e  $\epsilon_t$  é uma matriz  $N \times 1$  de erros não independentes entre si e não correlacionados serialmente. A matriz  $B$  de dimensão  $N \times q$  contém em suas colunas os vetores de cointegração, sendo  $q$  o número de vetores de cointegração incluídos no modelo. A matriz  $A$  de dimensão  $N \times q$  contém os parâmetros que indicam a velocidade com que as séries reverterem às tendências comuns. Os vetores em  $B$  podem ser utilizados para formar a combinação linear  $B'X = S$  que é estacionária.

Como as séries trabalhadas são todas integradas de ordem um, os termos envolvendo as diferenças  $\Delta X$  são estacionários. Portanto, as tendências do modelo são definidas apenas pelo termo  $AB'X_{t-1}$ .

A metodologia para formular os VECM envolve a identificação do número de vetores de cointegração (ou o número de tendências comuns) e testes de cointegração entre as séries. Os parâmetros são estimados por máxima verossimilhança.

## 4.7 AIC e BIC

O critério de informação de Akaike e o critério de informação Bayesiano (AIC e BIC, respectivamente) são medidas de seleção de modelos para uma determinada amostra. Estes critérios medem a informação que um modelo consegue capturar, favorecendo modelos com melhor ajuste e, ao mesmo tempo, penalizando aqueles que precisam de um número elevado de variáveis explicativas. Os critérios são dados por

$$\text{AIC} = -2 \log L + 2p \quad (4.5)$$

e

$$\text{BIC} = -2 \log L + p \log N, \quad (4.6)$$

onde  $p$  é o número de parâmetros estimados no modelo. O modelo com o menor AIC ou BIC (eles não necessariamente irão coincidir) é a escolha indicada por aquele critério de seleção.

## 4.8 Teste de arbitragem estatística

Para definir se uma série de retornos é resultado da existência de arbitragem estatística, pode ser utilizado o teste proposto por (JARROW et al., 2012), daqui em diante chamado de JTTW. Dada uma estratégia de *trading*, são computados os ganhos totais acumulados  $v_t$  e partir deles são calculados os incrementos  $\Delta v_t = v_t - v_{t-1}$  para um horizonte  $[t-1; t]$  fixo, trazidos a valor presente pela taxa livre de risco.

Os autores definem arbitragem estatística como uma estratégia onde os seguintes itens são conjuntamente verdadeiros:

1.  $v(0) = 0$
2.  $\lim_{t \rightarrow \infty} E^P[v(t)] > 0$
3.  $\lim_{t \rightarrow \infty} P(v(t) < 0) = 0$
4.  $\lim_{t \rightarrow \infty} \text{Var}[\Delta v(t) | v(t) < 0] = 0$ .

A estratégia deve ter custo inicial zero, esperança de retornos acumulados positiva, uma probabilidade de perda e variância ponderada no tempo convergindo a zero. Assume-se que os incrementos dos retornos podem ser expressados por um processo da forma

$$\Delta v_t = \mu t^\theta + \sigma t^\lambda z_t, \quad (4.7)$$

sendo  $z_t \sim N(0, 1)$  *i.i.d.*. Os parâmetros  $\mu$  e  $\theta$  determinam a esperança dos incrementos enquanto os parâmetros  $\sigma$  e  $\lambda$  definem a variância do processo. Nota-se que o processo é não estacionário se  $\theta \neq 0$  ou  $\lambda \neq 0$ .

Considerando este modelo, (JARROW et al., 2012) propõe duas abordagens: se for razoável supor que a esperança dos ganhos é constante ao longo do tempo, então  $\theta = 0$  e o modelo é chamado de média restrita (CM). Caso contrário,  $\theta \neq 0$  e o modelo é de média irrestrita (UM). Os autores também propõem extensões a estes modelos, nos quais não são impostas restrições de normalidade de independência aos resíduos  $z_t$ . No entanto, eles mostram também que as conclusões entre os testes são em geral parecidas.

A hipótese a ser testada é  $H_0$ : Não há arbitragem estatística. Para rejeitar  $H_0$  e concluir que há arbitragem estatística precisam ocorrer, simultaneamente:

1.  $R_1 : \mu \geq 0$
2.  $R_2 : -\lambda > 0$  or  $\theta - \lambda > 0$
3.  $R_3 : \theta - \lambda + 1/2 > 0$  and
4.  $R_4 : \theta + 1 > 0$

Se as hipóteses fossem testadas separadamente, poderia ser feito um teste-t separado para cada hipótese. No entanto, o teste deve ser feito sob uma união de hipóteses, por isso a menor das estatísticas t serve como critério de rejeição. Por isso, o teste JTTW utiliza a estatística Min-t definida por

$$\text{Min-t} = \min \left\{ t(\hat{\mu}), t(\hat{\theta} - \hat{\lambda} + 1/2), t(\hat{\theta} + 1), \max \left[ t(-\hat{\lambda}), t(\hat{\theta} - \hat{\lambda}) \right] \right\}. \quad (4.8)$$

Os parâmetros são estimados por máxima verossimilhança e o t-crítico para um determinado nível de significância  $\alpha$ , é obtido através de simulações Monte Carlo. Para realizar as simulações, os parâmetros são fixados nos valores da fronteira entre a arbitragem e não arbitragem estatística.

Durante a estimação do modelo especificado em 4.7 para as simulações realizadas, nota-se que o parâmetro  $\theta$  é na grande maioria dos casos não significativo. Para poder testar de forma direta todas as estratégias com uma mesma metodologia neste trabalho será abordada apenas a versão de média restrita (CM) do teste.

## 5 Seleção de portfolio

### 5.1 Os dados

A amostra é composta pelos preços de fechamento diários do mercado de ações local BMF&BOVESPA. O histórico foi definido levando em consideração primeiramente o tempo de amadurecimento do mercado e às incidências de crises. Grandes marcos da economia brasileira aconteceram até o ano 2000, como o plano real implantado em 1996, ou a política monetária com metas para a inflação introduzida em 1999. Assim, dados mais antigos não seriam uma amostra representativa do mercado atual. Outro marco importante é a crise de 2008. Durante este período, houve um comportamento atípico de queda de preços de ativos e aumento de volatilidade, influenciado por fatores externos.

Para a escolha dos ativos na amostra, um fator limitante deriva do fato que estratégias *longshort* para todo um portfólio precisam de um número elevado de ativos com liquidez. Se o modelo for estimado com um histórico muito longo, o número de ativos disponível na amostra seria limitado, pois são poucos os ativos que apresentam boa liquidez durante longos períodos de tempo. Além disso, o tamanho da preamostra (isto é, os dados usados apenas para estimação) fará com que o número observações restantes disponíveis para simular as estratégias de *trading* seja reduzido.

O tamanho total da amostra foi definido em nove anos: janeiro de 2007 a dezembro de 2015. As observações dos primeiros quatro anos formam a preamostra e serão usadas unicamente para o primeiro ajuste do modelo, enquanto as estratégias serão simuladas de 2010 a 2015. O período da crise de 2008, que apresenta uma queda significativa nos índices da bolsa (ver figura 1) será depois abordado em um *backtest* no capítulo 7.

A quantidade de observações e de ativos da amostra foi separada levando em consideração os seguintes critérios: 1) O ativo deve existir para negociação durante os nove anos incluídos na amostra, 2) para todo o histórico, ter pelo menos uma transação ao dia, 3) ter boa liquidez para contratos de aluguel, visto que é uma ferramenta essencial das estratégias *longshort* e 4) ter pertencido, em algum momento dos 9 anos da amostra, ao índice IBX100.

O critério 3), referente à liquidez de contratos de aluguel, foi analisado de forma julgamental devido à falta de informações sobre estes contratos. Ativos que não tiveram nenhum registro de contrato de aluguel por semanas seguidas ao longo de todo o histórico foram desconsiderados. O critério 4) é um critério adicional de liquidez, e optou-se pelo índice IBX100 em vez do IBOVESPA apenas pelo número maior de ativos, já que ambos os índices são compostos apenas por ativos líquidos. Depois de aplicar os quatro critérios,

foi determinado um universo de 76 ativos.



Figura 1 – Valores diários dos índices IBOVESPA e IBX100

## 5.2 Determinando a existência de tendências comuns

Como será explicado no capítulo 6, o funcionamento das estratégias de *trading* propostas limita o tipo de tratamento que pode ser feito aos dados. Além de não tirar diferenças, também não foi aplicado qualquer procedimento de *smoothing* pois isso alteraria a natureza da cointegração. Trabalhar com logpreços também não seria possível pois altera a relação linear dos vetores de cointegração com as variáveis originais. Por isso, logpreços foram utilizados apenas nos casos em que os vetores de cointegração não são diretamente utilizados para definir pesos de alocação, que é o caso das estratégias em que os pesos são definidos pelos *forecasts*.

Aplicando PCA em todo o histórico da amostra (2007-2015), percebe-se que há fatores claramente não estacionários (que podem definir tendências comuns), enquanto outros aparentam ser estacionários. Ver figura 2 para a representação gráfica dos oito primeiros fatores patentes extraídos por PCA. Como será mostrado pelos testes a seguir, não é necessário analisar todos os 76 fatores comuns que podem ser extraídos, já que há evidências de que o número de tendências comuns é de apenas três, sendo o restante dos fatores estacionários.

É de interesse verificar se a estrutura dos dados se mantém a mesma ao longo do tempo, já que o período abordado na amostra é longo. Para cada ano separadamente, foram extraídos os fatores latentes pelas técnicas PCA e ICA e testes univariados KPSS e de raiz unitária foram utilizados para identificar a quantidade de fatores não estacionários e assim estimar o número total de tendências comuns. O ideal para se trabalhar com apenas um tipo de parametrização é que o número de tendências comuns seja constante ao longo do tempo.

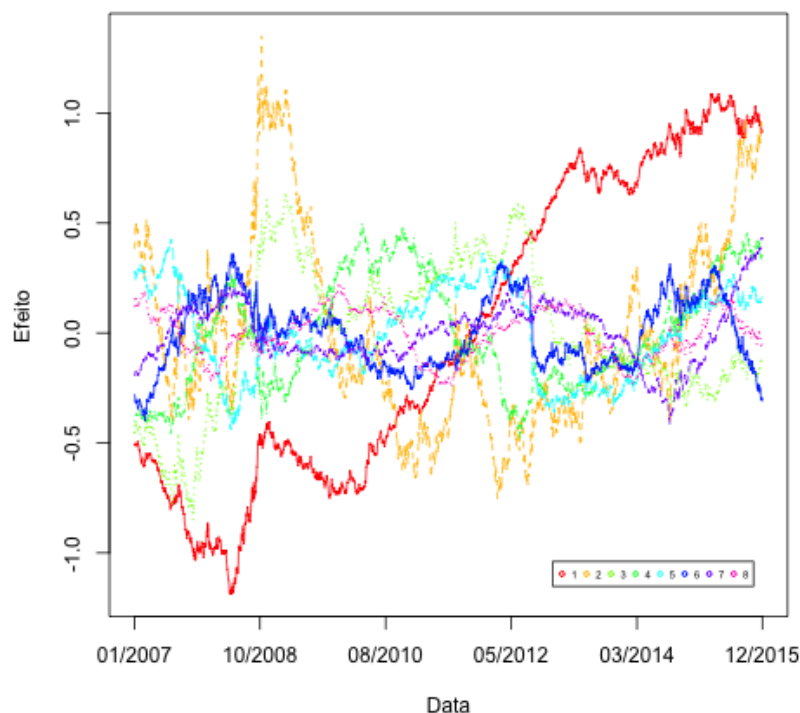


Figura 2 – Fatores latentes extraídos por PCA, utilizando todo o período da amostra (2007-2015).

O resultado do teste KPSS para cada um dos primeiros oito fatores, para intervalos de um ano está especificado na tabela 1. Observa-se que, para um nível de significância de 5%, os três primeiros fatores são não estacionários para a maioria dos anos na amostra. Utilizando o teste de raiz unitária, é possível identificar quais dos fatores não estacionários são integrados. (Lembrando que uma série integrada sempre será não estacionária, mas uma série não estacionária não é necessariamente integrada, pois a série pode ter uma tendência não estocástica.) Os resultados do teste de raiz unitária são apresentados na tabela 2. Para os primeiros oito fatores mostrados na tabela, na maioria dos anos o teste não rejeita a hipótese de raiz unitária. Para os fatores extraídos por ICA, os testes KPSS e raiz unitária apresentaram resultados semelhantes àqueles obtidos por PCA, como mostram as tabelas 3 e 4. Conclui-se que podem existir até três tendências comuns no universo estudado.

Pelas tabelas 5 e 6, vê-se que os três primeiros fatores estimados por PCA são suficientes para explicar 80% da variância para a maioria dos anos. Pelos fatores estimados por ICA, no entanto, seriam necessários seis fatores para conseguir este resultado.

Além dos testes univariados, também foi testada a existência cointegração pelo teste KPSS generalizado proposto por Niblom-Harvey. Para todos os anos, rejeitou-se a hipótese nula de que não há tendências comuns entre as séries. Aplicando novamente o



teste sob as séries diferenciadas (logretornos) a hipótese nula não é rejeitada. Conclui-se então que os dados são de fato cointegrados, e integrados de ordem um, como era esperado.

Tabela 1 – Valor-p do teste KPSS para fatores estimados por PCA, por ano.

|         | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---------|------|------|------|------|------|------|------|------|------|
| Fator 1 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Fator 2 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Fator 3 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |
| Fator 4 | 0.04 | 0.07 | 0.09 | 0.02 | 0.10 | 0.01 | 0.03 | 0.05 | 0.08 |
| Fator 5 | 0.02 | 0.10 | 0.10 | 0.10 | 0.08 | 0.10 | 0.09 | 0.09 | 0.10 |
| Fator 6 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| Fator 7 | 0.10 | 0.05 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| Fator 8 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |

Resultados que rejeitam  $H_0$  para um nível de significância de 5%, indicando que o fator é não estacionário, foram marcados em vermelho.

Tabela 2 – Valor-p do teste de raiz unitária para fatores estimados por PCA, por ano.

|         | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---------|------|------|------|------|------|------|------|------|------|
| Fator 1 | 0.79 | 0.78 | 0.02 | 0.45 | 0.60 | 0.49 | 0.80 | 0.66 | 0.68 |
| Fator 2 | 0.93 | 0.86 | 0.99 | 0.65 | 0.81 | 0.96 | 0.78 | 0.92 | 0.87 |
| Fator 3 | 0.89 | 0.61 | 0.47 | 0.96 | 0.93 | 0.41 | 0.64 | 0.22 | 0.94 |
| Fator 4 | 0.33 | 0.32 | 0.28 | 0.70 | 0.10 | 0.89 | 0.64 | 0.38 | 0.65 |
| Fator 5 | 0.76 | 0.45 | 0.33 | 0.79 | 0.70 | 0.95 | 0.34 | 0.99 | 0.13 |
| Fator 6 | 0.09 | 0.08 | 0.47 | 0.02 | 0.57 | 0.02 | 0.26 | 0.58 | 0.09 |
| Fator 7 | 0.04 | 0.78 | 0.06 | 0.06 | 0.11 | 0.09 | 0.11 | 0.06 | 0.01 |
| Fator 8 | 0.02 | 0.03 | 0.01 | 0.08 | 0.08 | 0.15 | 0.07 | 0.01 | 0.06 |

Resultados que rejeitam  $H_0$  para um nível de significância de 5%, indicando que não há raiz unitária e portanto o fator é estacionário, foram marcados em azul.

Tabela 3 – Valor-p do teste KPSS para fatores estimados por ICA, por ano.

|         | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---------|------|------|------|------|------|------|------|------|------|
| Fator 1 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Fator 2 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 |
| Fator 3 | 0.02 | 0.03 | 0.01 | 0.02 | 0.01 | 0.02 | 0.01 | 0.01 | 0.01 |
| Fator 4 | 0.04 | 0.07 | 0.09 | 0.02 | 0.10 | 0.01 | 0.03 | 0.05 | 0.08 |
| Fator 5 | 0.02 | 0.10 | 0.10 | 0.10 | 0.08 | 0.10 | 0.09 | 0.09 | 0.10 |
| Fator 6 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| Fator 7 | 0.10 | 0.05 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |
| Fator 8 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 | 0.10 |

Resultados que rejeitam  $H_0$  para um nível de significância de 5%, indicando que o fator é não estacionário, foram marcados em vermelho.

Tabela 4 – Valor-p do teste de raiz unitária para fatores estimados por ICA, por ano.

|         | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|---------|------|------|------|------|------|------|------|------|------|
| Fator 1 | 0.79 | 0.78 | 0.02 | 0.45 | 0.60 | 0.49 | 0.80 | 0.66 | 0.68 |
| Fator 2 | 0.93 | 0.86 | 0.99 | 0.65 | 0.81 | 0.96 | 0.78 | 0.92 | 0.87 |
| Fator 3 | 0.89 | 0.61 | 0.47 | 0.96 | 0.93 | 0.41 | 0.64 | 0.22 | 0.94 |
| Fator 4 | 0.33 | 0.32 | 0.28 | 0.70 | 0.10 | 0.89 | 0.64 | 0.38 | 0.65 |
| Fator 5 | 0.76 | 0.45 | 0.33 | 0.79 | 0.70 | 0.95 | 0.34 | 0.99 | 0.13 |
| Fator 6 | 0.09 | 0.08 | 0.47 | 0.02 | 0.57 | 0.02 | 0.26 | 0.58 | 0.09 |
| Fator 7 | 0.04 | 0.78 | 0.06 | 0.06 | 0.11 | 0.09 | 0.11 | 0.06 | 0.01 |
| Fator 8 | 0.02 | 0.03 | 0.01 | 0.08 | 0.08 | 0.15 | 0.07 | 0.01 | 0.06 |

Resultados que rejeitam  $H_0$  para um nível de significância de 5%, indicando que não há raiz unitária e portanto o fator é estacionário, foram marcados em azul.

