

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE CIÊNCIAS ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA**

Joao Fróis Caldeira

Ensaio em Econometria Financeira

**Porto Alegre
2010**

Joao Fróis Caldeira

Ensaio em Econometria Financeira

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Doutor em Economia, com ênfase em economia aplicada.

Orientador: Marcelo Savino Portugal

**Porto Alegre
2010**

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO
(CIP) Responsável: Biblioteca Gládis W. do Amaral, Faculdade de Ciências Econômicas da
UFRGS

C146e Caldeira, João Fróis
Ensaaios em econometria financeira / João Fróis Caldeira. - Porto Alegre, 2010.
000 f. : il.

Orientador: Marcelo Savino Portugal

Ênfase em Economia Aplicada.

Tese (Doutorado em Economia) - Universidade Federal do Rio Grande do Sul,
Faculdade de Ciências Econômicas, Programa de Pós-Graduação em Economia,
Porto Alegre, 2010.

1. Mercado financeiro : Modelo econométrico : Volatilidade. 2. Econometria :
Modelo de previsão. I. Portugal, Marcelo Savino. II. Universidade Federal do Rio
Grande do Sul. Faculdade de Ciências Econômicas. Programa de Pós-Graduação em
Economia. III. Título.

CDU 336.761

Joao Fróis Caldeira

ENSAIOS EM ECONOMETRIA FINANCEIRA

Tese submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como quesito parcial para obtenção do título de Doutor em Economia, com ênfase em economia aplicada.

Aprovada em: Porto Alegre, ____ de ____ de 2010.

Prof. Dr. Marcos Eugênio da Silva
USP

Prof. Dr. Gilberto de Oliveira Kloechner
UFRGS

Prof. Dr. Paulo Renato Terra Soares
UFRGS

Prof. Dr. Marcelo Savino Portugal
UFRGS

Agradecimentos

Agradeço a todos que me ajudaram. Obrigado.

Resumo

Os modelos de otimização de carteiras baseados na análise média-variância apresentam dificuldades para estimação das matrizes de covariância, usadas no processo de otimização, o que leva a necessidade de métodos *ad hoc* para limitar ou suavizar as alocações eficientes recomendadas pelo modelo. Embora as carteiras obtidas por este método sejam eficientes, não é assegurado que o *tracking error* seja estacionário, podendo a carteira se distanciar do benchmark, exigindo frequentes recomposições. Neste artigo é empregada a metodologia de cointegração para otimização de carteiras no âmbito de duas estratégias: *index tracking* e estratégia *long-short*. A estabilidade das carteiras otimizadas através da cointegração em diferentes cenários de mercado, diminuindo custos relativos a frequentes recomposições da carteira, e níveis de retorno e volatilidade superiores aos benchmarks, mostram que a metodologia é uma ferramenta eficiente e capaz de gerar resultados robustos, se caracterizando como uma atraente ferramenta para a gestão quantitativa de recursos. Modelar a estrutura a termo da taxa de juros é extremamente importante para macroeconomistas e participantes do mercado financeiro em geral. Neste artigo é empregada a formulação de Diebold-Li para ajustar e fazer previsões da estrutura a termo da taxa de juros brasileira. São empregados dados diários referentes às taxas dos contratos de DI Futuro negociados na BM&F que apresentaram maior liquidez para o período de Janeiro de 2006 a Fevereiro de 2009. Diferentemente da maior parte da literatura sobre curva de juros para dados brasileiros, em que o modelo de Diebold-Li é estimado pelo método de dois passos, neste trabalho o modelo é colocado no formato de estado espaço, e os parâmetros são estimados simultaneamente, de forma eficiente, pelo Filtro de Kalman. Os resultados obtidos tanto para o ajuste, mas principalmente no que diz respeito à previsão, mostram que a estimação do modelo através do Filtro de Kalman é a mais adequada, gerando melhores previsões para todas as maturidades quando é considerado horizontes de previsão de um mês, três meses e seis meses. No terceiro artigo nós propomos estimar o modelo dinâmico da estrutura a termo da curva de juros de Nelson e Siegel (1987) considerando duas especificações alternativas. Na primeira, nós consideramos os pesos dos fatores como variantes no tempo e tratamos a heterocedasticidade condicional via um modelo volatilidade estocástica com fatores comuns. No segundo caso, consideramos um modelo onde os fatores latentes seguem individualmente processos autoregressivos com volatilidade estocástica. Os assim chamados fatores de volatilidade buscam capturar a incerteza ao longo do tempo associada ao nível, inclinação e curvatura da curva de juros. A estimação é realizada através de métodos de inferência bayesiana, por Markov Chain Monte Carlo. Os resultados mostram que os fatores de volatilidade são altamente persistentes, dando suporte ao fato estilizado de que os choques na volatilidade das taxas de juros são altamente persistentes, e também indicam que o uso de estruturas de volatilidade estocástica levam a melhores ajustes dentro da amostra para a curva de juros observada.

Palavras-chave: Cointegração. *Index Tracking*. *Long-Short*. Estratégia Neutra ao Mercado

Abstract

The traditional models to optimize portfolios based on mean-variance analysis aim to determine the portfolio weights that minimize the variance for a certain return level. The covariance matrices used to optimize are difficult to estimate and ad hoc methods often need to be applied to limit or smooth the mean-variance efficient allocations recommended by the model. Although the method is efficient, the tracking error isn't certainly stationary, so the portfolio can get distant from the benchmark, requiring frequent re-balancements. We used the cointegration methodology to devise two quantitative strategies: index tracking and long-short market neutral. We aim to design optimal portfolios acquiring the asset prices' co-movements. We used Ibovespa's index and stocks from Jan-2000 to Dec-2008. The results show that the devise of index tracking portfolios using cointegration generates goods results, replicating the benchmark's return and volatility. The long-short strategy generated stable returns under several market circumstances, presenting low volatility. Modeling the term structure of interest rate is very important to macroeconomists and financial market practitioners in general. In this paper, we used the Diebold-Li interpretation to the Nelson Siegel model in order to fit and forecast the Brazilian yield curve. The data consisted of daily observations of the most liquid future ID yields traded in the BM&F from January 2006 to February 2009. Differently from the literature on the Brazilian yield curve, where the Diebold-Li model is estimated through the two-step method, the model herein is put in the state-space form, and the parameters are simultaneously and efficiently estimated using the Kalman filter. The results obtained for the fit and for the forecast showed that the Kalman filter is the most suitable method for the estimation of the model, generating better forecast for all maturities when we consider the forecasting horizons of one and three months. In the third essay we propose to estimate the dynamic Nelson-Siegel model of yield curve considering two alternative specifications. At first, we consider the factor loadings such as time-varying conditional heteroskedasticity and treat via a common factors of stochastic volatility models. In the second case, we consider a model where the latent factors individually following autorregressive process with stochastic volatility. The volatility factors seek to capture the uncertainty over time associated with level, slope and curvature of yield curve. The estimation is performed through bayesian inference, Markov Chain Monte Carlo. The volatility factors showed high persistence, supporting the stylized fact that shocks in the volatility of interest rate are highly persistent, and also indicate that the used of structures of stochastic volatility lead to better in-sample fits of the observed yield curve.

Keywords: Cointegration. Index Tracking. Long Short. Market Neutral Strategy

Sumário

1	Introdução	11
1.1	Contexto e Motivação	11
1.2	Conteúdo	13
2	Estratégias Long-Short, Neutras ao Mercado, e Index Tracking Baseadas em Portfólios Cointegrados	15
2.1	Introdução	15
2.2	Estratégias Neutras ao Mercado <i>Long-Short</i> e Carteiras <i>Index Tracking</i>	20
2.3	Dados	23
2.4	Construção das Carteiras	26
2.5	Avaliação Empírica das Estratégias com Dados do Mercado Brasileiro	28
2.6	Conclusões	40
3	Estimação e Previsão Eficientes da Curva Juros para Dados do Brasil	42
3.1	Introdução	42
3.2	Modelo de Fatores da Curva de Juros	44
3.3	Função de Desconto, Curva <i>Forward</i> e Curva de Juros	44
3.4	O Modelo da Curva de Juros de Diebold e Li	46
3.5	Estimação e previsão usando o formato espaços de estados	50
3.6	Dados e Análise dos Resultados	51
3.7	Dados	52
3.8	Avaliação Empírica do Modelo para Dados do DI-Futuro	53
3.9	Conclusão	61
4	Inferência Bayesiana Aplicada ao Modelo Dinâmico de Nelson-Siegel com Volatilidade Estocástica nos Fatores	65
4.1	Introdução	65
4.2	Modelo de fatores de Nelson e Siegel para estrutura a termo da taxa de juros	70
4.3	Modelos de Nelson-Siegel Generalizados	71
4.4	Estimação Bayesiana baseado em MCMC	74
4.5	Dados, Estimação e Resultados	77
4.6	Dados	77
4.7	Análise dos Resultados	78
4.8	Conclusão	87
5	Conclusão	89
	Referências	92
	APÊNDICE A - Coeficientes de correlação de Spearman	98
	APÊNDICE B - Composição das carteiras	103

1 Introdução

Nos anos recentes tem se observado a expansão de uma área de pesquisa multidisciplinar conhecida como "Finanças Computacionais", impulsionados de um lado pelos avanços computacionais, disponibilidade de dados e métodos de modelagem econométricos. Por outro lado, a competitividade e contínua evolução natural dos mercados financeiros, com destaque para o crescimento da indústria de *hedge funds*. O desenvolvimento da indústria de *hedge funds* e acirramento da concorrência impulsionam a busca por alta rentabilidade com baixos níveis de risco, criando uma demanda por melhores técnicas de modelagem da dinâmica do mercado. Já os avanços computacionais, maior disponibilidade de dados e metodologia cria uma plataforma fértil para desenvolvimento e oferta de tais modelos e ferramentas para o mercado.