Tabela 5 – Variância explicada acumulada dos fatores estimados por PCA, por ano.

|         | 2007        | 2008        | 2009        | 2010        | 2011        | 2012        | 2013        | 2014        | 2015        |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Fator 1 | <b>0.59</b> | <b>0.72</b> | <b>0.80</b> | <b>0.49</b> | <b>0.54</b> | <b>0.46</b> | <b>0.38</b> | <b>0.42</b> | <b>0.55</b> |
| Fator 2 | <b>0.79</b> | <b>0.84</b> | 0.85        | <b>0.71</b> | <b>0.73</b> | <b>0.72</b> | <b>0.62</b> | <b>0.71</b> | <b>0.74</b> |
| Fator 3 | <b>0.84</b> | 0.88        | 0.89        | <b>0.79</b> | <b>0.81</b> | <b>0.81</b> | <b>0.76</b> | <b>0.81</b> | <b>0.83</b> |
| Fator 4 | 0.87        | 0.91        | 0.92        | <b>0.85</b> | 0.86        | 0.87        | <b>0.82</b> | 0.85        | 0.87        |
| Fator 5 | 0.90        | 0.92        | 0.94        | 0.88        | 0.88        | 0.90        | 0.86        | 0.88        | 0.89        |
| Fator 6 | 0.91        | 0.94        | 0.96        | 0.91        | 0.90        | 0.92        | 0.89        | 0.90        | 0.91        |
| Fator 7 | 0.93        | 0.94        | 0.96        | 0.92        | 0.92        | 0.93        | 0.90        | 0.91        | 0.93        |
| Fator 8 | 0.94        | 0.95        | 0.97        | 0.93        | 0.93        | 0.94        | 0.92        | 0.93        | 0.94        |

Os fatores necessários para representar pelo menos 80% da variância foram marcados em negrito.

Tabela 6 – Variância explicada acumulada dos fatores estimados por ICA, por ano.

|         | 2007        | 2008        | 2009        | 2010        | 2011        | 2012        | 2013        | 2014        | 2015        |
|---------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| Fator 1 | <b>0.29</b> | <b>0.30</b> | <b>0.42</b> | <b>0.17</b> | <b>0.21</b> | <b>0.29</b> | <b>0.16</b> | <b>0.20</b> | <b>0.44</b> |
| Fator 2 | <b>0.49</b> | <b>0.57</b> | <b>0.71</b> | <b>0.33</b> | <b>0.36</b> | <b>0.46</b> | <b>0.31</b> | <b>0.37</b> | <b>0.55</b> |
| Fator 3 | <b>0.61</b> | <b>0.69</b> | <b>0.81</b> | <b>0.48</b> | <b>0.48</b> | <b>0.60</b> | <b>0.45</b> | <b>0.53</b> | <b>0.64</b> |
| Fator 4 | <b>0.72</b> | <b>0.74</b> | 0.86        | <b>0.60</b> | <b>0.60</b> | <b>0.70</b> | <b>0.58</b> | <b>0.64</b> | <b>0.73</b> |
| Fator 5 | <b>0.78</b> | <b>0.80</b> | 0.90        | <b>0.70</b> | <b>0.70</b> | <b>0.80</b> | <b>0.69</b> | <b>0.72</b> | <b>0.80</b> |
| Fator 6 | <b>0.84</b> | 0.86        | 0.93        | <b>0.79</b> | <b>0.79</b> | 0.88        | <b>0.79</b> | <b>0.80</b> | 0.86        |
| Fator 7 | 0.89        | 0.91        | 0.95        | <b>0.86</b> | <b>0.87</b> | 0.91        | <b>0.88</b> | 0.87        | 0.91        |
| Fator 8 | 0.94        | 0.95        | 0.97        | 0.93        | 0.93        | 0.94        | 0.92        | 0.93        | 0.94        |

Os fatores necessários para representar pelo menos 80% da variância foram marcados em negrito.

### 5.3 Seleção de portfólios candidatos

O fato dos testes terem detectado tendências comuns e cointegração não significa que todos os ativos são influenciados por todas estas tendências ao mesmo tempo. Se fosse assim, empresas de diferentes setores e tamanhos seriam influenciados pelas mesmas fontes, e na prática não é isto o que acontece. Surge assim a necessidade de um método de identificação de grupos de ativos dentro do qual as tendências sejam comuns a todos os ativos do grupo. Uma forma de fazer isto é analisar os *factor loadings*,  $\Lambda$ , estimados pela

técnica de PCA, que apresentaram resultados mais consistentes ao longo dos anos do que as estimativas feitas por ICA.

Ao trabalhar com uma amostra padronizada, os *factor loadings* refletem a influência de cada ativo na composição das tendências. Assim, é razoável supor que ativos com maiores pesos  $\Lambda$  são os mais influenciados pelos verdadeiros fatores latentes. Com o intuito de não utilizar informação futura, a matriz  $\Lambda$  estimada na preamostra foi usada como critério de seleção para os portfólios (a) e (b) descritos a seguir. Para fins de referência, também foram selecionados os portfólios (c) e (d) muito utilizados em estratégias de *pairs trading*:

- (a) Portfólio com um fator latente

Para montar este portfólio, foram selecionados os  $N$  ativos com os maiores elementos de  $\Lambda$  correspondentes ao primeiro fator. O objetivo deste portfólio é analisar se é mais vantajoso para estratégias de *trading* utilizar uma amostra reduzida, cuja influência depende majoritariamente de apenas um fator, ou se seria mais vantajoso ampliar a amostra para estimar o efeito de todas as tendências comuns. Foram montados portfólios com  $N = 10, 20$  e  $30$  ativos. Os ativos selecionados foram, na ordem dos pesos dados por  $\Lambda$ : RAPT4, ITUB3, BBAS3, BBDC4, BBDC3, SLED4, LAME4, BRAP4, CYRE3, RENT3, ITUB4, VALE5, LREN3, VALE3, BRKM5, ETER3, POMO4, ITSA4, VLID3, RSID3, CCRO3, TBLE3, PSSA3, ODPV3, SBSP3, MMXM3, GFSA3, CSMG3, KLBN4, GRND3. O mais chamativo deste grupo de ativos, é a diversidade dos setores das empresas. Um detalhe importante é que em um princípio, foram testados também portfólios estruturados pelos subsequentes fatores estimados por PCA. No entanto, eles não tiveram bom desempenho em nenhuma estratégia testada. Percebe-se que selecionar um grupo de ativos ignorando o primeiro fator é desvantajoso já que ele é o mais influente em toda a amostra.

- (b) Portfólio com  $r$  fatores latentes

Este portfólio é composto pelos ativos correspondentes aos 10 maiores elementos de  $\Lambda$  relacionados aos  $r = 3$  fatores latentes especificados na seção anterior. O portfólio é composto pelos seguintes 30 ativos: RAPT4, BBAS3, ITUB3, BBDC4, BBDC3, LAME4, SLED4, BRAP4, CYRE3, RENT3, UNIP6, BTOW3, PCAR4, EMBR3, VAGR3, VIVT4, VIVT3, NATU3, TOTS3, TRPL4, PETR3, PETR4, CSNA3, CARD3, OIBR4, GOAU4, GGBR4, ELPL4, USIM5 e GGBR3. Adiantando alguns dos resultados das conclusões, o portfólio de um fator apresentou os melhores resultados quando  $N = 10$ . Portanto, para montar este portfólio foram escolhidos também 10 ativos por fator considerado.

- (c) Portfólio com ativos da mesma empresa

Portfólio contendo dois ativos da mesma empresa, PETR3 e PETR4. Estes ativos

são conhecidamente cointegrados, e são o par mais usualmente selecionado para *pairs trading* no mercado brasileiro.

- (d) Portfólio com ativos do mesmo setor

Portfólio contendo ativos do mesmo setor de atuação. Por causa da boa liquidez, foram selecionados os ativos do setor bancário BBAS3, BBDC3, BBDC4, ITUB4 e ITUB3. Utilizar ativos do mesmo setor é uma outra prática usual para identificar ativos cointegrados, visto que é razoável supor que haverá tendências comuns entre eles.

Como mostra a figura 3, os portfólios separados têm uma composição muito diferente. Aqueles escolhidos através de PCA apresentam a maior diferença de escalas entre preços e as tendências comuns são menos evidentes.

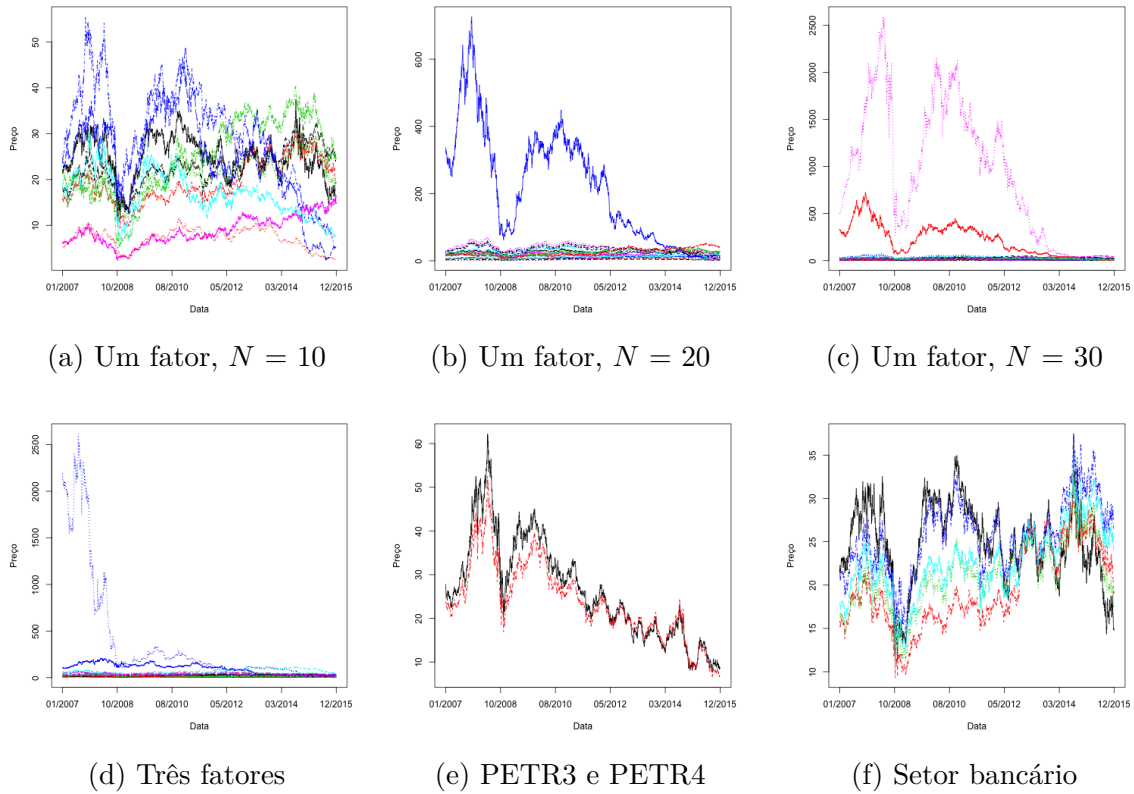


Figura 3 – Conjunto de séries  $X_t$  no período 2007-2015 para diferentes tipos de seleção de portfólio

## 6 As estratégias de *trading*

A seguir são apresentados 5 métodos para geração de sinais e determinação dos pesos para estratégias *longshort*. As posições deste tipo de estratégia podem ser compradas (*long*) ou vendidas a descoberto (*short*, indicadas por pesos negativos). Quando uma estratégia assume posições em *short*, os ativos são alugados e vendidos em seguida. Não é abordado o uso de derivativos.

Estratégias *longshort* são muitas vezes calibradas para serem *self-financing*. Isto é possível ao usar venda a descoberto como a única fonte de recursos para assumir as posições compradas. Como será explicado nas seções seguintes, isto não sempre será possível.

### 6.1 Proposição de estratégias definidas em função dos vetores de cointegração

As estratégias aqui sugeridas exploram o fato que o conjunto de séries  $B'X = S$  é estacionário. Quanto mais distante  $S_t$  estiver da sua média histórica, maior a reversão que deverá acontecer nas próximas observações. Já que os vetores de cointegração são assumidos constantes, um valor extremo de  $S_t$  é um indício de que há desvios na precificação de um ou mais ativos naquele momento. Este processo de reversão de  $S_t$  permite prever o movimento de um portfólio de ativos como um todo, sem de fato precisar fazer *forecasts* das séries.

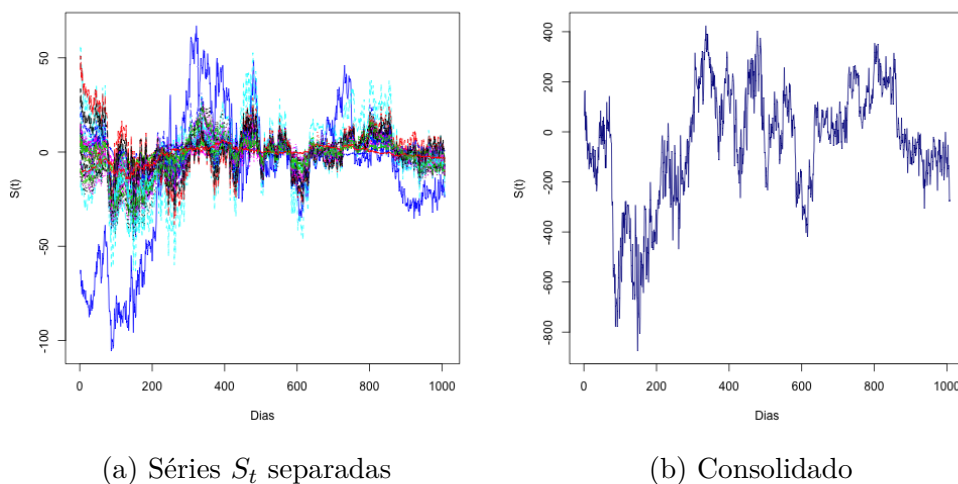


Figura 4 –  $S_t$  estimadas na preamostra, considerando o VECM com 1 *lag* e 27 vetores de cointegração.

As estratégias (1) e (2) descritas a seguir baseiam-se nos valores de  $S_t$  a cada nova observação para tomar as decisões de *enter* e *exit* no *trading*. No caso de modelos com apenas um vetor de cointegração, a existência de falhas de precificação é determinada diretamente pelos valores extremos de  $S_t$ . Se o valor de  $S_t$  estiver abaixo da média histórica, então os pesos dos ativos no portfólio devem ser proporcionais ao vetor de cointegração. Da mesma forma, se o valor de  $S_t$  estiver acima da sua média histórica, os pesos devem ser contrários ao vetor de cointegração. Desta forma, o portfólio irá imitar o movimento de reversão observado em  $S_t$ . Se houver mais de um vetor de cointegração estimado, eles são somados para criar um novo vetor  $B_{cons} = \sum_{i=1}^q B_i$  e a lógica da estratégia continuará a mesma. Como qualquer combinação linear dos vetores de cointegração é vetor de cointegração também, sabe-se que a nova série  $S = B'_{cons}X$  será estacionária. Este comportamento de  $S_t$  para a preamostra do portfólio de 3 fatores é ilustrado na figura 4.

Por utilizar os vetores de cointegração diretamente para definir os pesos do portfólio, quando estas estratégias forem ser utilizadas, o modelo não pode ser estimado utilizando logpreços. Apesar de séries de logpreços serem mais suaves do que as séries de preços propriamente ditas, a natureza linear do equilíbrio estimado por  $B$  se perderia e isso impediria definir os pesos corretamente.

Outra particularidade desta metodologia é que o total das posições em *long* não necessariamente irá coincidir com o total das posições em *short*. Isto é uma característica relevante, pois a estratégia não tem garantia de ser *self-financing*. Para contornar este problema foi imposto um controle sobre o volume negociado: os pesos são sempre ajustados para que o total comprado seja igual a um capital predeterminado (desconsiderando os custos de transação). Isto é, dado o capital predeterminado  $C$ , os pesos  $w_{t,i}$  são dados por  $w_{t,i} = \frac{B_{t,i}}{B^+} \times C$  onde  $B^+$  é a soma dos valores positivos em  $B_{t,1}, \dots, B_{t,N}$ . O procedimento de reescalar os pesos é análogo para quando os pesos forem contrários aos vetores de cointegração. Os recursos para a compra, fixada sempre no montante  $C$ , e para custos de transação adicionados posteriormente são obtidos pela venda a descoberto e, se mais recursos forem necessários, serão tomados da taxa livre de risco.

Os detalhes das estratégias são dados a seguir.

### 1) Estratégia de dois pontos

Para cada observação  $t$ , estimar o modelo VECM, obter o vetor  $B_{cons}$  e calcular a série estacionária  $S_t$ . Dada uma porcentagem  $k$ , calcular o  $k\%$  percentil histórico de  $S$ ,  $S_{k\%}$ . Lembrando que um ativo com peso positivo indica uma posição comprada (*long*) e peso negativo indica uma posição vendida (*short*), as regras de *trading* são dadas por:

- Se  $S_t < S_{k\%}$ : *Exit* e *Enter*

Desfaz-se a posição anterior, recolhendo os ganhos ou perdas. Toma-se nova posição com pesos proporcionais a  $B_{cons}$ .

- Se  $S_t > S_{(100-k)\%}$ : *Exit* e *Enter*

Desfaz-se a posição anterior, recolhendo os ganhos ou perdas. Toma-se nova posição com pesos contrários a  $B_{cons}$ .

- Se nenhum dos anteriores, manter a posição atual, se houver.

## 2) Estratégia de três pontos

A diferença desta estratégia com a anterior é que as posições são desfeitas no momento em que  $S_t$  atinge novamente sua mediana histórica. Mais especificamente, a posição é desfeita quando  $S_t$  estiver dentro no intervalo  $[S_{(50-\frac{k}{2})\%}; S_{(50+\frac{k}{2})\%}]$ . Neste caso, as regras são dadas por:

- Se  $S_t < S_{k\%}$ : *Enter*

Toma-se posição com pesos proporcionais a  $B_{cons}$ . Se já havia posição tomada, manter a posição anterior.

- Se  $S_t \in [S_{(50-\frac{k}{2})\%}; S_{(50+\frac{k}{2})\%}]$ : *Exit*

Desfaz-se a posição anterior, recolhendo os ganhos ou perdas.

- Se  $S_t > S_{(100-k)\%}$ : *Enter*

Toma-se posição com pesos proporcionais ao contrário de  $B_{cons}$ . Se já havia posição tomada, manter a posição anterior.

- Se nenhum dos anteriores, manter a posição atual, se houver.

A velocidade de reversão à média de  $S_t$  irá definir a quantidade de transações a serem feitas em cada estratégia. Se  $S_t$  tiver uma reversão lenta, trocar a posição no meio fará com que os ganhos, ainda que menores, sejam recolhidos antes. Esta estratégia possibilita não tomar nenhuma posição enquanto não forem observados valores extremos de  $S_t$ , o que permite não adquirir risco enquanto erros de precificação não forem detectados.

## 6.2 Estratégias definidas em função de *forecasts*

As duas estratégias descritas acima foram comparadas com 3 estratégias usuais baseadas em *forecasts*, na mesma linha daquelas propostas por (FOCARDI; FABOZZI; MITOV, 2016). Nesta abordagem, as estimativas e atualizações nas posições são feitas a cada horizonte de *forecast*  $hf$ . Para cada atualização, novos *forecasts* são gerados para o próximo horizonte  $[t, t + hf]$  e as posições são sempre definidas com base nos retornos esperados  $R_{t+hf,i} = (X_{t+hf,i} - X_{t,i})/X_{t,i}$ . Como neste tipo de abordagem os pesos não dependem diretamente dos vetores de cointegração, a estimação do modelo foi feita sob

logpreços e os *forecasts* obtidos foram devidamente transformados. Além disso, os pesos sempre são controlados para que a estratégia seja *self-financing*.

As estratégias são descritas a seguir.

### 3) Estratégia de pesos iguais (*Dummy*)

Dada uma porcentagem limite  $l$ , selecionar os  $l\%$  ativos com maiores retornos esperados e alocar posições compradas com pesos iguais. Selecionar os  $l\%$  ativos com menores retornos esperados e alocar posições vendidas com pesos iguais.

### 4) Otimização *Max Sharpe*

Os pesos de alocação devem ser definidos de forma a maximizar a estatística *Sharpe Ratio* dos retornos esperados. A estatística é dada por  $Sharpe Ratio = \mu/\theta$ , onde  $\mu$  e  $\theta$  são estimados pela média e desvio padrão dos retornos esperados, respectivamente. Consideram-se os  $l\%$  ativos com maiores retornos para as posições em *long* e os  $l\%$  ativos com menores retornos para as posições em *short*.

### 5) Estratégia de pesos proporcionais ao retorno esperado: (*Benchmark*)

Alocação é feita com pesos proporcionais à diferença dos retornos esperados em relação à média dos retornos esperados no portfólio. Isto é, os pesos são proporcionais às diferenças

$$R_{t,i} - \frac{1}{N} \sum_{i=1}^N R_{t,i}.$$

## 6.3 Implementação

As simulações descritas a seguir foram feitas através de um código na linguagem **R**.

Foram descritas até aqui 5 diferentes metodologias de *trading* para serem simuladas em 6 tipos de portfólios diferentes. Para todas as simulações, foi fixado um volume de negociação de R\$1 milhão, que determina o capital máximo a ser tomado nas posições *long*. Assume-se que não há reinvestimento dos retornos: a cada iteração, os ganhos são recolhidos e investidos na taxa livre de risco. Caso houver prejuízo, a quantidade equivalente será tomada emprestada também pela taxa livre de risco.

Para a estimação dos modelos, foi utilizado 1 *lag*, como sugerido pelos AIC e BIC na preamostra. A quantidade de vetores de cointegração que devem ser incluídos no modelo poderia ser definida também por critérios de informação, mas o número ótimo para o modelo não necessariamente será o número ótimo de vetores a ser considerado para *trading*. Portanto, todas as estratégias serão testadas diferentes quantidades de vetores de cointegração, sendo que o máximo é sempre  $q = N - r$  vetores. Também foram testados diferentes valores para o horizonte de *forecast*  $hf$  e para os parâmetros  $k$  e  $l$  que guiam as decisões do *trading*. Os parâmetros considerados para todas as estratégias podem ser observados na tabela 7.



Tabela 7 – Combinação de parâmetros para cada estratégia

| Nr. | Descrição da estratégia  | Parâmetros testados          |
|-----|--|------------------------------|
| 1   | Posições são tomadas e desfeitas nos extremos de $S_t$ . (2 pontos)                                    | $k = 10, 30, 50$             |
| 2   | Posições são tomadas nos extremos de $S_t$ e desfeitas no seu centro. (3 pontos)                       | $k = 10, 20, 30$             |
| 3   | Pesos iguais para os 1% ativos com maiores e menores retornos. ( <i>Dummy</i> )                        | $l = 30, 50, hf = 1, 10, 30$ |
| 4   | Otimização <i>Max Sharpe</i> para os 1% ativos com maiores e menores retornos. ( <i>Max Sharpe</i> )   | $l = 30, 50, hf = 1, 10, 30$ |
| 5   | Pesos proporcionais às diferenças $R_{t,i} - \frac{1}{N} \sum_{i=1}^N R_{t,i}$<br>( <i>Benchmark</i> ) | $hf = 1, 10, 30$             |

### 6.3.1 Custos de operação e a taxa livre de risco

No caso de estruturas *longshort* com transações possivelmente diárias, há gastos elevados de compra e aluguel de ativos, de modo que os custos de operação não poderiam ser considerados irrisórios. Para simular corretamente o comportamento das estratégias no mercado, foram incluídos nos cálculos os custos aproximados para um investidor pessoa jurídica.

O preço de aluguel varia de acordo com o ativo e a data da negociação. Para transações feitas a partir de setembro de 2014, foi utilizado o preço real de aluguel de um dia para aquele ativo naquele momento. Para transações anteriores, como não há dados disponíveis sobre os preços de aluguel, foi utilizada uma média por ativo. Os custos referentes à taxa anualizada de aluguel  $a$  são calculados diariamente por:

$$Aluguel = Volume\ negociado * \left[ (1 + a)^{1/252} - 1 \right].$$

Uma suposição relevante é que o contrato de aluguel nunca pode ser terminado pelo doador, e que este sempre estará disposto a estender o contrato de aluguel a critério do tomador. Além da taxa de aluguel paga ao doador, paga-se a taxa de liquidação para a BMF&BOVESPA. Considera-se que o empréstimo é feito de forma automática, de modo que seu custo de liquidação dos aluguéis para cada dia cobrado pela BMF&BOVESPA é dado por:

$$Liq.\ fin.\ do\ empréstimo = Volume\ negociado * \left( 1,005^{1/252} - 1 \right).$$

Os emolumentos e liquidação da compra de ações foram fixados em 0,025% por transação. A corretagem operacional foi fixada em 0,5% do volume da operação, com devolução de 80%. Assim, a taxa de corretagem final resulta em  $(0,5 * 0,8)\% = 0,1\%$ . A taxa de custódia é desconsiderada, pois ela não está diretamente relacionada à operação, nem foram considerados descontos referentes a impostos.

O custo do capital emprestado é a taxa livre de risco. No mercado brasileiro, o

conceito de taxa livre de risco é objeto de debate, pois não há ativos no mercado nacional que apresentem zero risco de inadimplência e ao mesmo tempo zero risco de reinvestimento. Aqui será utilizada a aproximação usual da taxa livre de risco para grandes investidores, que é 100% do CDI. É importante salientar que o CDI possui volatilidade alta como taxa livre de risco, diferente das taxas utilizadas por (JARROW et al., 2012) ou (FOCARDI; FABOZZI; MITOV, 2016), por exemplo.

## 7 Resultados

Neste capítulo são apresentados os resultados das simulações das estratégias de *trading* apresentadas, para todos os portfólios selecionados. São em total cinco estratégias aplicadas em seis portfólios: os portfólios de um fator latente com  $N = 10, 20$  e  $30$ , portfólio de  $r = 3$  fatores latentes, portfólio composto por PETR3 e PETR4, e o portfólio composto por ativos do setor bancário. As operações de *trading* são simuladas no período de cinco anos de janeiro de 2010 a dezembro de 2015, sempre utilizando um histórico de 4 anos para estimação do modelo.

O modelo utilizado para estimação de  $S_t$  e dos *forecasts* em todos os casos foi VECM com um *lag*, como sugerido pelos critérios AIC e BIC na preamostra. Para as estratégias (1) e (2), cujos pesos são definidos pelos vetores de cointegração, o modelo foi estimado utilizando preços, enquanto para as estratégias (3), (4) e (5), o modelo foi estimado utilizando logpreços e os *forecast* foram depois devidamente transformados.

O maior interesse dos resultados a seguir é a quantidade de vetores de cointegração utilizados nos diferentes portfólios. Gráficos com resultados detalhados para os outros parâmetros (limites de *enter* e *exit*, e horizonte de *forecast*) assim como tabelas com as principais medidas resumo de todas as estratégias estão disponíveis no apêndice A, e medidas resumo das estratégias estão disponíveis no apêndice B.

### 7.1 Portfólio de um fator latente

Os resultados para os três tamanhos de portfólio podem ser vistos na figura 5. Percebe-se que o uso de um grupo menor de ativos ( $N = 10$ , representados no gráfico pela cor vermelha) é o mais adequado para todas as estratégias: em geral, as simulações com maiores ganhos vêm dos portfólios com este número de ativos, e mesmo quando um portfólio com  $N = 10$  resulta em prejuízo, é por causa da escolha inapropriada dos parâmetros  $k$ ,  $l$  ou  $hf$ . A escolha equivocada dos parâmetros  $k$  nas estratégias (1) e (2), por exemplo, pode levar a uma quantidade muito reduzida de transações. Simulações feitas com  $k = 10\%$  resultam em apenas 10 ou 11 transações em todo o período de cinco anos da simulação. (Ver apêndice B.) O comportamento de um grupo maior de ativos não é explicado de forma satisfatória por apenas um fator latente, portanto este tipo de seleção cria um grupo em que as influências de todas as tendências comuns não estão sendo propriamente representadas.

Na figura 6 são detalhados os resultados pela quantidade de vetores de cointegração de forma separada para o portfólio com  $N = 10$ . Para as estratégias (1) e (2), não é

necessário utilizar mais do que uma série  $S_t$  para conseguir ganhos relevantes. É recolhido em ganhos mais da metade do volume máximo negociado (R\$ 1 milhão) em cinco anos de simulação. Por causa do portfólio ser dominado por apenas uma tendência comum, o processo de reversão capturado por apenas um vetor de cointegração parece ser suficiente. Além disso, ao ser considerado um número baixo de séries, a estimação dos vetores de cointegração adicionais pode não ser precisa o bastante.

Por outro lado, os resultados em relação aos vetores de cointegração são o oposto para as estratégias baseadas em *forecasts*. Aquelas em que os retornos esperados são obtidos pelo VECM com  $N - r$  vetores de integração tiveram os maiores ganhos, enquanto o uso de apenas um vetor de cointegração trouxe principalmente prejuízos. Em particular, nota-se que estratégia (4) (*Max-Sharpe*) têm perdas constantes para o horizonte de um dia, já que transações diárias não são adequadas para uma otimização que depende da variância e demora para refletir as alterações nos preços.

## 7.2 Portfólio de três fatores latentes

Dentre todos os portfólios selecionados, este é aquele que concentra a maior quantidade de informação. O portfólio é composto por um número grande de ativos ( $N = 30$ ), e sua escolha foi baseada em todas as tendências comuns encontradas na amostra. Ao contrário de técnicas usuais de *pairs trading* que procuram isolar os efeitos da cointegração em um grupo pequeno de ativos, com este portfólio a intenção é conseguir uma boa representação das principais tendências do mercado.

Os resultados detalhados pela quantidade de vetores de cointegração considerados são mostrados na figura 7. Para este tipo de portfólio, os melhores resultados estão claramente nas estratégias (1) e (2) utilizando a maior quantidade de vetores de cointegração. Este resultado não é uma surpresa já que o portfólio foi montado de forma a refletir as três tendências comuns. As estratégias (1) e (2) aplicadas neste portfólio utilizam a maior quantidade de informação disponível dentre todas as metodologias testadas. É possível que isto as torne as estratégias mais consistentes ao longo de toda a simulação. Assim como no caso do portfólio de um fator, as simulações com os piores resultados são aquelas com parâmetros  $k$  baixos e consequentemente, com baixo número de transações.

O uso de um número maior de vetores de cointegração permite capturar melhor as oportunidades de arbitragem. A estratégia (2) por exemplo, apresenta baixa volatilidade e ganhos praticamente contínuos quando  $k = 50\%$  e a quantidade de vetores de cointegração é  $N - r = 27$ . Com quase duzentas transações em cinco anos, a estratégia realiza aproximadamente uma transação a cada sete dias úteis, gerando mais ganhos com menos transações do que se apenas um vetor de cointegração fosse utilizado.

Comparando os resultados deste portfólio com os do portfólio de um fator, há

diferenças evidentes. Como as estratégias (1) e (2) são altamente dependentes da estimação dos vetores de cointegração, elas apresentaram uma performance muito superior neste portfólio. As estratégias de *forecasts*, por outro lado, são muito mais voláteis aqui do que no portfólio de um fator e performam melhor quando apenas um vetor de cointegração é utilizado. Uma possível explicação é que, como séries de preços são muito imprevisíveis, aumentar o tamanho do portfólio não é uma vantagem para este tipo de estratégia, e que fazer *forecasts* utilizando um modelo mais simples acaba sendo mais eficiente.

### 7.3 Portfólios de ativos da mesma empresa

Os ativos PETR3 e PETR4 são conhecidos por serem cointegrados. Porém, um portfólio contendo apenas este par é particularmente complicado porque a empresa é muito influenciada por diversos fatores externos dificilmente previsíveis, como o cenário político. Os ativos estão em queda durante praticamente todo o histórico da amostra, diferentemente do que os primeiros fatores latentes estimados para toda a amostra indicam (as tendências explicadas pelos fatores foram mostradas na figura 2).

As estratégias (3), (4) e (5) apresentam prejuízos que indicam pouca capacidade de previsão. As estratégias (1) e (2), que não dependem de *forecasts*, realizam de 2 a 27 transações no total dos cinco anos de simulação dependendo dos parâmetros escolhidos. Durante o período de queda dos ganhos particularmente, elas não realizam nenhuma transação. Portanto, apesar da recuperação dos ganhos no final do histórico, a simulação produz resultados que não podem ser inferidos facilmente.

Como este par de ativos é extremamente líquido, é mais difícil que aconteçam erros de precificação detectáveis em uma frequência diária. Além disso, não há nenhum mecanismo de *stop-loss* implementado, o que impede de uma posição tomada ser desfeita fora das ordens da estratégia mesmo se houver indícios de que a relação de cointegração é diferente da estimada quando a posição foi tomada.

### 7.4 Portfólios de ativos do mesmo setor

Os resultados para o portfólio de empresas do setor bancário são apresentados na figura 9. Observa-se que este portfólio teve uma performance similar em todas as estratégias ao portfólio de um fator latente. O maior ponto de atenção é que a melhor performance obtêm-se através do modelo com apenas um vetor de cointegração. É a segunda evidência de que, para amostras menores e com uma tendência comum que se sobressai, compensa centralizar o *trading* em volta de um vetor de cointegração apenas.

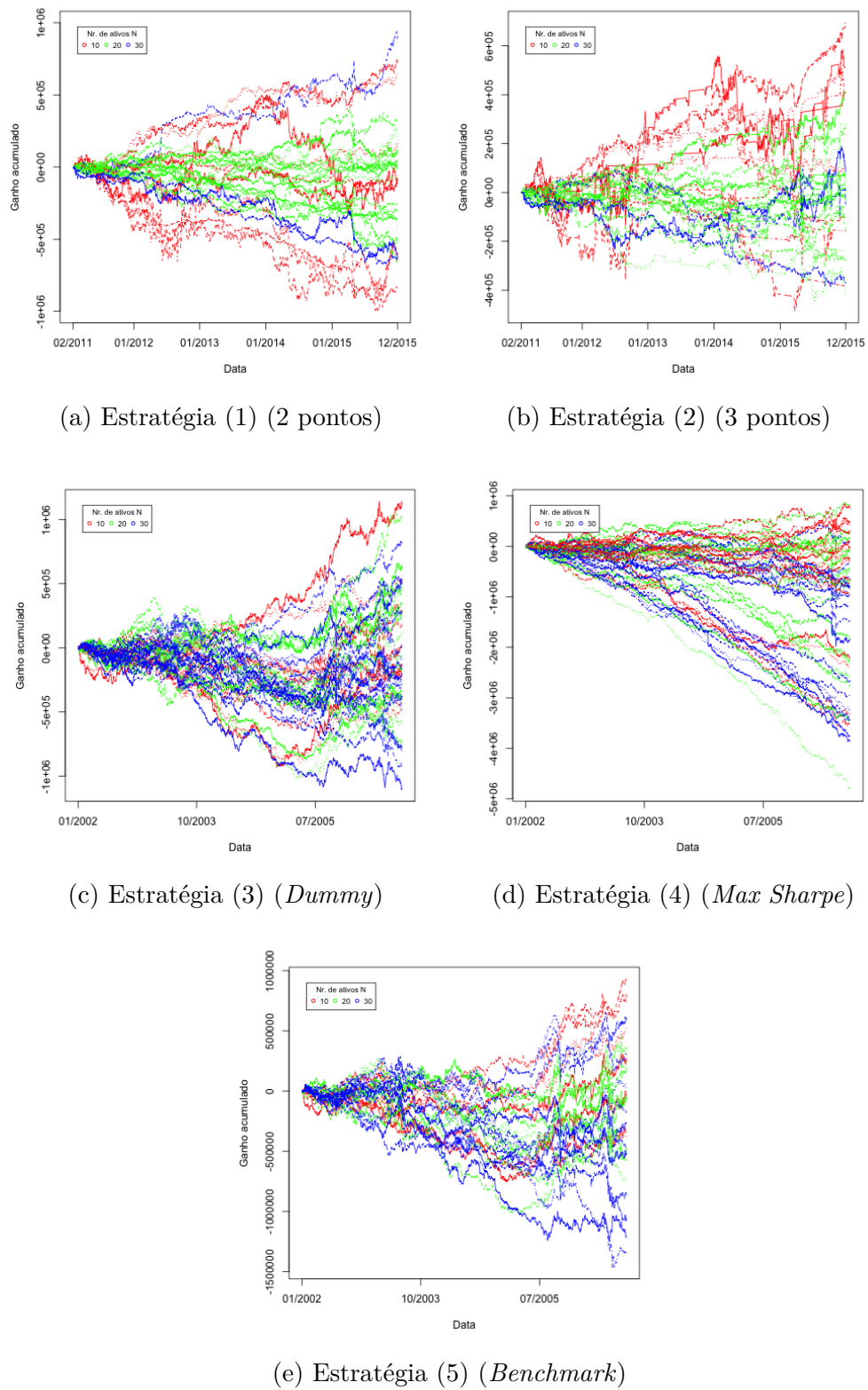


Figura 5 – Ganhos para os portfólios de um fator, detalhados pelo número de ativos que compõem os portfólios.