1.1 Contexto e Motivação

A filosofia que norteia esta tese é que os recentes avanços nas técnicas de modelagem econométrica e computacionais, juntamente com o desenvolvimento da indústria de *hedge funds*, abrem novas oportunidades para identificar e modelar regularidades nas dinâmicas dos preços dos ativos. Assim, a motivação principal é empregar técnicas econométricas no desenvolvimento de estratégias quantitativas para alocação de recursos em *hedge funds*, e na modelagem e previsão da estrutura a termo da taxa de juros.

Hedge funds são geralmente considerados como instrumentos de investimentos privados para grandes aplicadores ou investidores institucionais, tipicamente organizados como sociedades limitadas, que podem assumir posições compradas (*long*) e vendidas (*short*) em vários mercados, usando diversas estratégias de investimento. Se por um lado a definição é precisa, por outro é muito ampla, tendo em vista que estes fundos podem usar vários tipos de títulos em vários mercados. Esta parte da definição é muito mais aberta e permite que uma grande quantidade de fundos seja classificada como *hedge fund*, desde que assuma posições compradas e vendidas¹. Anson (2006) define *hedge fund* como veículos de investimentos privados para gerenciar carteiras concentradas de títulos públicos e privados, e seus respectivos instrumentos de derivativos, que podem assumir posições tanto compradas quanto vendidas e podem se utilizar de alavancagem. Investimentos em *hedge funds* requerem carteiras diversificadas e bem construídas, com diferentes estratégias para assegurar que os riscos, tanto a nível individual quanto sistêmico sejam minimizados. As diferentes estratégias dos *hedge funds* são expostas a diferentes riscos sistêmicos, e portanto, os retornos mudam com mudanças no ambiente de mercado. Esses diferentes parâmetros de risco/retorno fornecem possibilidades

¹*Hedge funds* investem em um conjunto heterogêneo de ativos financeiros: ações, títulos, swaps, derivativos sofisticados, moedas, títulos hipotecários, conversíveis, entre outros.

quase infinitas para criar valor através de alocações em carteiras ativas, onde as estratégias interagem para gerar consistentes retornos absolutos.

De acordo com Caldwell (1995), o primeiro hedge fund foi constituído por Alber Wislow Jones em 1949. A principal estratégia usada consistia em posições compradas e vendidas em ações (*long-short*) e alavancagem. O fundo também adotava um sistema de taxa de performance. Os hedge funds se mantiveram relativamente obscuros para o mundo do investimento até 1966, quando um artigo da Fortune escreveu que os fundos de Jones tinham retornos (líquidos) substancialmente maiores que os fundos mútuos com melhores performances (ver Loomis, 1966). Isto levou a uma onda de interesse por hedge funds e muitos foram constituídos nos dois anos seguintes. Segundo Caldwell (1995), a SEC afirma em levantamento que no fim de 1968 existiam 215 sociedades de investimento, e concluiu que 140 deles eram hedge funds, sendo a maioria aberta naquele ano.

Depois da rápida expansão de 1967-68, a indústria de *hedge funds* experimentou um substancial revés durante os anos de 1969-70 e 1973-74, quando muitos fundos sofreram expressivas perdas e fuga de capital. Novamente os *hedge funds* caíram na obscuridade até 1986, quando um artigo da *Institutional Investor* reportou que o fundo Julian Robeston's Tiger obteve retornos líquidos anuais de 43% durante seus primeiros seis anos de existência. Isto reacendeu o interesse pelos hedge funds, com a formação de muitos novos *hedge funds*.

Desde o início dos anos de 1990 quando existia aproximadamente 2000 *hedge funds*, administrando ativos da ordem de 60 bilhões dólares, o subsequente crescimento no número e na base patrimonial dos *hedge funds* é expressiva. A indústria sofreu uma redução na velocidade de crescimento em 1998, desde então desfrutou de uma vitalidade renovada, com o número total estimado de fundos de 10000, administrando mais que um trilhão de dólares no final de 2006. Mesmo durante o colapso do mercado de ações iniciado em março de 2000, quando o índice NASDAQ Composite saiu de 5132 alcançando 1253 pontos três anos mais tarde, o volume total administrado pelos *hedge funds* continuou a crescer a uma taxa estável de 10,6% (*Van Hedge Funds Advisors International*, 2002), contrastando com um recuo mundial da indústria de fundos mútuos de 2,7% (*Investment Company Institute*, 2003). Mais recentemente, Capocci D. (2004) estimaram que em 2001 existiam 6000 *hedge funds* administrando aproximadamente 400 bilhões de dólares. Em 2007, Capocci (2007) estima que existiam 10000 *hedge funds* administrando em torno de 1 trilhão de dólares. Isto representa um crescimento de 11% no número de fundos e 26% no volume de ativos sob gestão durante seis anos.

Desenvolver estratégias de alocação baseadas em métodos quantitativos, e técnicas para fazer previsões de séries econômicas e financeiras assumem papel relevante nesse contexto e tem recebido cada mais atenção de pesquisadores e participantes dos mercados em geral. Em grande parte isto se deve ao desenvolvimento da indústria de *hedge funds*, citado acima, abrindo um vasto campo para pesquisas acadêmicas e aplicadas. Por outro lado, os avanços computacionais e das técnicas econométricas verificados nos últimos anos, propiciam as ferramentas necessárias para esta crescente área.

1.2 Conteúdo

Neste sentido, esta tese consiste de ensaios em econometria e finanças, tendo como objetivo empregar técnicas econométricas avançadas no desenvolvimento de estratégias quantitativas de alocação de recursos no contexto de *hedge funds*. Também são empregados os métodos do filtro de Kalman e técnicas de inferência bayesiana Markov chain Monte Carlo - MCMC para ajustar e fazer previsões da estrutura a termo da taxa de juros, com base no modelo clássico da curva de juros de Nelson e Siegel. Assim, a tese assume um caráter interdisciplinar, envolvendo a aplicação de técnicas econométricas no desenvolvimento de estratégias para o mercado financeiro e no ajuste e previsão da estrutura a termo da taxa de juros, que são do interesse tanto de economistas, formuladores de políticas econômicas e de participantes do mercado em geral.

A tese é composta por três ensaios. No primeiro ensaio: "Cointegração e Estratégias de Alocação no Mercado de Ações Brasileiro: Carteiras Index Tracking e Estratégias Neutras ao Mercado Long-Short"^{2, 3} nós propomos o uso da metodologia de cointegração na otimização de carteiras no âmbito de duas estratégias de alocação para *hedge funds*⁴.

No segundo ensaio "Estimação e Previsão Eficientes da Curva de Juros com Dados do Brasil"⁵⁶, nós empregamos dados dos contratos de DI-futuro negociados na BM&F e propomos uma especificação do modelo da estrutura a termo de Nelson e Siegel no formato de espaços de estados e empregamos o algoritmo do filtro de Kalman, de tal forma que o parâmetro que determina o peso dos fatores é estimado conjuntamente com os demais parâmetros, não sendo determinado a priori por algum critério.

Já no terceiro ensaio, "O Modelo Dinâmico da Estrutura a Termo de Nelson e Siegel com Volatilidade Estocástica nos Fatores", nosso foco principal é o tratamento da heterocedasticidade condicional presente na estrutura a termo das taxas de juros. Assim, nós apresentamos uma generalização das estruturas de volatilidade apresentadas em Koopman et al. (2010), Laurini and Hotta (2008) e Hautsch and Ou (2009). Utilizando dados dos contratos de DI-Futuro negociados na BM&F, nós estimaremos o modelo dinâmico de Nelson e Siegel com os pesos dos fatores e volatilidade variando no tempo. Nós consideramos duas especificações para modelar a volatilidade condicional. Na primeira, análoga a utilizada em Koopman et al. (2010) o grau de risco inerente à estrutura a termo é capturado via um componente comum de volatilidade agregada. Na segunda alternativa propomos modelar diretamente a volatilidade estocástica de cada fator latente do modelo de Nelson e Siegel. Assim, o terceiro ensaio avança em relação ao segundo na medida em que o modelo de fatores de Nelson e Siegel da

² Artigo revisado e reenviado para publicação na Revista Brasileira de Finanças.

³ Esse artigo se beneficiou de comentários decorrentes de apresentações no XXXI Encontro Brasileiro de Econometria, Research in Options 2009 e IX Encontro Brasileiro de Finanças.

⁴ No Brasil, é comum usar a expressão fundos multimercado em substituição a *hedge funds*.

⁵ Aceito para publicação na Revista EconomiA da ANPEC.