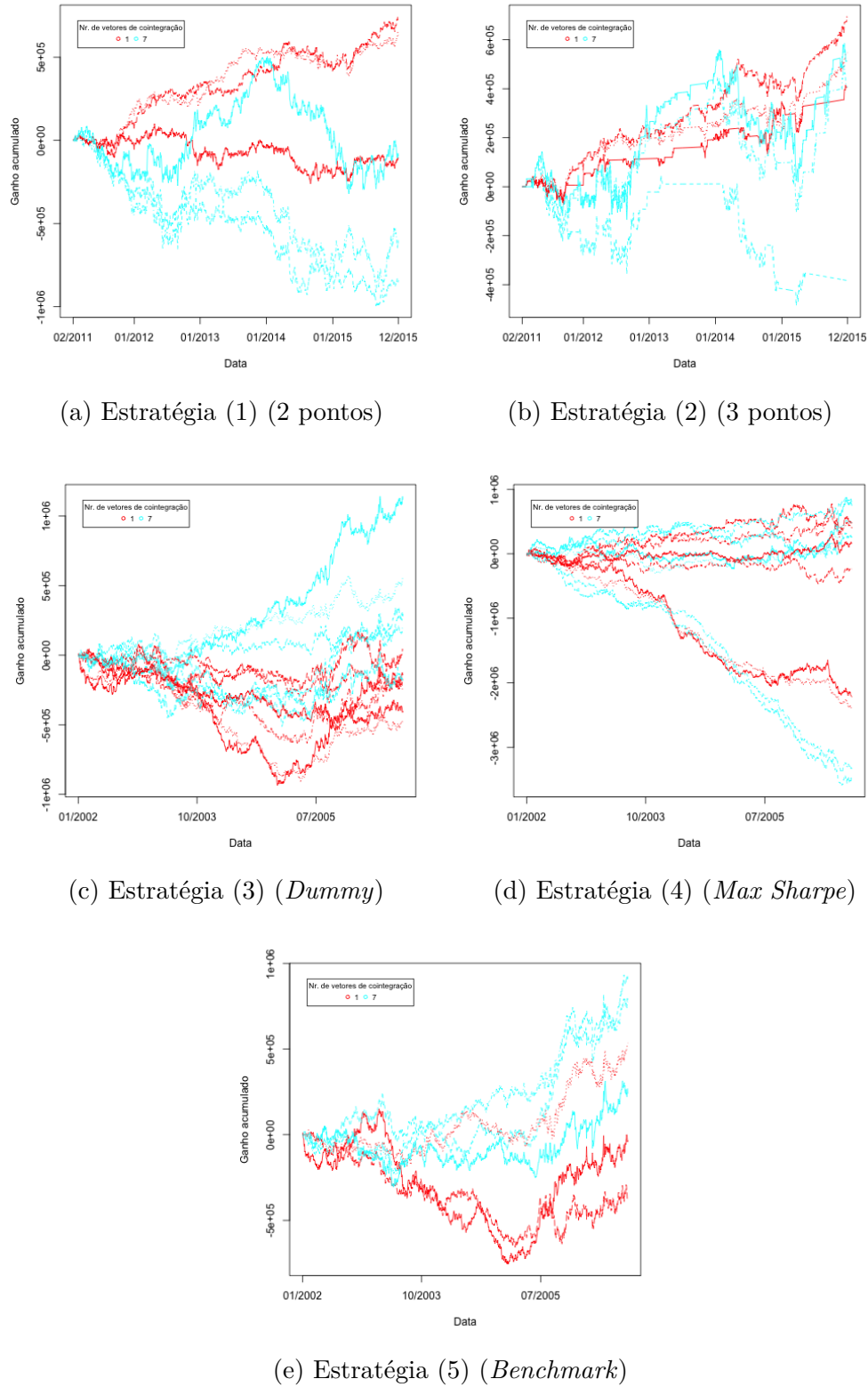
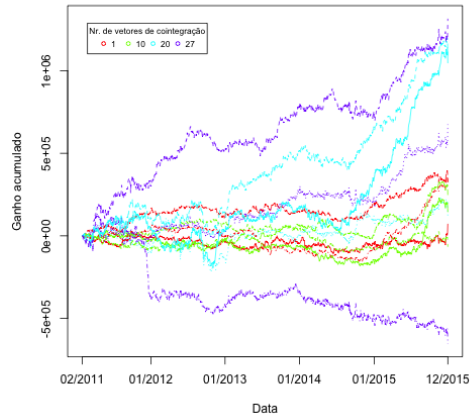
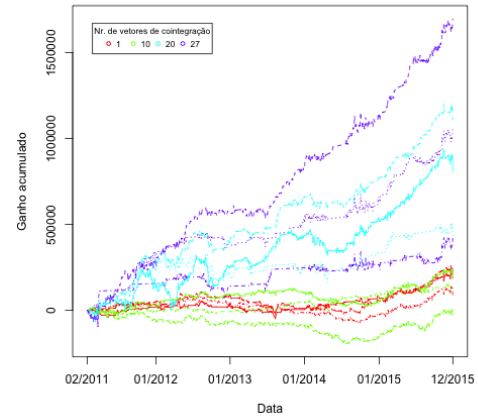


Figura 6 – Ganhos para os portf lios de um fator com  $N = 10$ , detalhados pelo n mero de vetores de cointegra  o considerados.



(a) Estrat gia (1) (2 pontos)



(b) Estrat gia (2) (3 pontos)

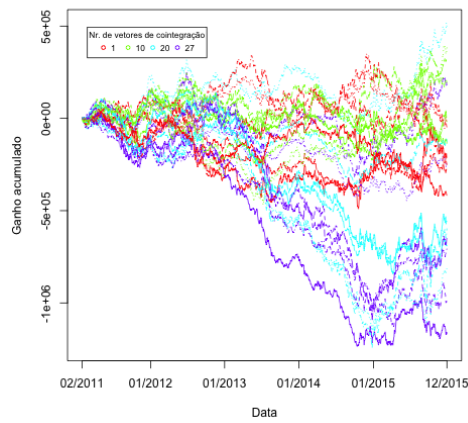
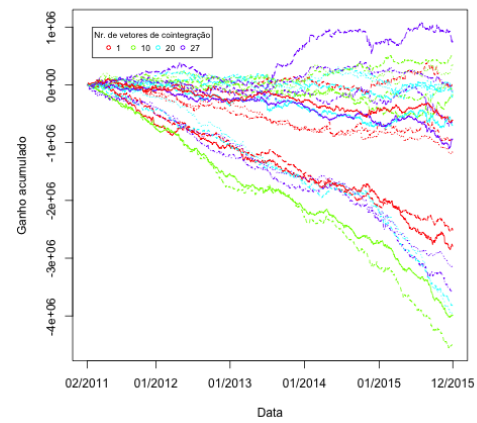
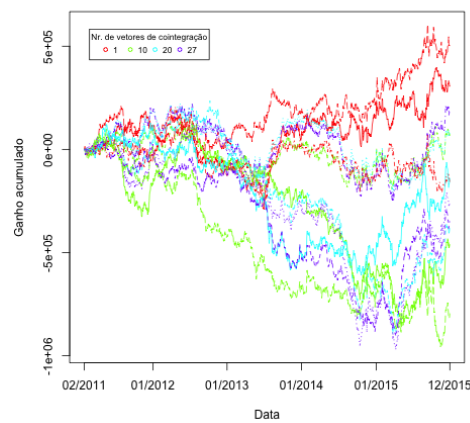
(c) Estrat gia (3) (*Dummy*)(d) Estrat gia (4) (*Max Sharpe*)(e) Estrat gia (5) (*Benchmark*)

Figura 7 – Ganhos para o portf lio de tr s fatores



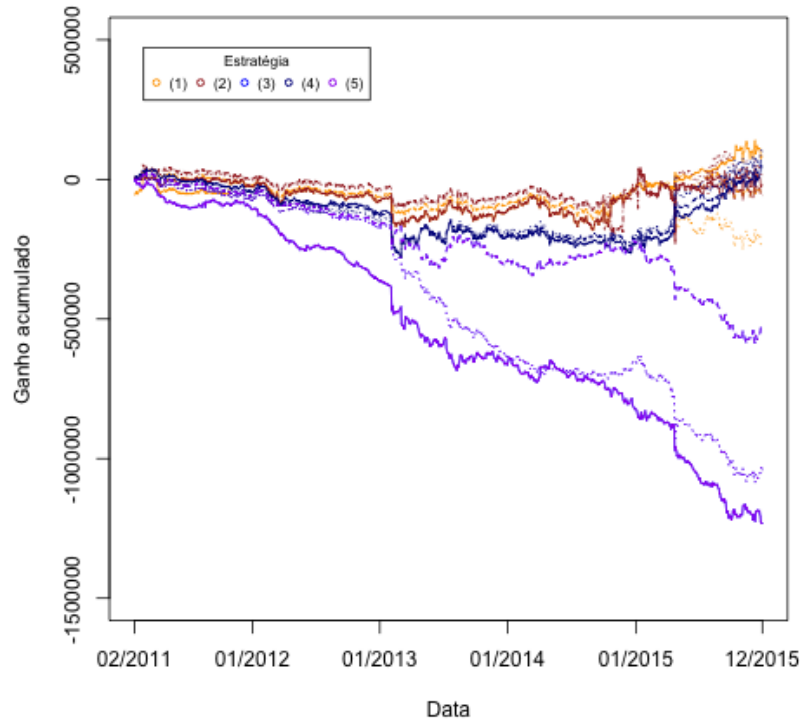


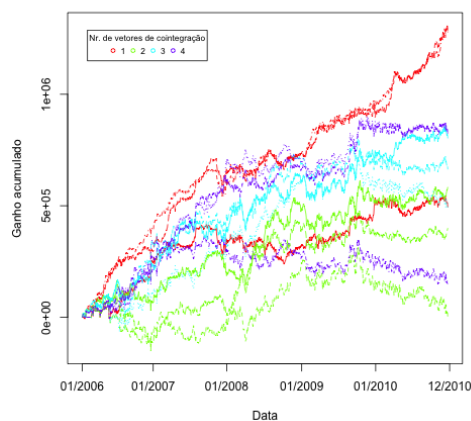
Figura 8 – Ganhos para o portfólio com ativos da mesma empresa, PETR3 e PETR4

## 7.5 Teste de arbitragem estatística para o portfólio de três fatores

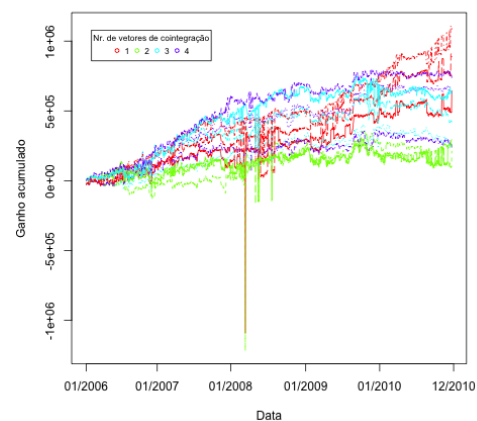
Nesta seção serão aprofundados os resultados do portfólio de três fatores, que utiliza a maior quantidade de informação disponível e obteve boa performance nas estratégias propostas. Para determinar se houve arbitragem estatística nas estratégias testadas foi utilizado o teste de arbitragem estatística proposto por (JARROW et al., 2012) e previamente descrito no capítulo 4.

No momento de estimar os parâmetros do modelo para os ganhos acumulados, originalmente foram abordados tanto os modelos CM (onde  $\theta = 0$ ) e UM ( $\theta \neq 0$ ), mas após as estimativas dos parâmetros, ficou evidente que o parâmetro  $\theta$  era não significativo para a maioria das simulações. Por isso, somente o modelo CM foi considerado para realização do teste.

Os resultados para o teste JTTW para um nível de significância de 5% são mostrados na figura 10. Pelo teste, apenas realizações das estratégias (1) e (2) apresentaram arbitragem estatística em alguma simulação.



(a) Estrat gia (1) (2 pontos)



(b) Estrat gia (2) (3 pontos)

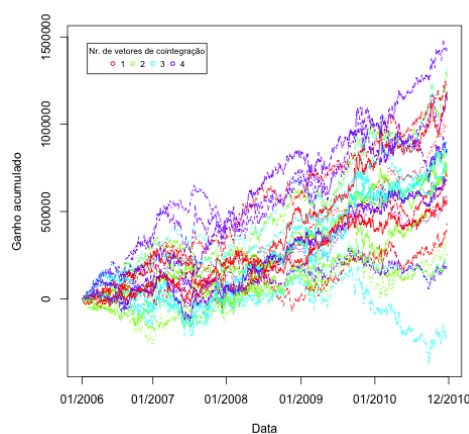
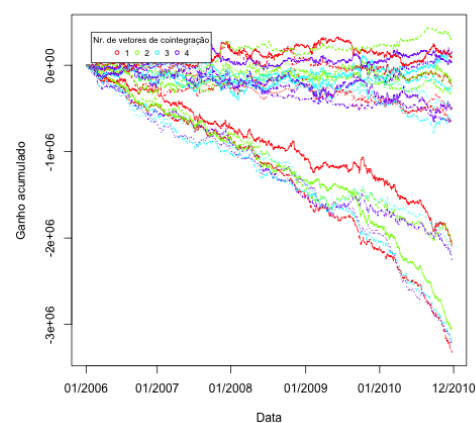
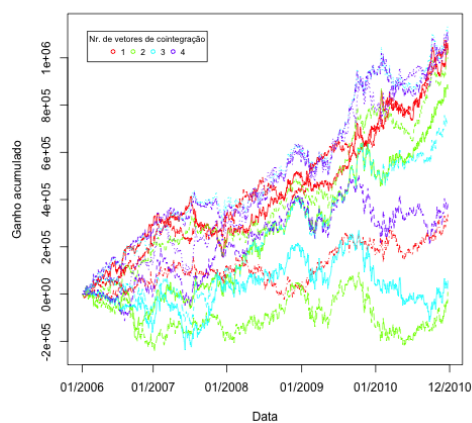
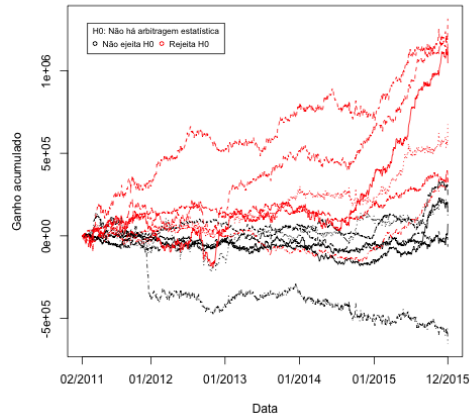
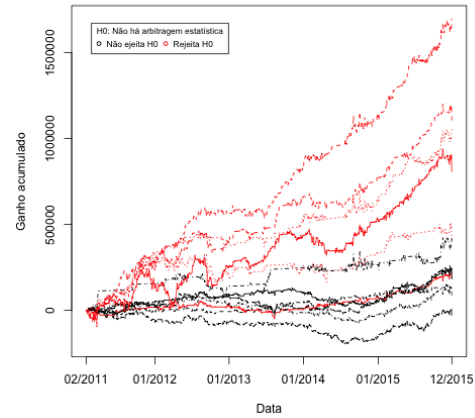
(c) Estrat gia (3) (*Dummy*)(d) Estrat gia (4) (*Max Sharpe*)(e) Estrat gia (5) (*Benchmark*)

Figura 9 – Ganhos para o portf lio do setor banc rio



(a) Estrat gia (1) (2 pontos)



(b) Estrat gia (2) (3 pontos)

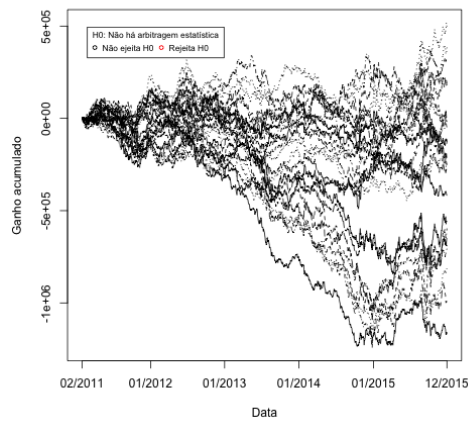
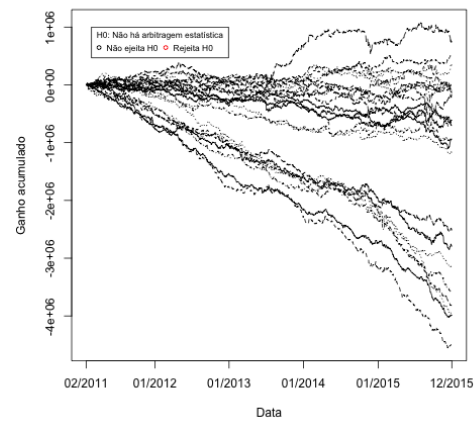
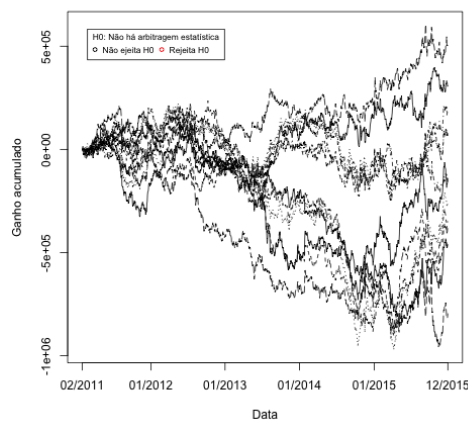
(c) Estrat gia (3) (*Dummy*)(d) Estrat gia (4) (*Max Sharpe*)(e) Estrat gia (5) (*Benchmark*)

Figura 10 – Resultado do teste JTTW para todas as estrat gias no portf lio de tr s fatores

## 7.6 Backtest

Até agora foi visto que os ganhos mais consistentes são provenientes do portfólio de três fatores latentes, utilizando todos os vetores de cointegração. Para identificar se há efeito do momento do mercado nestas conclusões, foi feita uma segunda simulação utilizando uma amostra no período de 2002 a 2010.

Os ativos foram selecionados pelo mesmo método da amostra principal. No entanto, o universo considerado nesta amostra de acordo com os critérios de liquidez explicados previamente é de apenas 38 ativos. Por isso, foram selecionados apenas os cinco ativos com maiores pesos para cada tendência comum. A amostra final é composta então por 15 ativos com os maiores pesos na *loading matrix*  $\Lambda$  na preamostra: TRPL4, PETR4, PETR3, BRKM5, ITSA4, CPLE6, INEP4, FJTA4, VALE5, VALE3, EMBR3, CLSC4, VIVT4, PCAR4 e BBDC4. Novamente, o setor de atuação das empresas selecionadas é muito diversificado.

Assim como na simulação anterior, os primeiros quatro anos foram utilizados para estimação apenas, enquanto os anos 2006 a 2010 foram utilizados tanto para estimação como para simulação das estratégias. No histórico desta amostra está incluído o período da crise de 2008-2009. Como pode se observar na figura 11, neste período há uma forte queda em todos os preços da amostra e é de interesse analisar seu efeito nos ganhos das simulações.

Os resultados das simulações, considerando o mesmo volume de negociação de R\$ 1 milhão e custos de operação como nas simulações anteriores, são mostrados na figura 12. Percebe-se que a presença da crise torna os ganhos das estratégias mais sensíveis à calibração do modelo: as estratégias (1) e (2) performam bem quando utilizados todos os vetores de cointegração, mas o desempenho dos métodos e relação aos outros parâmetros (limites de *enter* e *exit*) se alterou respeito simulação anterior. Estratégias (1) e (2) se beneficiam agora de limites menores para tomar as posições ( $k = 10\%$  e  $20\%$ ).

Diferentemente da outra amostra, as estratégias baseadas em *forecasts* agora apresentaram ganhos superiores e com menos volatilidade. Novamente, uma amostra de menor tamanho mostra-se mais eficaz ao trabalhar com *forecasts*. Este tipo de estratégia tem, portanto, um comportamento mais imprevisível: é mais dependente do tamanho do portfólio e do momento do mercado.

No momento da crise em particular, há saltos em praticamente todas as simulações que tiveram resultados positivos no final do período. Isto é indício de que há erros de precificação durante a crise dos quais as estratégias *longshort* se beneficiaram se bem calibradas. Conclui-se que estes métodos não são dependentes do momento econômico do mercado nem de um grupo de ativos em particular.

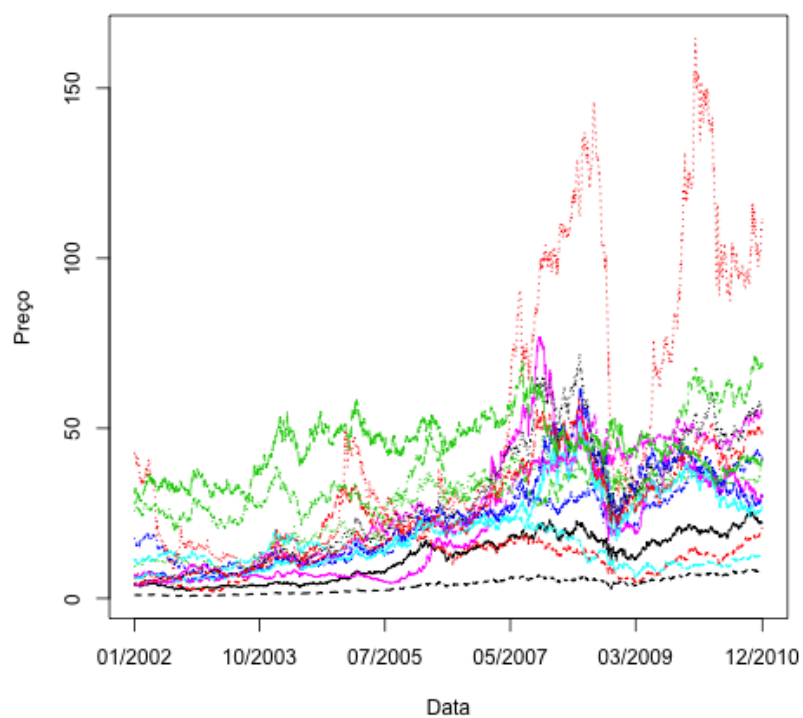


Figura 11 – *Backtest*: Preços diários dos ativos do portfólio de três fatores

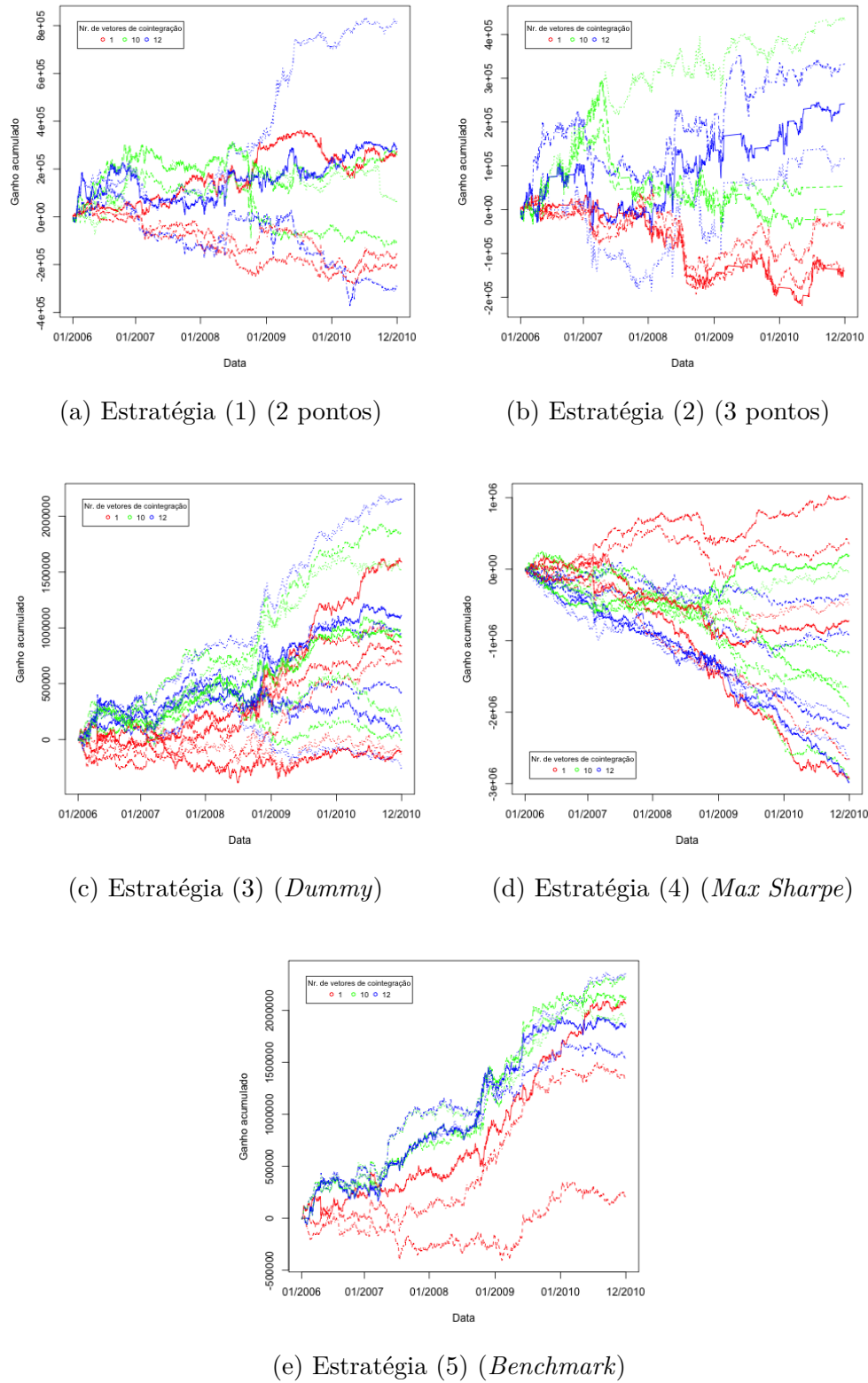


Figura 12 – *Backtest*: Resultado das estratégias no portfólio de três fatores

## 8 Considerações finais

A prática mais comum para trabalhar com séries de preços de ações é tomar suas diferenças e estudar o resultado com modelos voltados para séries estacionárias. Ao fazer isto, no entanto, perde-se a informação sobre tendências comuns e cointegração das séries. Neste trabalho foi abordado o uso de VECM para modelar séries de preços do mercado brasileiro de ações sem diferenciá-las. Foram sugeridas duas estratégias de *trading* do tipo *longshort* onde os pesos de alocação são diretamente determinados pelos vetores de cointegração. Estas estratégias foram simuladas em diferentes tipos de portfólios e comparadas a outras três estratégias de *trading* comuns, onde os pesos de alocação são determinados por *forecasts*. Ao todo foram testadas cinco tipos de estratégias de *trading* em seis tipos de portfólios diferentes.

Quando bem calibradas, as estratégias propostas apresentaram ganhos que superam o volume máximo negociado em cinco anos de operação. Estes ganhos são particularmente interessantes porque nenhuma das estratégias testadas precisa de aporte de capital próprio. Elas são *self-financing* até onde possível, e o baixo custo marginal pode ser tomado da taxa livre de risco se preciso.

Quando o portfólio é composto por ativos com uma tendência comum muito forte, como acontece quando a seleção de ativos é feita considerando apenas o primeiro fator latente ou ao considerar a tendência de um setor do mercado em particular, considerar apenas um vetor de cointegração para gerar sinais de *trading* para as estratégias propostas é suficiente. Isto é, a reversão dos preços em relação àquela tendência é suficientemente bem explicada por apenas um vetor de cointegração, permitindo trabalhar com uma forma mais simplificada do modelo.

Por outro lado, ótimos resultados com baixa volatilidade são obtidos ao trabalhar com um portfólio maior e mais complexo, que é o caso do portfólio selecionado considerando o total de  $r = 3$  tendências comuns. Neste caso, é possível utilizar todos os  $N - r$  vetores de cointegração estimados para gerar ganhos consistentemente ao longo do tempo. Neste portfólio, as duas estratégias propostas, quando bem calibradas, apresentam oportunidades de arbitragem estatística pelo teste JTTW.

Definir os pesos em função dos vetores de cointegração, no entanto, cria uma sensibilidade maior das estratégias em relação a mudanças na estrutura dos dados. Como visto no *backtest*, em períodos de crise, é necessário que os parâmetros do modelo que definem quando há erros de precificação sejam monitorados e recalibrados se preciso. Além disto, a imprevisibilidade da frequência de *trading* deve ser bem estudada: como tomar e desfazer as posições não dependem de um prazo fixo, a estrutura do portfólio e o

momento do mercado podem fazer com que uma posição fique tomada por muito tempo. Para mitigar este risco, é necessário procurar um portfólio de ativos que viabilize gerar ganhos com uma frequência compatível com as necessidades do investidor. Medidas de *stop-loss*, por exemplo, também podem ser implementadas para agilizar as possibilidades de transação.

Por último, os resultados mostram que utilizar VECM para modelar as séries de preços e logpreços gera boas estimativas dos vetores de cointegração e também bons *forecasts*. Por isso, as estratégias que dependem da boa qualidade das previsões também apresentaram ganhos significativos.

Apesar dos bons resultados, não é suficiente apenas a identificação de um grupo cointegrado para poder aproveitar oportunidades de arbitragem. Quando há poucas evidências de erros de precificação ou quando há efeito de influências externas (como no portfólio composto por PETR3 e PETR4, por exemplo), nenhuma metodologia abordada foi capaz de gerar ganhos relevantes. Destaca-se neste caso a necessidade de identificar não somente portfólios cointegrados, mas também portfólios em que erros de precificação estejam presentes de forma compatível com o modelo e frequência de *trading* escolhidas.

Os resultados podem ser aperfeiçoados com o uso de outras técnicas para determinação de tendências além de PCA, como recentes extensões de modelos de fatores dinâmicos para séries cointegradas ou o modelo de espaço estado sugerido por (CHANG; JIANG; PARK, 2008). Por último, seria interessante também aprimorar os critérios de todas as estratégias de modo a permitir ações de *stop-loss* e medidas semelhantes aplicadas na prática, tornando-as mais flexíveis.



# Referências

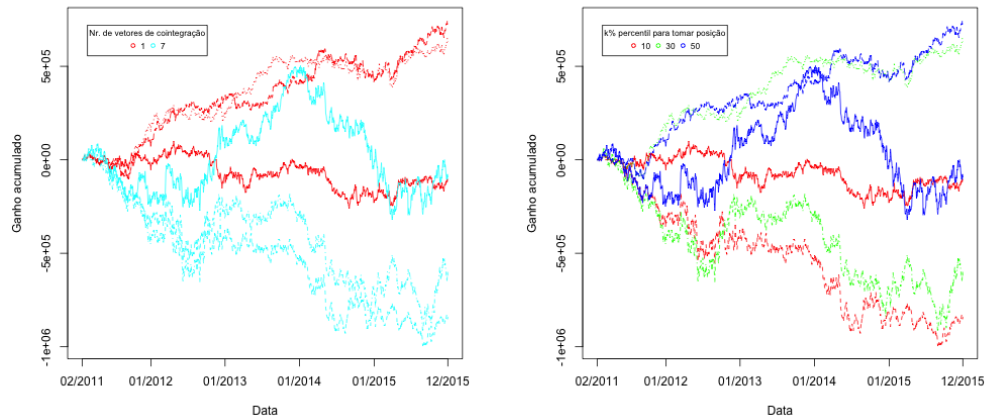
- AMENGUAL, D.; WATSON, M. W. *Consistent Estimation Of The Number Of Dynamic Factors In A Large N And T Panel*. Dissertação (Mestrado) — Princeton University, May 2006.
- BAI, J. Estimating cross-section common stochastic trends in nonstationary panel data. *Journal of Econometrics*, v. 122, p. 137–183, October 2004.
- BAI, J.; NG, S. Determining the number of factors in approximate factor models. *Econometrica*, v. 70, n. 1, p. 191–221, Janeiro 2002.
- BAI, J.; NG, S. Principal components estimation and identification of static factors. *Journal of Econometrics*, 2013.
- BERTRAM, W. K. Optimal trading strategies for itô diffusion processes. *PhysicaA*, March 2009.
- BOSSAERTS, P. L. Common nonstationary components of asset prices. *Journal of Economic Dynamics and Control*, v. 12, p. 347–364, 1988.
- CALDEIRA, J. F.; MOURA, G. V. Seleção de uma carteira de pares de ações usando cointegração: uma estratégia de arbitragem estatística. *Rev. Bras. Finanças (Online)*, v. 11, n. 1, p. 49–80, Março 2013.
- CARVALHO, J. P. de. *Modelo de fatores dinâmicos: aplicação à estrutura a termo da taxa de juros*. Dissertação (Mestrado) — UFSC, 2013.
- CHANG, Y.; JIANG, B.; PARK, J. Y. Using kalman filter to extract and test for common stochastic trends. In: *Midwest Econometrics Conference*. [S.l.: s.n.], 2008.
- CHANG, Y.; MILLER, J. I.; PARK, J. Y. Extracting a common stochastic trend: Theories with some applications. *Symposium on Econometric Theory and Applications*, 2005.
- DIEBOLD, F. X.; LI, C. Forecasting the term structure of government bond yields. *Journal of Econometrics*, v. 130, p. 337–364, 2006.
- ELLIOTT, R. J.; HOEK, J. V. D.; MALCOLM, W. P. Pairs trading. *Quantitative Finance*, v. 5, n. 3, p. 271–276, June 2005.
- FABOZZI, F. J.; FOCARDI, S. M.; KOLM, P. N. *Quantitative Equity Investing*. [S.l.]: Wiley, 2010.
- FOCARDI, S. M.; FABOZZI, F. J.; MITOV, I. K. A new approach to statistical arbitrage: Strategies based on dynamic factor models of prices and their performance. *Journal of Banking & Finance*, April 2016.
- GÖNCÜ, A. Statistical arbitrage in the black-scholes framework. Última revisão em 30 de Agosto. 2014.
- GÖNCÜ, A. Existence of statistical arbitrage portfolios in the black-scholes framework. Última revisão em Fevereiro. 2015.