⁶ Esse artigo se beneficiou de comentários decorrentes de apresentações no XXXVII Encontro Nacional da ANPEC e IX Encontro Brasileiro de Finanças.

curva de juros, é estimado como um modelo com quatro fatores latentes estados, em que o parâmetro que determina o peso dos fatores é tratado como um quarto fator latente, sendo estimado conjuntamente com os demais fatores. Além disso, nós consideramos duas especificações para a volatilidade condicional, e empregamos inferência bayesiana, Markov chain Monte Carlo, para estimar o modelo em formato não-linear. Vale destacar, porém, que neste caso, o principal objetivo não é fazer previsões para estrutura a termo da curva de juros, mas sim analisar o ajuste e modelar a volatilidade inerente à estrutura a termo.

2 Estratégias Long-Short, Neutras ao Mercado, e Index Tracking Baseadas em Porfólios Cointegrados

Os modelos de otimização de carteiras baseados na análise média-variância apresentam dificuldades para estimação das matrizes de covariância, usadas no processo de otimização, o que leva a necessidade de métodos *ad hoc* para limitar ou suavizar as alocações eficientes recomendadas pelo modelo. Embora as carteiras obtidas por este método sejam eficientes, não é assegurado que o *tracking error* seja estacionário, podendo a carteira se distanciar do benchmark, exigindo frequentes recomposições. Neste artigo é empregada a metodologia de cointegração para otimização de carteiras no âmbito de duas estratégias: *index tracking* e estratégia *long-short*. A estabilidade das carteiras otimizadas através da cointegração em diferentes cenários de mercado, diminuindo custos relativos a frequentes recomposições da carteira, e níveis de retorno e volatilidade superiores aos benchmarks, mostram que a metodologia é uma ferramenta eficiente e capaz de gerar resultados robustos, se caracterizando como uma atraente ferramenta para a gestão quantitativa de recursos.

2.1 Introdução

Os modelos tradicionais de otimização de carteiras baseados na análise média-variância buscam determinar a composição que minimize a volatilidade da carteira para um dado nível de retorno. Dificuldades com a estimação das matrizes de covariância, usadas no processo de otimização, levam à necessidade de métodos *ad hoc* para limitar ou suavizar as alocações eficientes recomendadas pelo modelo de média-variância. Embora as carteiras obtidas por este método sejam eficientes, não é assegurado que o tracking error seja estacionário, podendo a carteira se distanciar do benchmark, exigindo frequentes rebalanceamentos. Neste artigo empregamos a metodologia de cointegração para otimização de carteiras no âmbito de duas estratégias: *index tracking* e estratégia neutra ao mercado, *long-short*. Para a primeira estratégia, o objetivo principal é minimizar o risco de uma estratégia de investimento e entregar lucros sob todas as circunstâncias de mercado, podendo ser descrita como não-direcional ou neutra ao mercado. Busca-se correlação mínima com o mercado, e gerar um alfa positivo independente da direção do mercado. Para a segunda estratégia o objetivo é replicar um benchmark em termos de retorno e volatilidade, e entregar um alfa positivo. As estratégias apresentadas aqui usam as séries históricas dos preços para projetar o comportamento futuro de uma ação. A diferença entre a metodologia empregada e os modelos tradicionais é que a otimização é baseada na cointegração dos preços, ao invés da correlação dos retornos.

Estratégias *long-short* são amplamente empregadas nos mercados financeiros, seja por gestores de *hedge-funds* ou fundos de pensão, gestores quantitativos e participantes do mercado

em geral. Na maioria das vezes as estratégias *long-short* buscam obter retorno sem exposição direcional ao mercado, por isso são chamadas de estratégias neutras ao mercado. As estratégias *long-short* podem aparecer sob vários formatos: i) operações com pares de ações¹, também conhecidas como *pair trading*, que podem ser construídas tanto com base em modelos matemáticos que buscam capturar desvios das relações de longo prazo entre os preços de ativos, ou com base em análise fundamentalista; ii) operações envolvendo uma ação e um índice de mercado do setor ao qual a empresa pertença², ou algum índice geral de mercado; iii) por fim, operações envolvendo duas carteiras de ações, onde assume-se posição vendida em uma das carteiras - a parte *short*, e com os recursos obtidos com a venda dessa carteira assume-se posição comprada na outra - a parte *long*.

De uma forma geral, uma estratégia *long-short* consiste na manutenção de uma carteira comprada (*long*), financiada por uma carteira vendida (*short*). A carteira *short*, além de funcionar como um *hedge* contra riscos de declínios dos mercados fornece ao gestor a possibilidade de adicionar valor pelo uso de algum critério de seleção de ações que apresentam desempenho inferior ao desempenho da carteira *long*. Na prática, em estratégias *long-short* envolvendo duas carteiras, a parte *long* não necessariamente precisa apresentar desempenho superior ao mercado, mas sim precisa apresentar desempenho superior à parte *short*, para que a estratégia obtenha resultados positivos.

Por muitas décadas, gestores e participantes do mercado em geral avaliam os co-movimentos e interdependência entre os mercados e ativos em geral. Entretanto, a tradicional otimização de carteiras quantitativas fundamenta-se principalmente na análise da correlação (conceito estritamente linear) para modelar as mais complexas interdependências entre ativos financeiros. A busca por técnicas adequadas para a construção de carteiras que visam replicar um índice de referência, ou para construção de estratégias *long-short* neutras ao mercado, não é algo recente. Atualmente grande parte das pesquisas com métodos quantitativos para construção de tais estratégias está envolvida com aplicações para *hedge funds*, que devido a flexibilidade operacional e baixas restrições, são adequados para aplicação dessas estratégias de alocação.

Nos modelos padrão de risco e retorno as séries de preços são diferenciadas antes de iniciar a análise, removendo *a priori* qualquer tendência de longo prazo nos dados. As tendências ficam implícitas nas séries de retornos, mas qualquer decisão baseada em tendências comuns de longo prazo nas séries de preços é excluída em modelagens padrões de risco e retorno. Uma vez que as correlações são altamente instáveis, sua variabilidade ao longo do tempo torna-a uma ferramenta difícil para uso prático na gestão quantitativa de carteiras, pois os frequentes rebalanceamentos podem implicar em custos elevados.

O *input* mais importante dos modelos baseados em otimização por média-variância é o ve-

¹Para as operações de *pair trading* o mais comum são operações de ações ON contra PN, ou de empresas que pertencem a um mesmo setor, onde é mais factível que preços das ações tenham uma relação estável ao longo do tempo.

²As operações de ações contra índices setoriais, conhecidos como ETF's (Exchange Trade Funds) são pouco utilizadas no mercado brasileiro devido à existência de poucos ETF's para o mercado brasileiro.

tor de retornos esperados. Autores como, por exemplo, Best and Grauer (1985) demonstram que um pequeno aumento no retorno esperado de um dos ativos da carteira pode tirar metade dos ativos da carteira. Com o objetivo de definir adequados pontos de partida no processo de otimização de carteiras Black and Litterman (1990) e He and Litterman (1999) consideram várias alternativas de previsão dos retornos esperados: retornos históricos, médias de retornos iguais para todos os ativos médias de retornos ajustadas pelo risco iguais. Concluem, porém, que todas as alternativas de previsão consideradas levam a portfólios extremos, com posições altamente concentradas em poucos ativos, tanto compradas quanto vendidas. Black and Litterman (1990, 1992) propõem uma metodologia baseada na análise média-variância para criar portfólios eficientes estáveis, baseados na visão individual de cada investidor, a qual supera o problema de sensibilidade extrema aos *inputs*, (para detalhes ver Meucci, 2005; Lee, 2000; Black and Litterman, 1990, 1992; Idzorek, 2004). Também existe uma ampla literatura que emprega métodos bayesianos no processo de otimização de carteiras, (a esse respeito ver Rachev et al., 2000; Michaud and Michaud, 1998).