- JARROW, R. et al. Testing market efficiency using statistical arbitrage with applications to momentum and value trading strategies. *Journal of economics*, v. 73, p. 525–565, 2004.
- JARROW, R. et al. An improved test for statistical arbitrage. *Journal of Financial Markets*, v. 15, n. 15, p. 47–80, February 2012. An improved test for statistical arbitrage.
- LEHMANN, B. Fads, martingales, and market efficiency. *Quarterly Journal of Economics*, v. 105, n. 1, p. 1–28, February 1990.
- MAYORDOMO, S.; PEÑA, J. I.; ROMO, J. A new test of statistical arbitrage with applications to credit derivatives markets. *Comisión Nacional del Mercado de Valores*, Junho 2011.
- MORETTIN, P. A. *Econometria financeira - Um curso em séries temporais financeiras*. 2. ed. [S.l.]: Blucher, 2011.
- NYBLOM, J.; HARVEY, A. Test of common stochastic trends. *Econometric Theory*, v. 16, n. 2, p. 176–199, 2000.
- PEÑA, D.; PONCELA, P. Nonstationary dynamic factor analysis. *Journal of Statistical Planning and Inference*, April 2006.
- PUCCIARELLI, A. J. *Estratégia de cointegração dinâmica empírica para arbitragem estatística e trading*. Dissertação (Mestrado) — Fundação Getúlio Vargas São Paulo, 2014.
- TRIAFYLOPOULOS, K.; MONTANA, G. Dynamic modeling of mean-reverting spreads for statistical arbitrage. *Computational Management Science*, v. 8, n. 1, p. 23–49, April 2011.
- ZHANG, W.; DAI, B. P. Z.; DJARBIROV, M. A multi-factor adaptive statistical arbitrage model. ArXiv:1405.2384. 2014.

## Apêndices

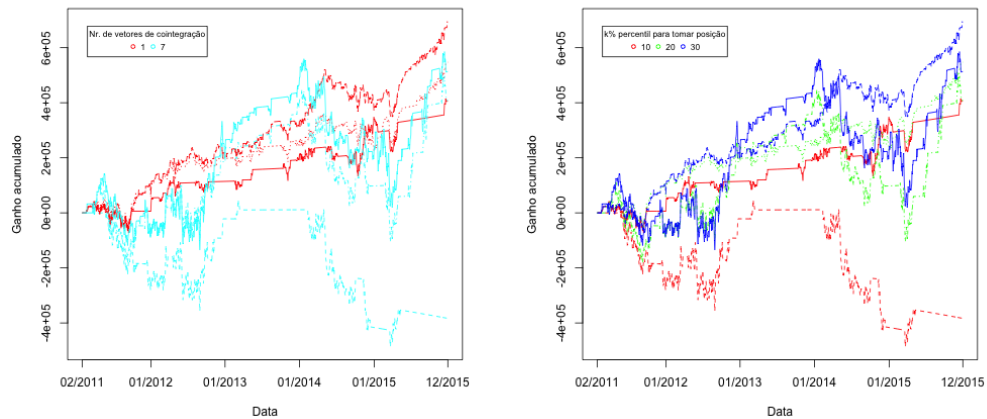
# APÊNDICE A – Gráficos dos ganhos acumulados das simulações

## A.1 Portfólio de um fator



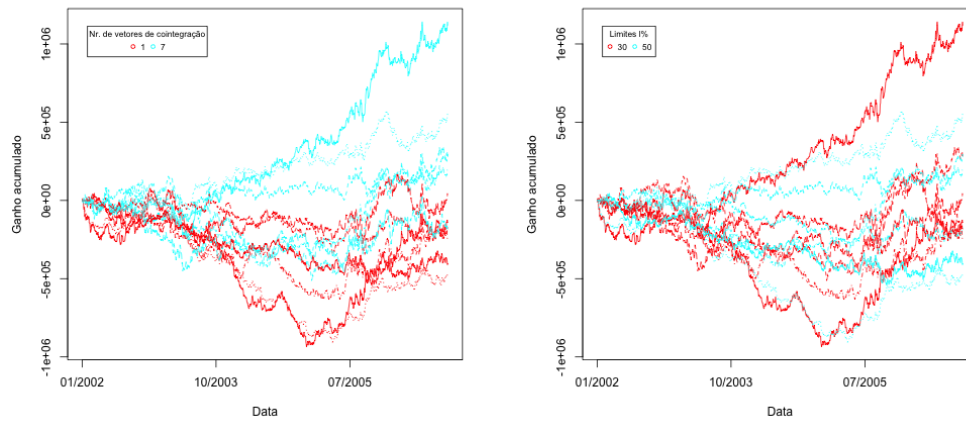
(a) Resultados por número de vetores de cointegração considerados (b) Resultados por percentis de  $S_t$  considerados

Figura 13 – Detalhamento do resultado para a estratégia (1) no portfólio de um fator com  $N = 10$  ativos



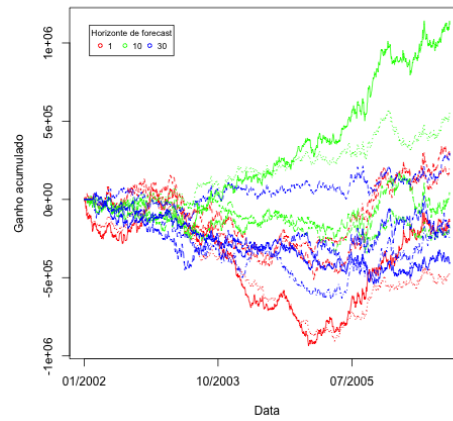
(a) Resultados por número de vetores de cointegração considerados (b) Resultados por percentis de  $S_t$  considerados

Figura 14 – Detalhamento do resultado para a estratégia (2) no portfólio de um fator com  $N = 10$  ativos



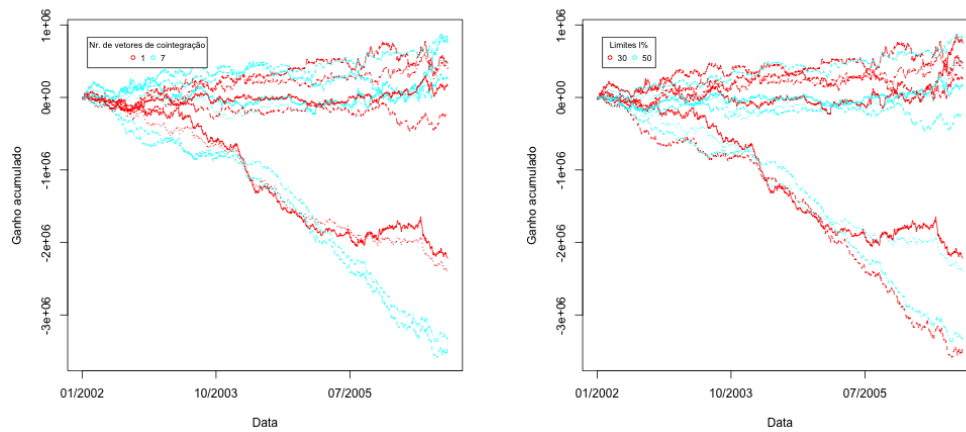
(a) Resultados por número de vetores de cointegração considerados

(b) Resultados por limites  $l$



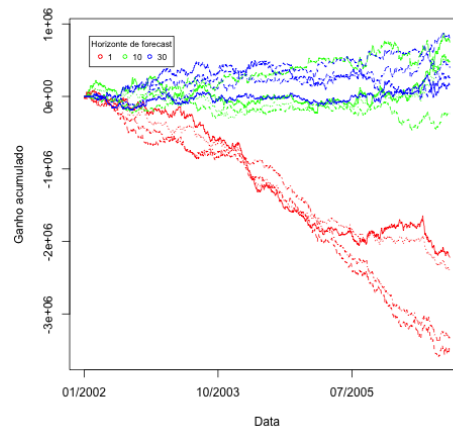
(c) Resultados por horizonte de *forecast*

Figura 15 – Detalhamento do resultado para a estratégia (3) no portfólio de um fator com  $N = 10$  ativos



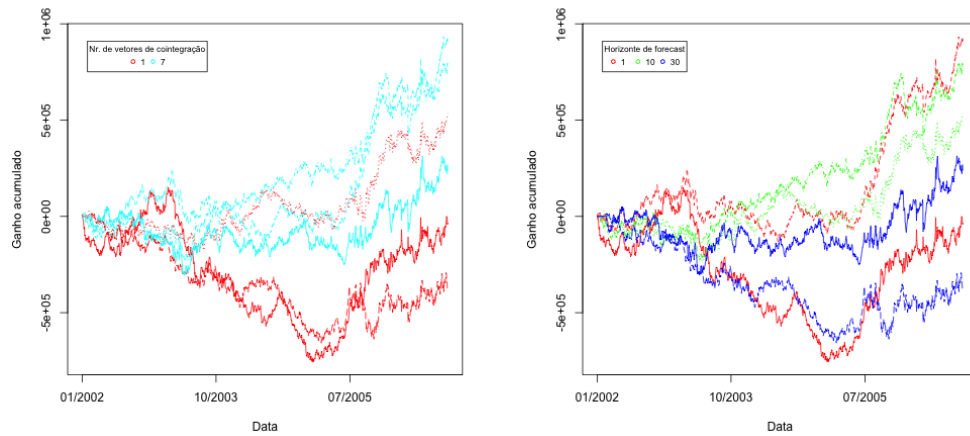
(a) Resultados por número de vetores de cointegração considerados

(b) Resultados por limites  $l$



(c) Resultados por horizonte de *forecast*

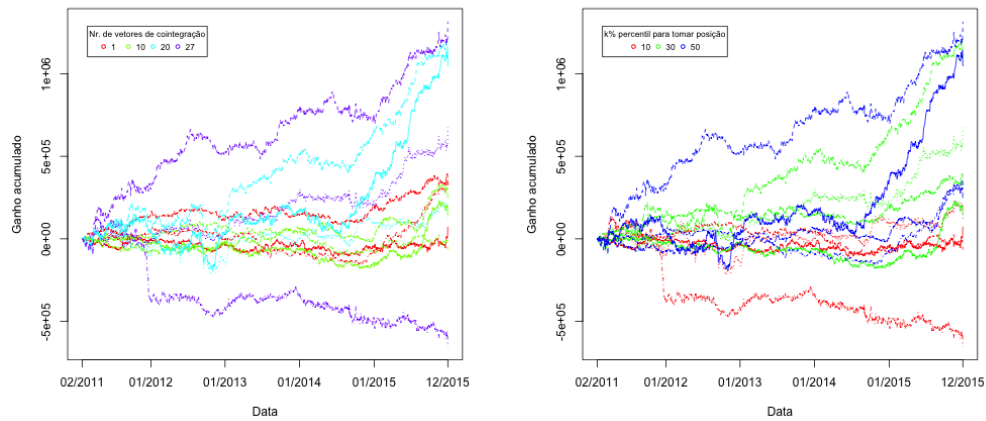
Figura 16 – Detalhamento do resultado para a estratégia (4) no portfólio de um fator com  $N = 10$  ativos



(a) Resultados por número de vetores de cointegração considerados (b) Resultados por horizonte de *forecast*

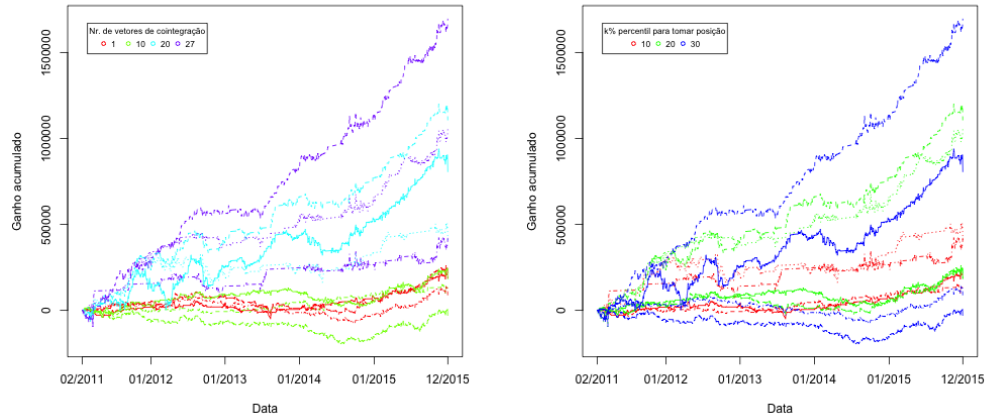
Figura 17 – Detalhamento do resultado para a estratégia (5) no portfólio de um fator com  $N = 10$  ativos

## A.2 Portfólio de três fatores



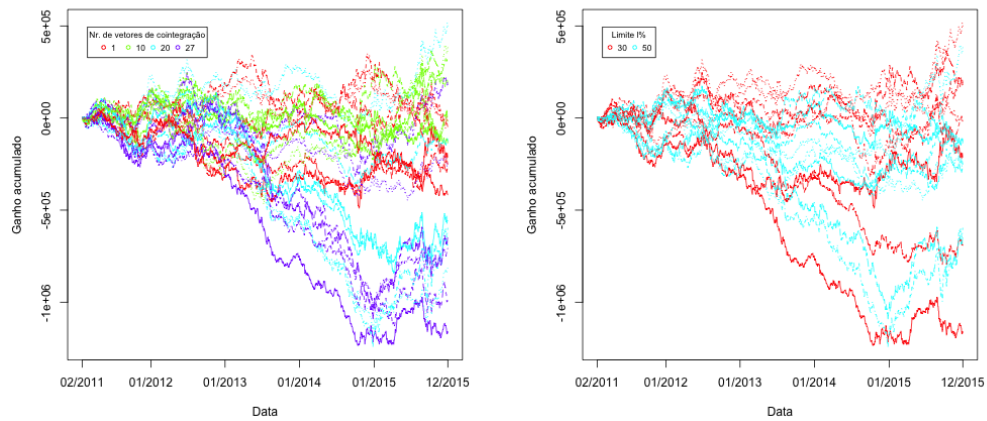
(a) Resultados por número de vetores de cointegração considerados (b) Resultados por percentis de  $S_t$  considerados

Figura 18 – Detalhamento do resultado para a estratégia (1) no portfólio de três fatores

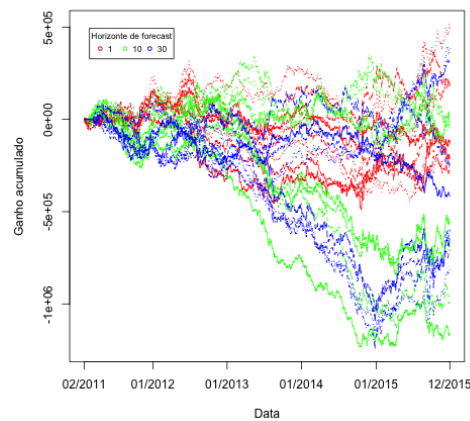


(a) Resultados por número de vetores de cointegração considerados (b) Resultados por percentis de  $S_t$  considerados

Figura 19 – Detalhamento do resultado para a estratégia (2) no portfólio de três fatores



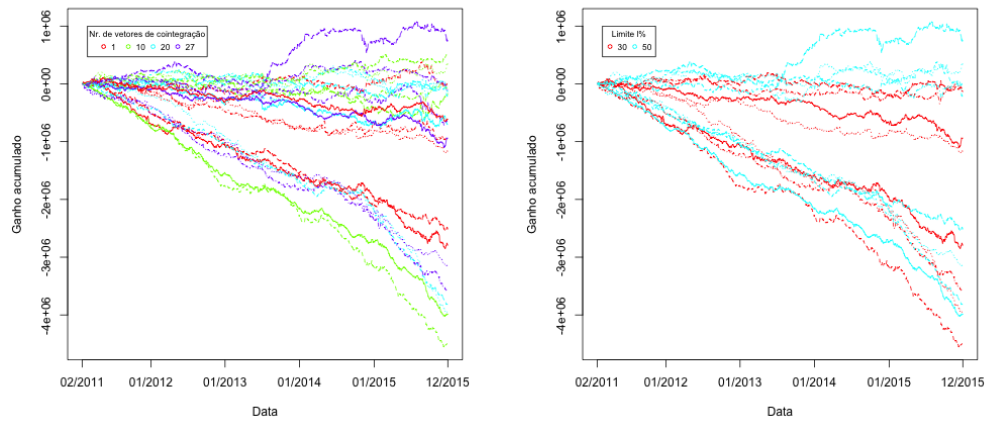
(a) Resultados por número de vetores de cointegração considerados (b) Resultados por limites  $l$



(c) Resultados por horizonte de *forecast*

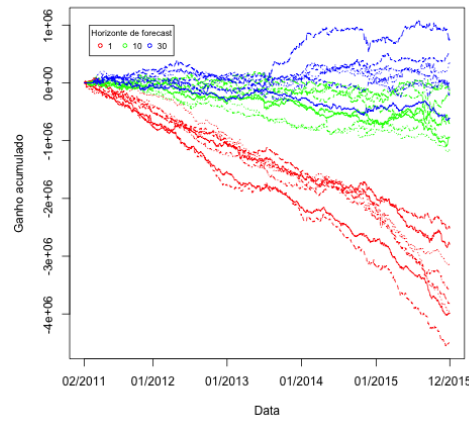
Figura 20 – Detalhamento do resultado para a estratégia (3) no portfólio de três fatores