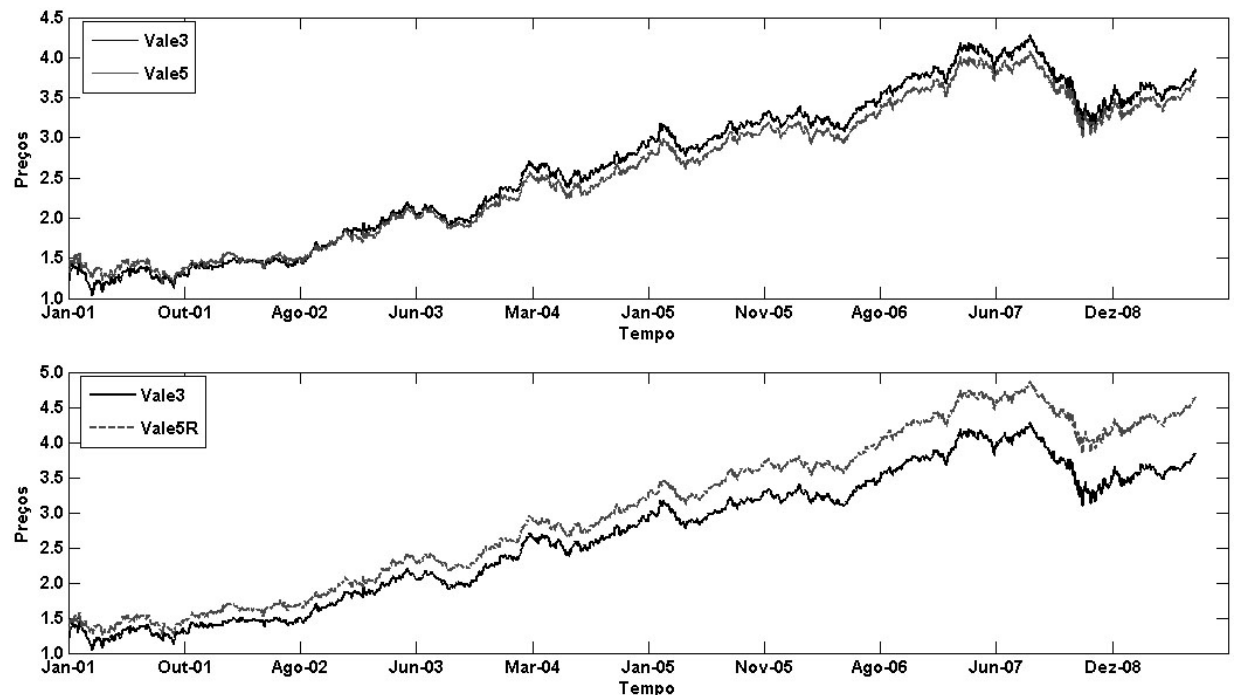
Porém, nos últimos anos, o conceito de cointegração tornou-se amplamente difundido na econometria de séries financeiras. A metodologia de cointegração tornou-se uma ferramenta extremamente poderosa, pois permite a aplicação de métodos de estimação simples para variáveis não estacionárias. Ainda assim, sua relevância para a análise de investimento tem sido limitada, principalmente pelo fato de que o padrão na gestão de carteiras e mensuração de risco se dá através da análise da correlação entre os retornos dos ativos.

Cointegração e correlação são conceitos relacionados, mas diferentes. Elevada correlação não implica em cointegração, bem como elevada cointegração não implica em alta correlação. Na verdade séries cointegradas podem apresentar momentos de baixa correlação. Uma carteira grande e diversificada composta por ações de algum índice, na qual as alocações são determinadas pelos respectivos pesos das ações no índice, deveria (a princípio) ser cointegrada com o índice. No longo prazo, a carteira deveria se mover junto com o índice. Contudo, podem ocorrer períodos em que as ações da carteira apresentam movimentos extremos nos preços, fazendo com que a correlação entre a carteira e o índice seja baixa nesse momento.

O exemplo a seguir, baseado em Alexander and Dimitriu (2002), ilustra que a cointegração não implica em correlação e vice-versa. Alta correlação dos retornos não implica em alta cointegração nos preços. A figura 2.1 apresenta, no primeiro painel, as séries de preços diários (em ln) das ações preferenciais e ordinárias da CVRD (VALE5 e VALE3), de janeiro de 2000 a outubro de 2009. Testes de cointegração revelam que as séries são cointegradas a 1% e os retornos diários das duas séries apresentam coeficiente de correlação de 0.99. No segundo painel são apresentados logaritmos da série de preços da VALE3 e da VALE5R³, testes de cointegração mostram que essas duas séries não são cointegradas nem a 10% de significância, a simples inspeção visual deixa claro que as séries divergem, porém os retornos

³A série VALE5R foi construída adicionando-se aos retornos diários da VALE5 10% ao ano distribuídos diariamente.

continuam apresentando coeficiente de correlação de 0.99.



Figural1: Preços diários das ações Vale5, Vale3 e Vale5 reconstruída.

Fonte: Elaboração própria.

Nota: No primeiros gráfico mostramos as séries diárias de Vale5 e Vale3 (em logaritmo), no segundo gráfico temos as séries diárias da Vale5 com adicional de retorno de 10% distribuídos diariamente e Vale3.

Correlação reflete co-movimentos nos retornos, os quais são suscetíveis a grande instabilidade ao longo do tempo. Trata-se de uma medida intrinsecamente de curto prazo, e estratégias baseadas em correlação comumente exigem frequentes rebalanceamentos. Por outro lado, cointegração é uma medida de co-movimentos de longo-prazo nos preços, os quais podem ocorrer mesmo em períodos quando a correlação estática parece baixa.

Os pioneiros na utilização da técnica de cointegração a séries financeiras são Lucas (1997) e Alexander (1999). Segundo Alexander (1999), a técnica de cointegração para modelagem de séries de tempo possui muitas aplicações nos mercados financeiros: quando *spreads* são revertidos à média, preços são cointegrados. Portanto, um modelo multivariado produzirá importantes informações do preço de equilíbrio dos ativos financeiros e da causalidade dos retornos dentro do sistema. Arbitragem entre preços à vista e futuros, modelagem da estrutura da curva de juros, negociações através da construção de *index tracking* e de *spreads*, construção de estratégias *long-short* e *pair tradings*, são algumas das aplicações de cointegração a dados financeiros.

O uso da metodologia de cointegração na construção de estratégias com ações é encon-

trado em Alexander and Dimitriu (2002), Dunis and Ho (2005), bem como em Lin (2006), que utilizam a metodologia de cointegração para desenvolver um procedimento que embute uma condição de lucro mínimo dentro de uma estratégia de *pair tradings*: uma forma comparativa-valor de arbitragem estatística designada para explorar divergências aleatórias temporárias entre o preço de equilíbrio de duas ações.

Existem também várias aplicações da metodologia de cointegração para modelagem da estrutura a termo da taxa de juros (ver Bhar, 1994; Luc Bauwens and Vandeuren, 1997; Jumah and Kunst, 2002). Em todos os casos é empregada a metodologia de cointegração para examinar relações intertemporais entre as taxas de juros para diferentes maturidades, usando dados da estrutura a termo para diferentes países.

O objetivo deste artigo é verificar a aplicabilidade da metodologia de cointegração para derivar carteiras ótimas para o mercado de ações brasileiro e empregá-las na construção de estratégias *long-short*, extrapolando a clássica estratégia *index tracking*. São utilizados os preços diários do Índice Bovespa e suas principais ações constituintes, para o período de janeiro de 2000 a dezembro de 2008, perfazendo um total de 2,230 observações diárias. São derivadas carteiras que buscam replicar índices de referência construídos a partir do índice Ibovespa, e o próprio Ibovespa.

Na maioria dos casos as carteiras ótimas são obtidas com aproximadamente dez ações e os resultados apresentados reforçam a utilização da cointegração nas estratégias explicitadas. Os resultados mostram a capacidade das carteiras cointegradas *index tracking* de replicar um benchmark em termos de retorno e volatilidade. Da mesma forma, as estratégias *long-short*, baseadas em cointegração, apresentaram resultados estáveis sob condições normais de mercado e em circunstâncias adversas, apresentando retornos superiores ao benchmark considerado e volatilidade em níveis baixos⁴.

Além desta introdução, este artigo é constituído de mais 3 seções. Na seção 2 são apresentadas as estratégias *index tracking* e *long short* neutra ao mercado⁵. Na seção 3 são apresentados os dados empregados, e metodologicamente explicadas como são construídas as carteiras usadas na implementação das estratégias *index tracking* e *long-short*. Na seção 4 são avaliados empiricamente, para dados do mercado brasileiro, as carteiras *index tracking* e a estratégia *long-short* com diferentes periodicidades de rebalanceamento e composições das carteiras. Por último, é apresentada uma conclusão dos resultados empíricos e indicação de passos que podem potencializar os resultados das estratégias.

⁴É comum fundos definidos como long-short no mercado brasileiro apresentarem volatilidades anualizadas inferiores às obtidas com a estratégia long-short aqui analisada, vale destacar que a maioria desses fundos, apesar de serem definidos como long-short, empregam uma diversidade de estratégias, sendo que apenas uma pequena parcela de seu patrimônio alocadas em estratégias estritamente long-shorts.

⁵Cabe destacar que uma estratégia long-short não necessariamente é neutra ao mercado. Já a estratégia long-short aqui desenvolvida, por construção, busca ser neutra ao mercado.

2.2 Estratégias Neutras ao Mercado Long-Short e Carteiras Index Tracking

A estratégia de investimento - *index tracking*, consiste em replicar um benchmark em termos de retorno e volatilidade. Nessa estratégia deve-se atentar principalmente para a análise do *tracking error*⁶. No processo de otimização de carteiras por meio da análise média-variância geralmente exige frequentes rebalanceamentos para que carteira se mantenha alinhada com o benchmark. Uma das razões para que isso ocorra é que a variância da carteira é mensurada através de uma matriz de covariância, que é difícil de ser estimada e pode mudar consideravelmente dia após dia, alterando as recomendações da fronteira eficiente. Portanto, o critério média-variância não assegura que o *tracking error* seja estacionário. Assim, a carteira poderia desviar arbitrariamente do benchmark. Quando as alocações são determinadas através da cointegração da carteira com um índice de referência, a mesma seguirá o índice no longo prazo, podendo até desviar no curto prazo, mas no longo prazo andarão juntos.

A habilidade da carteira para replicar o índice de referência é caracterizada por seu *tracking error*. O *tracking error* pode ser visto como uma medida da discrepância ou margem de erro que se pode esperar no processo de *tracking*. Numa situação ideal o *tracking error* deve ser um processo ruído branco - média zero e variância constante (invariante com o tempo t). Isto asseguraria que a carteira *index tracking* seria consistente, ou não apresentaria grandes desvios em relação ao benchmark. Outra propriedade desejável do *tracking error* é a baixa correlação com o retorno de mercado.

Diferentemente do processo de otimização de carteiras por média-variância, a otimização baseada em cointegração permite o uso total das informações contidas nos preços das ações, e a composição da carteira é baseada no comportamento de longo prazo das ações. Assim, a construção de carteiras *index tracking* com base em cointegração, ao invés da simples correlação, se deve às seguintes características dos sistemas cointegrados: a) o *tracking error* irá por construção convergir para a média; b) estabilidade dos pesos das ações na carteira e, conseqüentemente, reduzida quantidade de transações para reequilibrar a carteira; c) melhor uso das informações, em particular as informações contidas nos preços das ações.

Existe uma vasta literatura tratando dos diferentes métodos de cointegração⁷, começando com o clássico artigo de Engle and Granger (1987), posteriormente refinado em Engle and Yoo (1987). O teste de Engle-Granger consiste em estimar uma regressão de cointegração entre as séries integradas e testar estacionariedade dos resíduos. Existem vários testes de estacionariedade (Choi, 1992; Cochrane, 1991; Dickey and Fuller, 1979; Schmidt and Phillips,

⁶O *tracking error* é definido como uma medida de quão aproximadamente um portfólio replica um benchmark. O *tracking error* mede o desvio-padrão da diferença entre os retornos do portfólio e os do benchmark,

$$TE = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N-1}}, \text{ onde } X_i \text{ é a diferença entre os retornos do portfólio e do benchmark para o período } i.$$

⁷Para detalhes ver Hamilton (1994); Enders (2002); Hendry (2000).

1992; Phillips and Perron, 1988; Wang and Yau, 1994), sendo que o mais popular é o teste ADF (*Augmented Dickey Fuller*). Segundo Alexander and Dimitriu (2002), quando existem mais que duas séries o método de Engle-Granger pode sofrer de viés e o pesquisador precisa saber claramente qual série deve ser usada como variável dependente na regressão. Quando se trata de sistemas multivariados, a metodologia de cointegração de Johansen (ver Hendry, 2001; Johansen and Juselius, 1990; Johansen, 1991) é considerada superior. Os testes de Johansen são baseados nos autovalores de uma matriz estocástica e na verdade reduzem para um problema de correlação canônica similar ao dos problemas de componentes principais. Além de não ter viés, sua função de poder tem propriedades melhores: os testes de Johansen buscam a combinação mais estacionária, enquanto o teste de Engle-Granger busca a combinação com variância mínima.

Apesar da metodologia de Johansen ser comumente considerada superior a de Engle-Granger em um sistema multivariado, Alexander (1999) argumenta que para muitas aplicações financeiras de cointegração existem boas razões para este último método ser o preferido. Primeiro, é muito direta a sua forma de implementação. Segundo, em aplicações de administração do risco, geralmente é de maior importância o critério de Engle-Granger de variância mínima ao invés do critério de Johansen de máxima estacionariedade. Terceiro, existe frequentemente uma escolha natural da variável dependente em regressões de cointegração (por exemplo, em estratégias *index tracking*). E, finalmente, o viés de Engle-Granger em pequenas amostras não será necessariamente um problema, pois as amostras são geralmente muito grandes nas análises financeiras e o vetor de cointegração é super consistente.

A construção de uma carteira *index tracking* consiste em duas partes igualmente importantes. A primeira é a seleção das ações que irão compor a carteira, sendo que a qualidade do *index tracking* é altamente dependente desta seleção de ações. A segunda etapa visa determinar os pesos de cada ação na carteira. Esses pesos são determinados pelos coeficientes da equação de cointegração entre o logaritmo natural do índice de referência e o preço das ações (em \ln) como variáveis exógenas, onde as estimativas são feitas com base no período amostral escolhido. A equação da carteira *index tracking* é definida a seguir:

$$\ln ibov_t = \beta_0 + \sum_{k=0}^n \beta_k \ln P_{k,t} + \varepsilon_t \quad (2.1)$$

onde $ibov_t$ é o índice bovespa em t , $P_{k,t}$ são os preços das ações no tempo t e $\varepsilon \sim i.i.d.N(0, \sigma_\varepsilon^2)$. Tomar o \ln das séries produz séries homogêneas e uma vez que as variáveis são cointegradas em nível, os \ln s das variáveis também serão. Além disso, usar o \ln dos preços tem a vantagem que o *tracking error* ε_t assume a forma de retorno, e os coeficientes são os pesos de cada ativo na carteira. Esses coeficientes são normalizados para somar 1 (um), representando a participação percentual de cada ação selecionada na carteira. Os retornos da carteira são calculados como a soma, ponderada pelos β_k normalizados, dos retornos diários das ações que a compõem.

Grande parte das estratégias que não envolvem alocação de recursos busca explorar ineficiências do mercado, comprando ativos subvalorizados e vendendo ativos sobrevalorizados. Entretanto nem sempre é estabelecida uma relação entre os grupos de ativos na abertura da estratégia *long-short*. Espera-se que os ativos subvalorizados se valorizem mais, ou desvalorizem menos que os ativos sobrevalorizados, conseqüentemente diminuindo o diferencial de preços. Porém isto não implica em neutralidade ao mercado, a não ser que seja estabelecida alguma relação entre os grupos de ativos individualmente. Estratégias *long-short* baseadas em cointegração envolvem apenas ativos com comprovada interdependência.

O objetivo de estratégias neutras ao mercado é obter um beta neutro, no contexto do modelo CAPM, e gerar retorno através da performance superior das posições *long* sob as posições *short*, independente da direção que o mercado se mova, seja em tempos de baixa ou alta volatilidade.

Para entender como carteiras neutras ao mercado são construídas, considere uma carteira composta apenas por posições compradas nos ativos. Espera-se que o β dos ativos seja positivo. Portanto, retornos positivos do mercado resultam em retornos positivos dos ativos, conseqüentemente retornos positivos para a carteira, o que implicaria em um β positivo para a carteira. De forma similar, uma carteira com posições vendidas possivelmente apresentaria um beta negativo. Assim, construir uma carteira com β zero só é possível mantendo posições compradas (*long*) e vendidas (*short*) em diferentes ativos que compõem a carteira. Por esta razão, estas carteiras são chamadas de carteiras *long-short*. Outra característica das carteiras *long-short* é que os recursos obtidos com as posições vendidas são usados quase que integralmente para estabelecer as posições compradas, isto é, o valor líquido envolvido é próximo de zero. Não surpreendentemente, carteiras beta zero são também chamadas de carteiras *dollar neutral*.

Ao contrário de simples estratégias *long-short*, estratégias neutras ao mercado envolvem somente ações ou outros ativos com interdependência comprovada. Esta interdependência, que às vezes assume a forma de convergência, assegura que, ao longo do tempo, as ações ou outros ativos assumirão uma relação de preços. Segundo Alexander (1999), as vantagens de investir em estratégias *long-short* neutras ao mercado são perceptíveis pela independência da direção tomada pelo mercado e o uso mais eficiente das informações, quando comparado com estratégias de compra apenas. A independência da direção do mercado, ou baixa correlação com o retorno do mercado, é consequência direta da zeragem dos β 's entre as partes *long* e *short* da carteira.

Uma extensão natural do uso da cointegração na otimização de carteiras *index tracking* é empregá-la para o desenvolvimento de estratégias *long-short*. Nesse caso, o objetivo é construir carteiras que sejam capazes replicar um benchmark para a parte *long* e outro para replicar a parte *short* da estratégia, assumindo posição comprada no primeiro, *long*, e posição vendida no segundo, *short*. Os dois benchmarks são construídos somando e subtraindo do índice de referência um excesso de retorno anual de $\delta\%$, distribuídos uniformemente para os

retornos diários. A partir daí as séries de preços dos índices de referência "*plus*" e "*minus*" são reconstruídos, em relação aos quais as carteiras serão otimizadas. Geralmente, quanto maior for o diferencial de retorno usado na construção dos benchmarks *plus* e *minus*, mais difícil se torna a obtenção de carteiras cointegradas. É bastante razoável esperar que carteiras compostas por muitas ações cointegrem com o índice de referência, o mesmo pode não ser verdade para benchmarks construídos artificialmente. As novas equações de cointegração assumem o seguinte formato:

$$\ln ibov_t^+ = \beta_0 + \sum_{k=0}^n \beta_k \ln P_{k,t}^+ + \varepsilon_t^+ \quad (2.2)$$

onde $ibov_t^+$ é o índice Ibovespa "*plus*" construído para ser o benchmark para a carteira *long*, e $P_{k,t}^+$ são os preços das ações que constituem a carteira "*plus*" no tempo t . Se as séries $ibov_t^+$ e $P_{k,t}^+$ são cointegradas a série dos resíduos ε_t^+ é estacionária.

Para a carteira "*minus*",

$$\ln ibov_t^- = \beta_0 + \sum_{k=0}^n \beta_k \ln P_{k,t}^- + \varepsilon_t^- \quad (2.3)$$

onde $ibov_t^-$ é o índice Ibovespa "*minus*" construído para ser o benchmark para a carteira *short*, e $P_{k,t}^-$ são os preços das ações que compõem a carteira "*minus*" no tempo t . Se as séries $ibov_t^-$ e $P_{k,t}^-$ são cointegradas, a série dos resíduos ε_t^- é estacionária.