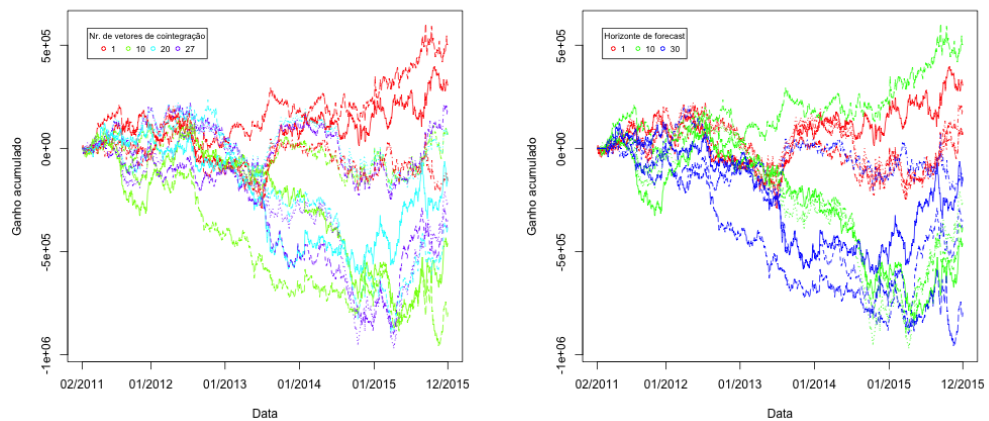
(a) Resultados por número de vetores de cointegração considerados

(b) Resultados por limites  $l$



(c) Resultados por horizonte de *forecast*

Figura 21 – Detalhamento do resultado para a estratégia (4) no portfólio de três fatores



(a) Resultados por número de vetores de cointegração considerados

(b) Resultados por horizonte de *forecast*

Figura 22 – Detalhamento do resultado para a estratégia (5) no portfólio de três fatores

# APÊNDICE B – Medidas resumo das simulações

## B.1 Portfólio de um fator com $N = 10$ ativos

Tabela 8 – Resumo da estratégia (1) no portfólio de um fator

| Estratégia                                     | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|--|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ , $N = 10$ | -0.103                           | 0.007  | -1.488             | 10                                  |
| Nr. vetores de coint. = 7, $k = 10$ , $N = 10$ | -0.840                           | 0.009  | -0.281             | 11                                  |
| Nr. vetores de coint. = 1, $k = 30$ , $N = 10$ | 0.631                            | 0.007  | 0.350              | 39                                  |
| Nr. vetores de coint. = 7, $k = 30$ , $N = 10$ | -0.620                           | 0.016  | -0.406             | 32                                  |
| Nr. vetores de coint. = 1, $k = 50$ , $N = 10$ | 0.725                            | 0.007  | 0.676              | 150                                 |
| Nr. vetores de coint. = 7, $k = 50$ , $N = 10$ | -0.093                           | 0.020  | -4.889             | 120                                 |

Tabela 9 – Resumo da estratégia (2) no portfólio de um fator

| Parâmetros  | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ , $N = 10$    | 0.406                            | 0.005  | 0.259              | 36                                  |
| Nr. vetores de coint. = 7, $k = 10$ , $N = 10$    | -0.383                           | 0.021  | -0.369             | 20                                  |
| Nr. vetores de coint. = 1, $k = 20$ , $N = 10$    | 0.547                            | 0.006  | 0.335              | 86                                  |
| Nr. vetores de coint. = 7, $k = 20$ , $N = 10$    | 0.409                            | 0.023  | 0.519              | 54                                  |
| Nr. vetores de coint. = the1, $k = 30$ , $N = 10$ | 0.694                            | 0.007  | 0.465              | 189                                 |
| Nr. vetores de coint. = 7, $k = 30$ , $N = 10$    | 0.513                            | 0.027  | 0.607              | 122                                 |

Tabela 10 – Resumo da estratégia (3) no portfólio de um fator

| Parâmetros   | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|--|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 1$ , $N = 10$  | -0.137                           | 0.015  | -14.525            | 1216                                |
| Nr. vetores de coint. = 7, $l = 30$ , $hf = 1$ , $N = 10$  | 0.270                            | 0.015  | 5.340              | 1216                                |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$ , $N = 10$  | -0.498                           | 0.011  | -2.993             | 1217                                |
| Nr. vetores de coint. = 7, $l = 50$ , $hf = 1$ , $N = 10$  | 0.173                            | 0.011  | 6.190              | 1217                                |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 10$ , $N = 10$ | 0.043                            | 0.013  | 7.013              | 120                                 |
| Nr. vetores de coint. = 7, $l = 30$ , $hf = 10$ , $N = 10$ | 1.118                            | 0.014  | 0.251              | 120                                 |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ , $N = 10$ | -0.178                           | 0.010  | -1.570             | 121                                 |
| Nr. vetores de coint. = 7, $l = 50$ , $hf = 10$ , $N = 10$ | 0.533                            | 0.011  | 0.447              | 121                                 |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 30$ , $N = 10$ | -0.237                           | 0.011  | -0.871             | 39                                  |
| Nr. vetores de coint. = 7, $l = 30$ , $hf = 30$ , $N = 10$ | -0.184                           | 0.014  | -1.219             | 39                                  |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ , $N = 10$ | -0.402                           | 0.008  | -0.502             | 39                                  |
| Nr. vetores de coint. = 7, $l = 50$ , $hf = 30$ , $N = 10$ | 0.270                            | 0.011  | 0.685              | 39                                  |

Tabela 11 – Resumo da estratégia (4) no portfólio de um fator

| Parâmetros   | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|--|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 1$ , $N = 10$  | -2.223                           | 0.017  | -1.186             | 1217                                |
| Nr. vetores de coint. = 7, $l = 30$ , $hf = 1$ , $N = 10$  | -3.457                           | 0.017  | -0.713             | 1217                                |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$ , $N = 10$  | -2.408                           | 0.014  | -0.979             | 1217                                |
| Nr. vetores de coint. = 7, $l = 50$ , $hf = 1$ , $N = 10$  | -3.315                           | 0.014  | -0.675             | 1217                                |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 10$ , $N = 10$ | 0.484                            | 0.018  | 0.748              | 121                                 |
| Nr. vetores de coint. = 7, $l = 30$ , $hf = 10$ , $N = 10$ | 0.743                            | 0.020  | 0.546              | 121                                 |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ , $N = 10$ | -0.239                           | 0.014  | -1.437             | 121                                 |
| Nr. vetores de coint. = 7, $l = 50$ , $hf = 10$ , $N = 10$ | 0.455                            | 0.014  | 0.796              | 121                                 |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 30$ , $N = 10$ | 0.393                            | 0.017  | 0.532              | 40                                  |
| Nr. vetores de coint. = 7, $l = 30$ , $hf = 30$ , $N = 10$ | 0.242                            | 0.013  | 1.076              | 40                                  |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ , $N = 10$ | 0.153                            | 0.012  | 1.340              | 40                                  |
| Nr. vetores de coint. = 7, $l = 50$ , $hf = 30$ , $N = 10$ | 0.807                            | 0.012  | 0.293              | 40                                  |

Tabela 12 – Resumo da estratégia (5) no portfólio de um fator

| Parâmetros                                      | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $hf = 1$ , $N = 10$  | -0.033                           | 0.016  | -60.093            | 1217                                |
| Nr. vetores de coint. = 7, $hf = 1$ , $N = 10$  | 0.908                            | 0.015  | 1.612              | 1217                                |
| Nr. vetores de coint. = 1, $hf = 10$ , $N = 10$ | 0.534                            | 0.012  | 0.567              | 121                                 |
| Nr. vetores de coint. = 7, $hf = 10$ , $N = 10$ | 0.771                            | 0.014  | 0.363              | 121                                 |
| Nr. vetores de coint. = 1, $hf = 30$ , $N = 10$ | -0.376                           | 0.012  | -0.557             | 40                                  |
| Nr. vetores de coint. = 7, $hf = 30$ , $N = 10$ | 0.239                            | 0.012  | 0.920              | 40                                  |

## B.2 Portfólio de três fatores

Tabela 13 – Resumo da estratégia (1) no portfólio de três fatores

| Parâmetros                                      | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ , $N = 30$  | 0.073                            | 0.005  | 3.376              | 36                                  |
| Nr. vetores de coint. = 10, $k = 10$ , $N = 30$ | -0.066                           | 0.004  | -3.148             | 36                                  |
| Nr. vetores de coint. = 20, $k = 10$ , $N = 30$ | 0.111                            | 0.007  | 1.991              | 28                                  |
| Nr. vetores de coint. = 27, $k = 10$ , $N = 30$ | -0.654                           | 0.019  | -0.257             | 19                                  |
| Nr. vetores de coint. = 1, $k = 30$ , $N = 30$  | 0.326                            | 0.004  | 1.263              | 122                                 |
| Nr. vetores de coint. = 10, $k = 30$ , $N = 30$ | 0.144                            | 0.005  | 2.451              | 119                                 |
| Nr. vetores de coint. = 20, $k = 30$ , $N = 30$ | 1.085                            | 0.012  | 0.321              | 98                                  |
| Nr. vetores de coint. = 27, $k = 30$ , $N = 30$ | 0.676                            | 0.016  | 0.299              | 61                                  |
| Nr. vetores de coint. = 1, $k = 50$ , $N = 30$  | 0.303                            | 0.005  | 2.788              | 308                                 |
| Nr. vetores de coint. = 10, $k = 50$ , $N = 30$ | 0.247                            | 0.005  | 2.914              | 289                                 |
| Nr. vetores de coint. = 20, $k = 50$ , $N = 30$ | 1.045                            | 0.012  | 0.675              | 264                                 |
| Nr. vetores de coint. = 27, $k = 50$ , $N = 30$ | 1.315                            | 0.021  | 0.394              | 199                                 |

Tabela 14 – Resumo da estratégia (2) no portfólio de três fatores

| Parâmetros                                      | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ , $N = 30$  | 0.209                            | 0.004  | 1.036              | 107                                 |
| Nr. vetores de coint. = 10, $k = 10$ , $N = 30$ | 0.133                            | 0.005  | 1.431              | 83                                  |
| Nr. vetores de coint. = 20, $k = 10$ , $N = 30$ | 0.403                            | 0.010  | 0.417              | 53                                  |
| Nr. vetores de coint. = 27, $k = 10$ , $N = 30$ | 0.427                            | 0.022  | 0.245              | 37                                  |
| Nr. vetores de coint. = 1, $k = 20$ , $N = 30$  | 0.186                            | 0.004  | 1.972              | 219                                 |
| Nr. vetores de coint. = 10, $k = 20$ , $N = 30$ | 0.181                            | 0.005  | 1.783              | 179                                 |
| Nr. vetores de coint. = 20, $k = 20$ , $N = 30$ | 1.094                            | 0.010  | 0.262              | 143                                 |
| Nr. vetores de coint. = 27, $k = 20$ , $N = 30$ | 1.058                            | 0.020  | 0.169              | 85                                  |
| Nr. vetores de coint. = 1, $k = 30$ , $N = 30$  | 0.072                            | 0.004  | 7.542              | 347                                 |
| Nr. vetores de coint. = 10, $k = 30$ , $N = 30$ | -0.048                           | 0.005  | -10.086            | 307                                 |
| Nr. vetores de coint. = 20, $k = 30$ , $N = 30$ | 0.804                            | 0.013  | 0.586              | 281                                 |
| Nr. vetores de coint. = 27, $k = 30$ , $N = 30$ | 1.696                            | 0.020  | 0.196              | 199                                 |

Tabela 15 – Resumo da estratégia (3) no portfólio de três fatores

| Parâmetros  | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 1$ , $N = 30$   | -0.213                           | 0.012  | -10.500            | 1217                                |
| Nr. vetores de coint. = 10, $l = 30$ , $hf = 1$ , $N = 30$  | 0.192                            | 0.013  | 9.118              | 1217                                |
| Nr. vetores de coint. = 20, $l = 30$ , $hf = 1$ , $N = 30$  | 0.497                            | 0.013  | 3.203              | 1217                                |
| Nr. vetores de coint. = 27, $l = 30$ , $hf = 1$ , $N = 30$  | 0.179                            | 0.013  | 8.669              | 1217                                |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$ , $N = 30$   | -0.294                           | 0.009  | -5.783             | 1216                                |
| Nr. vetores de coint. = 10, $l = 50$ , $hf = 1$ , $N = 30$  | -0.136                           | 0.009  | -9.566             | 1216                                |
| Nr. vetores de coint. = 20, $l = 50$ , $hf = 1$ , $N = 30$  | -0.127                           | 0.009  | -9.294             | 1216                                |
| Nr. vetores de coint. = 27, $l = 50$ , $hf = 1$ , $N = 30$  | -0.263                           | 0.009  | -4.438             | 1216                                |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 10$ , $N = 30$  | -0.051                           | 0.013  | -7.922             | 121                                 |
| Nr. vetores de coint. = 10, $l = 30$ , $hf = 10$ , $N = 30$ | -0.012                           | 0.012  | -32.480            | 121                                 |
| Nr. vetores de coint. = 20, $l = 30$ , $hf = 10$ , $N = 30$ | -0.557                           | 0.012  | -0.638             | 121                                 |
| Nr. vetores de coint. = 27, $l = 30$ , $hf = 10$ , $N = 30$ | -0.987                           | 0.013  | -0.357             | 121                                 |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ , $N = 30$  | -0.024                           | 0.009  | -14.703            | 120                                 |
| Nr. vetores de coint. = 10, $l = 50$ , $hf = 10$ , $N = 30$ | -0.024                           | 0.009  | -13.253            | 120                                 |
| Nr. vetores de coint. = 20, $l = 50$ , $hf = 10$ , $N = 30$ | -0.693                           | 0.009  | -0.434             | 120                                 |
| Nr. vetores de coint. = 27, $l = 50$ , $hf = 10$ , $N = 30$ | -1.169                           | 0.010  | -0.257             | 120                                 |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 30$ , $N = 30$  | -0.265                           | 0.012  | -1.031             | 40                                  |
| Nr. vetores de coint. = 10, $l = 30$ , $hf = 30$ , $N = 30$ | 0.384                            | 0.012  | 0.719              | 39                                  |
| Nr. vetores de coint. = 20, $l = 30$ , $hf = 30$ , $N = 30$ | -0.596                           | 0.015  | -0.437             | 39                                  |
| Nr. vetores de coint. = 27, $l = 30$ , $hf = 30$ , $N = 30$ | -0.645                           | 0.014  | -0.408             | 39                                  |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ , $N = 30$  | -0.419                           | 0.008  | -0.592             | 39                                  |
| Nr. vetores de coint. = 10, $l = 50$ , $hf = 30$ , $N = 30$ | 0.337                            | 0.009  | 0.718              | 39                                  |
| Nr. vetores de coint. = 20, $l = 50$ , $hf = 30$ , $N = 30$ | -0.813                           | 0.011  | -0.286             | 39                                  |
| Nr. vetores de coint. = 27, $l = 50$ , $hf = 30$ , $N = 30$ | -0.706                           | 0.010  | -0.337             | 39                                  |

Tabela 16 – Resumo da estratégia (4) no portfólio de três fatores

| Parâmetros  | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 1$ , $N = 30$   | -2.799                           | 0.015  | -0.986             | 1217                                |
| Nr. vetores de coint. = 10, $l = 30$ , $hf = 1$ , $N = 30$  | -4.481                           | 0.015  | -0.604             | 1217                                |
| Nr. vetores de coint. = 20, $l = 30$ , $hf = 1$ , $N = 30$  | -3.986                           | 0.015  | -0.672             | 1217                                |
| Nr. vetores de coint. = 27, $l = 30$ , $hf = 1$ , $N = 30$  | -3.567                           | 0.015  | -0.752             | 1217                                |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$ , $N = 30$   | -2.518                           | 0.012  | -1.000             | 1217                                |
| Nr. vetores de coint. = 10, $l = 50$ , $hf = 1$ , $N = 30$  | -3.978                           | 0.011  | -0.622             | 1217                                |
| Nr. vetores de coint. = 20, $l = 50$ , $hf = 1$ , $N = 30$  | -3.810                           | 0.011  | -0.645             | 1217                                |
| Nr. vetores de coint. = 27, $l = 50$ , $hf = 1$ , $N = 30$  | -3.143                           | 0.012  | -0.784             | 1217                                |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 10$ , $N = 30$  | -0.990                           | 0.014  | -0.469             | 121                                 |
| Nr. vetores de coint. = 10, $l = 30$ , $hf = 10$ , $N = 30$ | -0.164                           | 0.015  | -2.805             | 121                                 |
| Nr. vetores de coint. = 20, $l = 30$ , $hf = 10$ , $N = 30$ | -0.702                           | 0.014  | -0.650             | 121                                 |
| Nr. vetores de coint. = 27, $l = 30$ , $hf = 10$ , $N = 30$ | -0.641                           | 0.014  | -0.695             | 121                                 |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ , $N = 30$  | -1.181                           | 0.011  | -0.358             | 121                                 |
| Nr. vetores de coint. = 10, $l = 50$ , $hf = 10$ , $N = 30$ | -0.057                           | 0.012  | -7.388             | 121                                 |
| Nr. vetores de coint. = 20, $l = 50$ , $hf = 10$ , $N = 30$ | -0.140                           | 0.011  | -2.997             | 121                                 |
| Nr. vetores de coint. = 27, $l = 50$ , $hf = 10$ , $N = 30$ | -0.937                           | 0.010  | -0.447             | 121                                 |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 30$ , $N = 30$  | 0.010                            | 0.012  | 30.491             | 40                                  |
| Nr. vetores de coint. = 10, $l = 30$ , $hf = 30$ , $N = 30$ | 0.332                            | 0.014  | 0.873              | 40                                  |
| Nr. vetores de coint. = 20, $l = 30$ , $hf = 30$ , $N = 30$ | -0.212                           | 0.015  | -1.356             | 40                                  |
| Nr. vetores de coint. = 27, $l = 30$ , $hf = 30$ , $N = 30$ | 0.754                            | 0.016  | 0.366              | 40                                  |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ , $N = 30$  | -0.611                           | 0.010  | -0.444             | 40                                  |
| Nr. vetores de coint. = 10, $l = 50$ , $hf = 30$ , $N = 30$ | 0.533                            | 0.010  | 0.516              | 40                                  |
| Nr. vetores de coint. = 20, $l = 50$ , $hf = 30$ , $N = 30$ | -0.081                           | 0.010  | -3.185             | 40                                  |
| Nr. vetores de coint. = 27, $l = 50$ , $hf = 30$ , $N = 30$ | -0.102                           | 0.011  | -2.598             | 40                                  |

Tabela 17 – Resumo da estratégia (5) no portfólio de três fatores

| Parâmetros                                       | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|--|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $hf = 1$ , $N = 30$   | 0.307                            | 0.013  | 7.428              | 1217                                |
| Nr. vetores de coint. = 10, $hf = 1$ , $N = 30$  | 0.079                            | 0.013  | 23.426             | 1217                                |
| Nr. vetores de coint. = 20, $hf = 1$ , $N = 30$  | 0.080                            | 0.013  | 21.235             | 1217                                |
| Nr. vetores de coint. = 27, $hf = 1$ , $N = 30$  | 0.184                            | 0.014  | 9.032              | 1217                                |
| Nr. vetores de coint. = 1, $hf = 10$ , $N = 30$  | 0.501                            | 0.011  | 0.804              | 121                                 |
| Nr. vetores de coint. = 10, $hf = 10$ , $N = 30$ | -0.460                           | 0.014  | -0.846             | 121                                 |
| Nr. vetores de coint. = 20, $hf = 10$ , $N = 30$ | -0.383                           | 0.014  | -0.973             | 121                                 |
| Nr. vetores de coint. = 27, $hf = 10$ , $N = 30$ | -0.257                           | 0.014  | -1.393             | 121                                 |
| Nr. vetores de coint. = 1, $hf = 30$ , $N = 30$  | -0.158                           | 0.012  | -1.763             | 40                                  |
| Nr. vetores de coint. = 10, $hf = 30$ , $N = 30$ | -0.795                           | 0.013  | -0.342             | 40                                  |
| Nr. vetores de coint. = 20, $hf = 30$ , $N = 30$ | -0.139                           | 0.015  | -1.949             | 40                                  |
| Nr. vetores de coint. = 27, $hf = 30$ , $N = 30$ | -0.350                           | 0.016  | -0.743             | 40                                  |

### B.3 Portfólio composto por PETR3 e PETR4

Tabela 18 – Resumo da estratégia (1) no portfólio composto por PETR3 e PETR4

| Parâmetros                          | Ganhos / Capital<br>negociado | Desv. padrão dos<br>ganhos / Capital<br>negociado | Custos / Ganhos | Nr. de transações<br>completadas |
|-------------------------------------|-------------------------------|---|-----------------|----------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ | 0.084                         | 0.008   | 0.728           | 2                                |
| Nr. vetores de coint. = 1, $k = 30$ | -0.054                        | 0.003   | -1.133          | 2                                |
| Nr. vetores de coint. = 1, $k = 50$ | -0.226                        | 0.006   | -0.351          | 8                                |

Tabela 19 – Resumo da estratégia (2) no portfólio composto por PETR3 e PETR4

| Parâmetros                          | Ganhos / Capital negociado | Desv. padrão dos ganhos / Capital negociado | Custos / Ganhos | Nr. de transações completadas |
|-------------------------------------|----------------------------|---|-----------------|-------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ | -0.020                     | 0.008                                       | -2.680          | 3                             |
| Nr. vetores de coint. = 1, $k = 20$ | -0.058                     | 0.008                                       | -1.107          | 9                             |
| Nr. vetores de coint. = 1, $k = 30$ | 0.016                      | 0.014                                       | 5.417           | 27                            |

Tabela 20 – Resumo da estratégia (3) no portfólio composto por PETR3 e PETR4

| Parâmetros                                      | Ganhos / Capital negociado | Desv. padrão dos ganhos / Capital negociado | Custos / Ganhos | Nr. de transações completadas |
|---|----------------------------|---|-----------------|-------------------------------|
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$  | -1.233                     | 0.007                                       | -1.394          | 1214                          |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ | -0.532                     | 0.006                                       | -0.355          | 120                           |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ | -1.038                     | 0.007                                       | -0.148          | 39                            |

Tabela 21 – Resumo da estratégia (4) no portfólio composto por PETR3 e PETR4

| Parâmetros                                      | Ganhos / Capital negociado | Desv. padrão dos ganhos / Capital negociado | Custos / Ganhos | Nr. de transações completadas |
|---|----------------------------|---|-----------------|-------------------------------|
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$  | 0.044                      | 0.007                                       | 1.728           | 1214                          |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ | 0.069                      | 0.006                                       | 0.789           | 120                           |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ | 0.100                      | 0.007                                       | 0.479           | 39                            |

Tabela 22 – Resumo da estratégia (5) no portfólio composto por PETR3 e PETR4

| Parâmetros                           | Ganhos / Capital negociado | Desv. padrão dos ganhos / Capital negociado | Custos / Ganhos | Nr. de transações completadas |
|--------------------------------------|----------------------------|---|-----------------|-------------------------------|
| Nr. vetores de coint. = 1, $hf = 1$  | -1.233                     | 0.007                                       | -1.394          | 1217                          |
| Nr. vetores de coint. = 1, $hf = 10$ | -0.532                     | 0.006                                       | -0.355          | 121                           |
| Nr. vetores de coint. = 1, $hf = 30$ | -1.038                     | 0.007                                       | -0.148          | 40                            |

## B.4 Portfólio com ativos do setor bancário

Tabela 23 – Resumo da estratégia (1) no portfólio do setor bancário

| Parâmetros                          | Ganhos / Capital negociado | Desv. padrão dos ganhos / Capital negociado | Custos / Ganhos | Nr. de transações completadas |
|-------------------------------------|----------------------------|---|-----------------|-------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ | 0.492                      | 0.005                                       | 0.145           | 17                            |
| Nr. vetores de coint. = 2, $k = 10$ | -0.003                     | 0.006                                       | -17.875         | 8                             |
| Nr. vetores de coint. = 3, $k = 10$ | 0.480                      | 0.008                                       | 0.124           | 6                             |
| Nr. vetores de coint. = 4, $k = 10$ | 0.135                      | 0.006                                       | 0.380           | 3                             |
| Nr. vetores de coint. = 1, $k = 30$ | 1.281                      | 0.006                                       | 0.130           | 58                            |
| Nr. vetores de coint. = 2, $k = 30$ | 0.584                      | 0.008                                       | 0.179           | 31                            |
| Nr. vetores de coint. = 3, $k = 30$ | 0.663                      | 0.007                                       | 0.144           | 22                            |
| Nr. vetores de coint. = 4, $k = 30$ | 0.814                      | 0.057                                       | 0.121           | 20                            |
| Nr. vetores de coint. = 1, $k = 50$ | 1.268                      | 0.005                                       | 0.411           | 202                           |
| Nr. vetores de coint. = 2, $k = 50$ | 0.400                      | 0.008                                       | 0.907           | 137                           |
| Nr. vetores de coint. = 3, $k = 50$ | 0.802                      | 0.007                                       | 0.312           | 90                            |
| Nr. vetores de coint. = 4, $k = 50$ | 0.822                      | 0.073                                       | 0.295           | 78                            |

Tabela 24 – Resumo da estratégia (2) no portfólio do setor bancário

| Parâmetros                          | Ganhos / Capital<br>negociado | Desv. padrão dos<br>ganhos / Capital<br>negociado | Custos / Ganhos | Nr. de transações<br>completadas |
|-------------------------------------|-------------------------------|---|-----------------|----------------------------------|
| Nr. vetores de coint. = 1, $k = 10$ | 0.647                         | 0.005   | 0.128           | 49                               |
| Nr. vetores de coint. = 2, $k = 10$ | 0.214                         | 0.008   | 0.256           | 24                               |
| Nr. vetores de coint. = 3, $k = 10$ | 0.227                         | 0.010   | 0.249           | 23                               |
| Nr. vetores de coint. = 4, $k = 10$ | 0.237                         | 0.011   | 0.220           | 15                               |
| Nr. vetores de coint. = 1, $k = 20$ | 0.885                         | 0.004   | 0.187           | 111                              |
| Nr. vetores de coint. = 2, $k = 20$ | 0.297                         | 0.008   | 0.272           | 45                               |
| Nr. vetores de coint. = 3, $k = 20$ | 0.409                         | 0.011   | 0.275           | 61                               |
| Nr. vetores de coint. = 4, $k = 20$ | 0.623                         | 0.023   | 0.140           | 41                               |
| Nr. vetores de coint. = 1, $k = 30$ | 1.084                         | 0.005   | 0.286           | 223                              |
| Nr. vetores de coint. = 2, $k = 30$ | 0.230                         | 0.009   | 0.723           | 110                              |
| Nr. vetores de coint. = 3, $k = 30$ | 0.620                         | 0.011   | 0.375           | 155                              |
| Nr. vetores de coint. = 4, $k = 30$ | 0.736                         | 0.025   | 0.218           | 95                               |

Tabela 25 – Resumo da estratégia (3) no portfólio do setor bancário

| Parâmetros                                      | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 1$  | 1.139                            | 0.015  | 1.894              | 1217                                |
| Nr. vetores de coint. = 2, $l = 30$ , $hf = 1$  | 1.259                            | 0.015  | 1.624              | 1216                                |
| Nr. vetores de coint. = 3, $l = 30$ , $hf = 1$  | 1.281                            | 0.015  | 1.564              | 1216                                |
| Nr. vetores de coint. = 4, $l = 30$ , $hf = 1$  | 1.133                            | 0.015  | 1.734              | 1217                                |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$  | 0.699                            | 0.010  | 2.408              | 1217                                |
| Nr. vetores de coint. = 2, $l = 50$ , $hf = 1$  | 0.726                            | 0.010  | 2.129              | 1217                                |
| Nr. vetores de coint. = 3, $l = 50$ , $hf = 1$  | 0.864                            | 0.010  | 1.691              | 1217                                |
| Nr. vetores de coint. = 4, $l = 50$ , $hf = 1$  | 0.560                            | 0.010  | 2.505              | 1217                                |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 10$ | 1.217                            | 0.011  | 0.161              | 120                                 |
| Nr. vetores de coint. = 2, $l = 30$ , $hf = 10$ | 0.750                            | 0.013  | 0.251              | 120                                 |
| Nr. vetores de coint. = 3, $l = 30$ , $hf = 10$ | 0.861                            | 0.014  | 0.190              | 120                                 |
| Nr. vetores de coint. = 4, $l = 30$ , $hf = 10$ | 1.458                            | 0.015  | 0.115              | 120                                 |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ | 0.993                            | 0.009  | 0.171              | 120                                 |
| Nr. vetores de coint. = 2, $l = 50$ , $hf = 10$ | 0.829                            | 0.009  | 0.181              | 120                                 |
| Nr. vetores de coint. = 3, $l = 50$ , $hf = 10$ | 0.758                            | 0.009  | 0.178              | 120                                 |
| Nr. vetores de coint. = 4, $l = 50$ , $hf = 10$ | 0.682                            | 0.010  | 0.189              | 120                                 |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 30$ | 0.433                            | 0.008  | 0.212              | 40                                  |
| Nr. vetores de coint. = 2, $l = 30$ , $hf = 30$ | 0.359                            | 0.012  | 0.260              | 40                                  |
| Nr. vetores de coint. = 3, $l = 30$ , $hf = 30$ | -0.205                           | 0.010  | -0.469             | 40                                  |
| Nr. vetores de coint. = 4, $l = 30$ , $hf = 30$ | 0.871                            | 0.014  | 0.112              | 40                                  |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ | 0.568                            | 0.008  | 0.153              | 39                                  |
| Nr. vetores de coint. = 2, $l = 50$ , $hf = 30$ | 0.262                            | 0.008  | 0.335              | 39                                  |
| Nr. vetores de coint. = 3, $l = 50$ , $hf = 30$ | 0.219                            | 0.008  | 0.390              | 40                                  |
| Nr. vetores de coint. = 4, $l = 50$ , $hf = 30$ | 0.196                            | 0.009  | 0.404              | 40                                  |

Tabela 26 – Resumo da estratégia (4) no portfólio do setor bancário

| Parâmetros                                      | Ganhos /<br>Capital<br>negociado | Desv. padrão<br>dos ganhos /<br>Capital<br>negociado | Custos /<br>Ganhos | Nr. de<br>transações<br>completadas |
|---|----------------------------------|--|--------------------|-------------------------------------|
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 1$  | -2.095                           | 0.015  | -0.849             | 1217                                |
| Nr. vetores de coint. = 2, $l = 30$ , $hf = 1$  | -2.067                           | 0.015  | -0.866             | 1216                                |
| Nr. vetores de coint. = 3, $l = 30$ , $hf = 1$  | -2.213                           | 0.016  | -0.821             | 1216                                |
| Nr. vetores de coint. = 4, $l = 30$ , $hf = 1$  | -2.256                           | 0.016  | -0.815             | 1217                                |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 1$  | -3.328                           | 0.012  | -0.639             | 1217                                |
| Nr. vetores de coint. = 2, $l = 50$ , $hf = 1$  | -3.069                           | 0.012  | -0.674             | 1217                                |
| Nr. vetores de coint. = 3, $l = 50$ , $hf = 1$  | -3.207                           | 0.013  | -0.637             | 1217                                |
| Nr. vetores de coint. = 4, $l = 50$ , $hf = 1$  | -3.221                           | 0.013  | -0.640             | 1217                                |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 10$ | -0.671                           | 0.011  | -0.248             | 120                                 |
| Nr. vetores de coint. = 2, $l = 30$ , $hf = 10$ | -0.191                           | 0.013  | -0.928             | 120                                 |
| Nr. vetores de coint. = 3, $l = 30$ , $hf = 10$ | 0.129                            | 0.014  | 1.310              | 120                                 |
| Nr. vetores de coint. = 4, $l = 30$ , $hf = 10$ | 0.174                            | 0.015  | 1.036              | 120                                 |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 10$ | -0.500                           | 0.011  | -0.425             | 121                                 |
| Nr. vetores de coint. = 2, $l = 50$ , $hf = 10$ | -0.315                           | 0.010  | -0.682             | 121                                 |
| Nr. vetores de coint. = 3, $l = 50$ , $hf = 10$ | -0.271                           | 0.010  | -0.748             | 121                                 |
| Nr. vetores de coint. = 4, $l = 50$ , $hf = 10$ | 0.048                            | 0.010  | 4.096              | 121                                 |
| Nr. vetores de coint. = 1, $l = 30$ , $hf = 30$ | -0.250                           | 0.008  | -0.299             | 40                                  |
| Nr. vetores de coint. = 2, $l = 30$ , $hf = 30$ | -0.248                           | 0.012  | -0.345             | 40                                  |
| Nr. vetores de coint. = 3, $l = 30$ , $hf = 30$ | -0.634                           | 0.010  | -0.156             | 40                                  |
| Nr. vetores de coint. = 4, $l = 30$ , $hf = 30$ | -0.482                           | 0.014  | -0.181             | 40                                  |
| Nr. vetores de coint. = 1, $l = 50$ , $hf = 30$ | 0.083                            | 0.011  | 1.142              | 40                                  |
| Nr. vetores de coint. = 2, $l = 50$ , $hf = 30$ | 0.287                            | 0.009  | 0.367              | 40                                  |
| Nr. vetores de coint. = 3, $l = 50$ , $hf = 30$ | -0.459                           | 0.010  | -0.237             | 40                                  |
| Nr. vetores de coint. = 4, $l = 50$ , $hf = 30$ | -0.643                           | 0.009  | -0.152             | 39                                  |

Tabela 27 – Resumo da estratégia (5) no portfólio do setor bancário

| Parâmetros                           | Ganhos / Capital<br>negociado | Desv. padrão dos<br>ganhos / Capital<br>negociado | Custos / Ganhos | Nr. de transações<br>completadas |
|--------------------------------------|-------------------------------|---|-----------------|----------------------------------|
| Nr. vetores de coint. = 1, $hf = 1$  | 1.019                         | 0.011   | 1.865           | 1217                             |
| Nr. vetores de coint. = 2, $hf = 1$  | 0.996                         | 0.011   | 1.751           | 1217                             |
| Nr. vetores de coint. = 3, $hf = 1$  | 1.093                         | 0.012   | 1.546           | 1217                             |
| Nr. vetores de coint. = 4, $hf = 1$  | 1.074                         | 0.012   | 1.543           | 1217                             |
| Nr. vetores de coint. = 1, $hf = 10$ | 1.033                         | 0.009   | 0.172           | 121                              |
| Nr. vetores de coint. = 2, $hf = 10$ | 0.867                         | 0.011   | 0.180           | 121                              |
| Nr. vetores de coint. = 3, $hf = 10$ | 0.731                         | 0.012   | 0.197           | 121                              |
| Nr. vetores de coint. = 4, $hf = 10$ | 1.076                         | 0.012   | 0.130           | 121                              |
| Nr. vetores de coint. = 1, $hf = 30$ | 0.337                         | 0.006   | 0.259           | 40                               |
| Nr. vetores de coint. = 2, $hf = 30$ | 0.009                         | 0.009   | 9.687           | 40                               |
| Nr. vetores de coint. = 3, $hf = 30$ | 0.057                         | 0.011   | 1.548           | 40                               |
| Nr. vetores de coint. = 4, $hf = 30$ | 0.386                         | 0.012   | 0.218           | 40                               |