Geralmente, as estratégias *long-short* neutras a mercado consistem em comprar a carteira "*long*" e vender a carteira "*short*". O retorno global diário da estratégia é calculado como a soma dos retornos diários da carteira "*long*" e da carteira "*short*" (multiplicado por -1 no caso da carteira "*short*"), onde o retorno diário de cada carteira é a soma ponderada dos retornos diários das ações constituintes.

2.3 Dados

Os dados utilizados neste trabalho compreendem as séries de preços diários de fechamento do índice Ibovespa e das 50 principais ações constituintes do índice para o quadrimestre referente ao período de teste considerado. Por exemplo, na estimação das carteiras para janeiro de 2001, foi considerada a composição do índice referente ao último quadrimestre de 2000 e assim sucessivamente. Referem-se a preços diários de fechamento, no período de 02/01/2000 a 31/12/2008, totalizando 2,230 observações. A tabela 2.1 traz a relação das ações que compunham o índice Ibovespa no último quadrimestre de 2008 em ordem decrescente de participação no índice. O critério usado para selecionar as ações foi sua participação percentual no índice, bem como se foram negociadas em todo o período amostral. Como consta também na tabela 2.1, foram realizados testes de raiz unitária para os *lns* das séries dos preços das ações; todas as séries empregadas são $I(1)$. Vale ressaltar que as séries de preços de ações de algumas

empresas que faziam parte da amostra deixaram de existir ao longo do tempo, também ocorreram fusões, bem como novas ações passaram a fazer parte do índice ao longo do período de teste⁸.

Ao trabalhar apenas com as ações com maior participação no índice Bovespa, e que foram negociadas em todos os dias da amostra, evita-se a inclusão de ativos pouco líquidos, que estão sujeitos a apresentar maiores distorções entre o preço de fechamento e o preço ao qual o ativo poderia efetivamente ser negociado⁹. Também, elimina problemas com *missing values*, e a inclusão de ações de empresas que tiveram seu capital aberto recentemente, não apresentando séries históricas de preços suficientes para o tratamento econométrico desejado.

Outra razão para se trabalhar apenas com séries de preços de ações que apresentam elevada liquidez está na redução dos custos operacionais. Estratégias *long-short* envolvem a compra de uma carteira de ações e a venda a descoberto de outra carteira. Ao vender as ações a descoberto o agente precisa alugar essas ações no mercado pelo período em que mantém a operação em aberto. O custo de aluguel das ações é determinado, entre outras variáveis, por sua liquidez. Por outro lado, a venda de um ativo com baixa liquidez pode implicar em dificuldade para alugar o ativo no mercado, implicando no encerramento da operação antes do momento desejado, ou em elevados custos de aluguel do ativo. No apêndice A.1 apresentamos as quantidades disponíveis e os custos de aluguel de algumas ações negociadas no mercado. Conforme pode ser visto nas tabelas A.2 e A.3, as taxas de aluguel das ações variam entre 0.02% e 8.00% a.a, valores que prevaleceram na maioria do período analisado, sendo raros os casos em que o custo de aluguel de algum ativo supera a taxa de 4%a.a.

Nota-se na tabela 2.1, que traz as ações com maior participação no índice Bovespa, que todas as ações, excluindo-se BVMF3 e ALLL11, que não foram negociadas em todo o período analisado, todas as séries apresentaram uma raiz unitária em nível e estacionariedade em primeira diferença. Levando-se em conta que todas as séries de preços das ações do índice Ibovespa são $I(1)$, a metodologia de cointegração pode ser empregada na otimização das carteiras.

Um dos problemas com que deparamos ao fazer back-testes empregando algum índice de mercado é que a composição dos índices varia ao longo do tempo. Alexander and Dimitriu (2002) ao trabalhar com dados do índice Dow Jones adota a estratégia de reconstruir retroativa o índice a partir da composição mais recente da amostra. No caso do índice Dow Jones essa estratégia é bastante factível, dado que a carteira do índice é definida como a média aritmética das 30 ações que o compõem. No trabalho de Alexander and Dimitriu (2002) é utilizada uma base de dados que vai de janeiro de 1990 a dezembro de 2001, sendo que ao comparar a composição do Dow Jones no período inicial da amostra com o a composição do último período apenas 4 ações foram substituídas. No caso do índice Ibovespa, no último

⁸Neste caso, por sugestão de um parecerista anônimo, à medida que a série completava 252 dias de negociação a ação é incluída na amostra, desde que satisfeitas as demais condições.

⁹Ao trabalhar apenas com ações com maior liquidez minimiza problemas de distorções entre os preços de compra e preços de venda (*Bid and Ask spread*), tornando os resultados mais reais.

Tabela1: Ibovespa e Composição da Carteira Teórica Jan-Abr de 2009

Ativo	Peso %	QTDE Teórica	ADF-Nível	ADF-1 ^a Dif	Tipo
IBOV	-	-	0.40	0.00	-
PETR4	16.72	274.81	0.39	0.00	PN
VALE5	11.88	186.79	0.49	0.00	PNA
BVMF3	4.36	275.11	-	0.00	ON
BBDC4	3.70	61.50	0.17	0.00	PN
ITAU4	3.47	50.10	0.16	0.00	PN
CSNA3	3.45	44.72	0.34	0.00	ON
VALE3	3.27	44.34	0.48	0.00	ON
GGBR4	3.10	77.29	0.59	0.00	PN
PETR3	3.06	41.81	0.44	0.00	ON
USIM5	3.05	43.14	0.51	0.00	PNA
UBBR11	2.42	61.33	0.18	0.00	UNT
BBAS3	2.40	61.49	0.20	0.00	ON
ITSA4	2.24	105.46	0.09	0.00	PN
CMIG4	1.64	19.41	0.10	0.00	PN
CYRE3	1.26	51.29	0.58	0.00	ON
CESP6	1.22	30.23	0.77	0.00	PNB
LAME4	1.20	71.85	0.61	0.00	PN
BRAP4	1.20	23.47	0.46	0.00	PN
SDIA4	1.14	114.09	0.82	0.00	PN
AMBV4	1.12	4.16	0.21	0.00	PN
TNLP4	1.09	12.74	0.13	0.00	PN
ELET3	0.95	13.80	0.03	0.00	ON
ELET6	0.95	14.77	0.02	0.00	PNB
NETC4	0.94	26.72	0.58	0.00	PN
GOAU4	0.91	17.08	0.41	0.00	PN
BTOW3	0.90	14.26	0.82	0.00	ON
PRGA3	0.88	11.2	0.40	0.00	ON
LREN3	0.85	20.56	0.22	0.00	ON
ARCZ6	0.77	117.2	0.98	0.00	PNB
TCSL4	0.75	96.43	0.83	0.00	PN
GOLL4	0.73	27.93	0.98	0.00	PN
VIVO4	0.72	9.75	0.20	0.00	PN
TAMM4	0.71	13.98	0.93	0.00	PN
EMBR3	0.65	27.94	0.96	0.00	ON
CPLE6	0.65	10.17	0.05	0.00	PNB
NATU3	0.63	12.60	0.16	0.00	ON
USIM3	0.62	9.01	0.53	0.00	ON
CCRO3	0.61	9.7	0.26	0.00	ON
CSAN3	0.60	20.14	0.96	0.00	ON
RSID3	0.56	56.56	0.92	0.00	ON
PCAR4	0.55	6.7	0.08	0.00	PN
DURA4	0.55	14.39	0.70	0.00	PN
BRKM5	0.54	36.87	0.28	0.00	PNA
CPFE3	0.52	6.5	0.27	0.00	ON
CRUZ3	0.48	4.14	0.14	0.00	ON

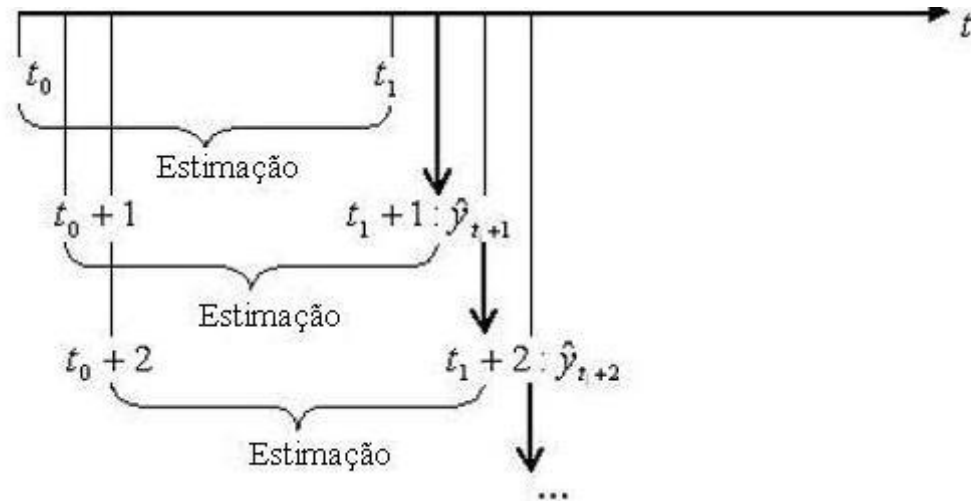


Figura 2: Janela móvel, esquema pelo qual a amostra é atualizada a cada nova estimação.

Fonte: Elaboração própria.

quadrimestre de 2000 o índice era composto por 57 ações e no último quadrimestre de 2008 o índice era composto por 66 ações. Das 57 ações que faziam parte do índice em 2000, 27 não faziam parte da carteira de 2008, e das 66 ações que compunham a carteira de 2008, 36 não estavam na carteira de 2000¹⁰. Além disso, a composição da carteira teórica do índice Ibovespa não é determinada por uma simples média linear, sendo que a participação de cada ação na carteira pode variar consideravelmente ao longo do tempo. Assim, optamos por trabalhar com o índice Ibovespa efetivo, sem reconstruir retroativamente com base na carteira teórica mais recente.

2.4 Construção das Carteiras

Os critérios da análise de cointegração são estacionariedade máxima e variância mínima do *tracking error*. Assim, através da cointegração é possível determinar carteiras ótimas que são aderentes ao benchmark e que apresentam risco mínimo relativo ao *tracking error*. Para a construção de uma carteira *index tracking*, primeiramente, é preciso definir uma carteira inicial cointegrada dentro da amostra - neste caso estimar as equações (2.2) e (2.3), de modo que, ε_t seja $I(0)$. Nesse sentido, foi utilizado o período de 02/01/2000 a 30/12/2000 (252 observações) - 01 ano de negociações diárias, para se definir a carteira inicial cointegrada (P0). Progressivamente, esta carteira foi expandida até dez/2008. A figura 2.2 ilustra o processo como a amostra é atualizada para estimação de cada nova carteira.

Isto é, para as duas estratégias, são definidas carteiras iniciais dentro da amostra com dados de janeiro de 2000 a dezembro de 2000, que são progressivamente expandidas men-

¹⁰Muitas das empresas tiveram os nomes alterados ou passaram por algum processo de fusão, passando a ser negociada sob outra denominação.

salmente até dezembro de 2008. A primeira carteira, construída com dados de janeiro de 2000 a dezembro de 2000 é testada fora da amostra a partir de janeiro de 2001, então a segunda carteira é estimada sobre o período de fevereiro de 2000 a janeiro de 2001 e aplicada fora da amostra para fevereiro de 2001. O procedimento é atualizado sucessivamente até dezembro de 2008. Portanto, 96 carteiras fora da amostra (P1-P96) são obtidas.

Os coeficientes da regressão de cointegração obtidos na carteira inicial (P0), assegurado que passa nos testes de cointegração, são subsequenteiramente re-estimados mensalmente para se definir a composição (os pesos) de cada carteira. A partir dos coeficientes estimados para cada carteira, *long* e *short*, constrói-se uma carteira com posições líquidas, uma vez que um ativo que faz parte da carteira *long* pode também estar presente na carteira *short*. Como interessa apenas o resultado líquido da estratégia e para evitar uma elevação nos custos operacionais, assumindo posições compradas e vendidas em um mesmo ativo, as operações de fato realizadas seguem a participação líquida dos ativos na carteira. No caso da estratégia *index tracking*, normaliza-se os coeficientes da regressão estimada, de tal forma que somem 1.

Questão crítica na construção de carteiras *index tracking* é o procedimento de seleção das ações. Para fins de determinação da carteira ótima são realizados testes de cointegração de Johansen e Engle-Granger, com todas as combinações entre os ativos da amostra para carteiras compostas entre 8 e 12 ativos. Apesar da composição da carteira ser determinada pela equação de cointegração de Engle-Granger, testes de cointegração de Johansen também são realizados e a carteira que apresenta os melhores critérios de cointegração são empregadas na construção das estratégias. Somente os pesos relativos da composição das ações na carteira são subsequenteiramente modificados¹¹. Outro ponto a ser destacado é que o uso de ativos mais líquidos minimiza possíveis erros na apuração dos resultados que surgem em decorrência dos preços de fechamento do mercado não refletirem necessariamente os preços aos quais a compra ou venda do ativo poderia ser realizada.

Como já destacado, a equação de cointegração de Engle-Granger permite que a composição da carteira seja determinada através dos coeficientes da regressão, normalizando-os para somar 1. Não existe nenhuma restrição específica na carteira: são admitidas tanto posições compradas como vendidas, tanto na carteira *long* quanto na carteira *short*. O coeficiente negativo do ativo na carteira significa que deve ser vendido no ato da ponderação da carteira, para posteriormente ser comprado a um preço relativo mais baixo. Ou seja, espera-se que o preço da ação vendida tenha uma queda relativa no tempo de maturidade da carteira, implicando em retornos positivos no ativo em questão.

¹¹Para a maioria dos casos a carteira ótima foi obtida com aproximadamente 10 ações, com esse número variando entre 8 e 12. Ao reestimar a carteira no momento do rebalanceamento a composição é alterada conforme os parâmetros da equação de cointegração de Engle-Granger estimada.

2.5 Avaliação Empírica das Estratégias com Dados do Mercado Brasileiro

Conforme destacado anteriormente, em modelos padrões de otimização de carteiras *index tracking*, os resíduos (*tracking error*) podem não ser estacionários, e a carteira construída para seguir um benchmark pode se distanciar do mesmo. A vantagem do uso da metodologia de cointegração para otimização de carteiras é a possibilidade de detectar tendências estocásticas comuns nas séries de preços dos ativos, ampliando sua estabilidade em relação ao índice de referência. O fato dos resíduos (*tracking error*) serem estacionários, por construção, assegura que a carteira se manterá próxima ao benchmark no longo prazo, independente dos movimentos individuais de curto prazo no preço das ações. As composições das carteiras *index tracking* e da estratégia *long-short* são apresentadas no apêndice A.1.

O retorno da estratégia *index tracking* é calculado como a soma do retorno das ações que compõem a carteira ponderada pelo seu peso na carteira, considerando que a composição percentual da mesma foi mantida constante até a data do próximo rebalanceamento. Porém, sabe-se que para manter a composição percentual exata da carteira ao longo de todo o período, seriam necessários ajustes diários, de tal forma a corrigir eventuais distorções na composição decorrentes de mudanças nos preços relativos dos ativos. Essas negociações diárias implicam em custos para a estratégia, e por mais que fossem feitos esses ajustes diários, dificilmente a composição percentual da carteira seria mantida exatamente a mesma ao longo do período.

Com o objetivo de eliminar os custos relativos aos ajustes diários necessários para manter a composição percentual da carteira constante e tornar a análise mais realista, foi analisada uma segunda alternativa, onde a carteira é construída no início de cada período com uma quantidade fixa de cada ação e mantida até o próximo rebalanceamento. A quantidade é determinada pela participação percentual da ação na carteira, multiplicada pelo volume de recursos a ser alocado na estratégia, divididos pelo preço da ação. Assim, o retorno da estratégia no período é obtido calculando-se o retorno dessa carteira com quantidades fixas de cada ação durante seu período de vigência. A figura 2.3 traz os retornos acumulados para a estratégia *index tracking* nas suas diferentes periodicidades de rebalanceamento e do Ibovespa.

Ao comparar os retornos acumulados das carteiras *index tracking* com os retornos do índice Ibovespa, pode-se observar que tanto as carteiras com composição percentual mantidas fixas quanto as carteiras com composição quantidades fixas apresentam comportamento similar ao índice de referência, em todas as periodicidades de rebalanceamento. Nota-se, porém, que a estratégia com composição quantidades fixas exibe maior aderência ao índice. Já a observação da figura 2.4, que traz os retornos mensais das carteiras *index tracking* e Ibovespa corrobora com a percepção de que a estratégia *index tracking* com carteiras cointegradas apresenta elevado grau de aderência ao benchmark, sendo que apenas em alguns meses apresentou descolamentos em relação ao índice de referência. Mesmo nos meses de maior oscilação as carteiras não exibem forte descolamento do índice de referência. Na tabela A.1, do apêndice

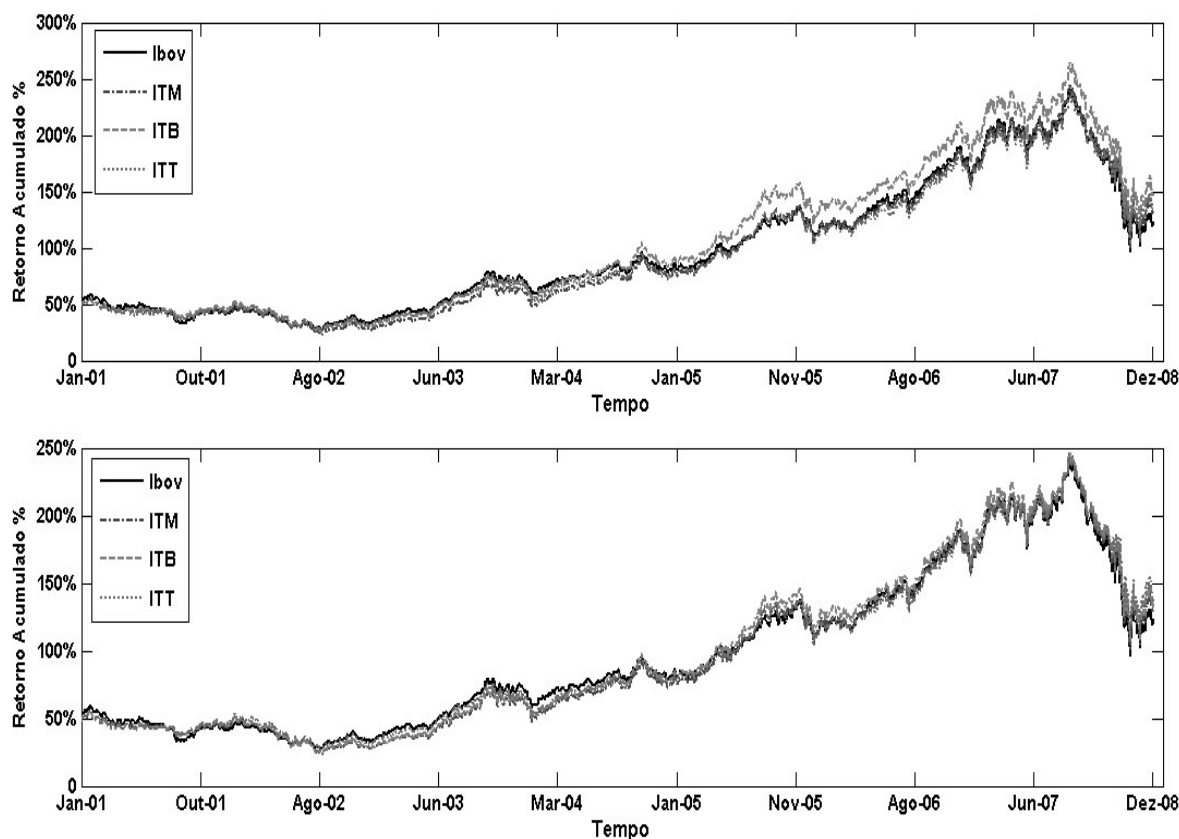


Figura 3: Resultados Acumulados da Estratégia Index tracking e Ibovespa (Jan-01 a Dez-08).

Fonte: Elaboração própria.

Nota: No primeiro painel estão as carteiras com pesos fixos e no segundo quantidades fixas.

A.1, apresentamos os retornos acumulados ano a ano da estratégia *index tracking* e Ibovespa. Também são apresentadas algumas estatísticas relativas aos retornos acumulados anuais.

A tabela 2.2 resume as estatísticas dos resultados acumulados da estratégia *index tracking* para o período entre janeiro de 2001 e dezembro de 2008, considerando todas as periodicidades de rebalanceamento analisadas, carteiras com composição percentual mantidas fixas e quantidades fixas. Além dos retornos acumulados, já descontados dos custos de transação, são apresentados volatilidade anualizada, coeficiente de correlação com Ibovespa e índice de Sharpe. Também são apresentadas estatísticas relativas ao *tracking error*, como curtose e assimetria. Por último, apresentamos autocorrelação de primeira ordem (ACF1) e teste de raiz unitária do *tracking error*. Gráficos mostrando a evolução dos coeficientes de correlação entre as carteiras *index tracking* e Ibovespa são apresentados no apêndice A.1.

Parte relevante da análise dos resultados da estratégia diz respeito aos custos de transação. Principalmente se levarmos em conta que uma das vantagens que se espera do uso de cointegração é a estabilidade dos parâmetros estimados, que determinam a composição da carteira.

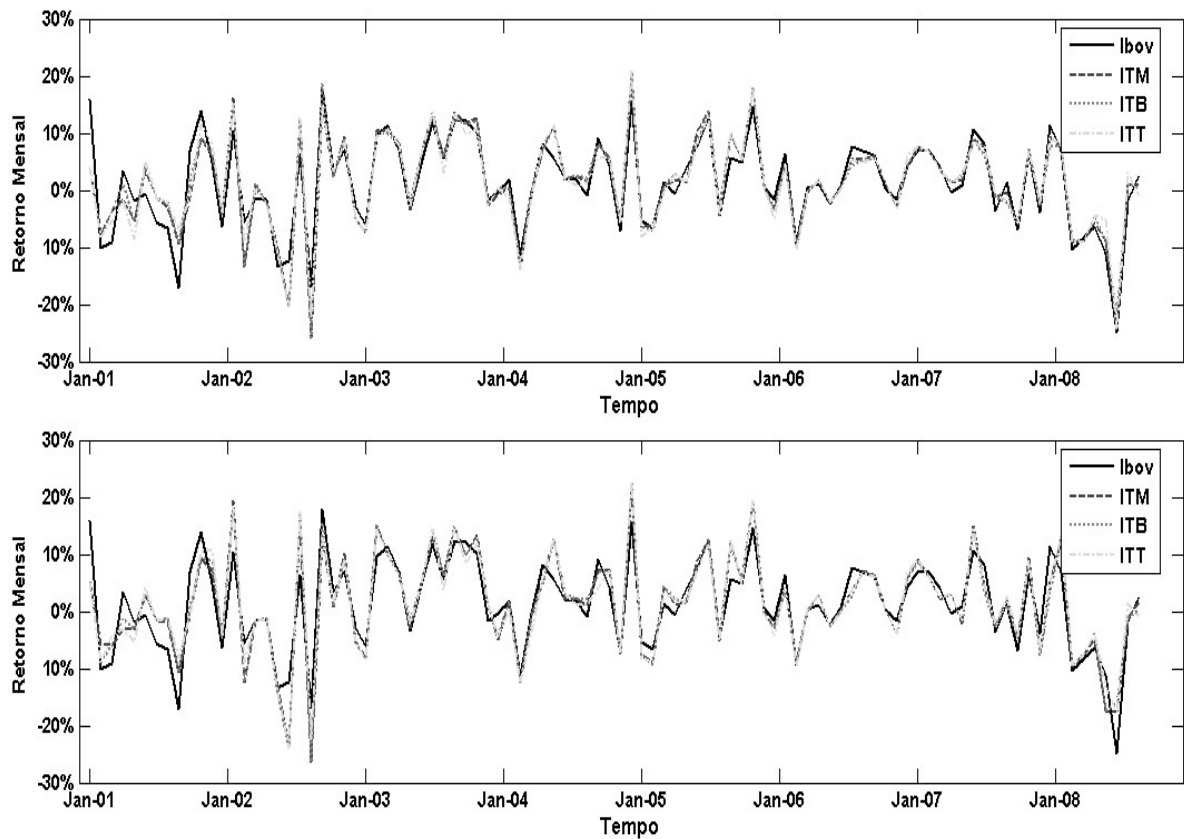


Figura 4: Retornos Mensais das Carteiras Index Tracking e Ibovespa (Jan-01 a Dez-008).

Fonte: Elaboração própria.

Nota: No primeiros gráfico carteiras com pesos fixos, no segundo quantidades fixas, considerando rebalanceamentos mensal, bimestral e trimestral nos dois casos.

Para a estratégia *index tracking*, os custos de transação são de aproximadamente 0.05%, sendo compostos por custos de operação e emolumentos. Assim, incluindo os custos de entrar na estratégia, a compra da carteira, e os custos de encerrar a estratégia, venda ou rebalanceamento da carteira, perfazem um custo total de 0.10%. Uma das principais diferenças entre as diversas estratégias de rebalanceamento é exatamente o custo de transação, que será menor quanto menor for a frequência de rebalanceamento.

Em termos de volatilidade, todas as carteiras *tracking* exibem padrão muito parecido com o Ibovespa, conforme se pode observar pela figura 2.5, que traz as volatilidades condicionais anualizadas¹² da estratégia *index tracking* e Ibovespa. As carteiras cointegradas compostas em média por 10 ações replicam de maneira satisfatória a volatilidade do benchmark.

Durante a maior parte do período analisado a volatilidade condicional das carteiras *index*

¹²As volatilidades condicionais foram obtidas através de suavização exponencial, EWMA, com lambda de 0.94, padrão nas séries de retornos de ativos financeiros. Esse método pode não ser o melhor para o propósito de ajuste e previsão de volatilidade, mas é o mais comumente empregado para comparação de volatilidades de diferentes ativos.

Tabela 2: Estatísticas das Carteiras *Index Tracking* e Ibovespa (Jan-01 a Dez-08).

	Ibovespa	ITWFM	ITWFB	ITWFT	ITQFM	ITQFB	ITQFT
Ret Médio Anual	18.77	19.38	20.48	18.01	19.50	19.85	18.92
Ret Acumulado	246.08	251.78	285.79	258.73	246.48	252.97	262.24
Vol Anual	32.34	34.70	34.80	34.52	34.710	35.04	34.60
Corr Ibov	1.000	0.947	0.944	0.946	0.946	0.942	0.944
Ind Sharpe	0.025	0.099	0.145	0.109	0.097	0.110	0.121
Assimetria	-	-0.05	-0.05	0.099	-0.100	-0.06	0.123
Curtose	-	3.168	2.935	3.113	3.315	3.262	3.300
ACF1	-	0.008	0.036	0.034	0.041	0.034	0.032
ADF	-	-21.29	-21.67	-21.83	-21.92	-21.70	-21.82

Fonte: Elaboração própria.

Nota: Resultados já descontados dos custos de transação. Retorno acumulado e volatilidade anualizada em %. WF = pesos fixos, QF = quantidades fixas, M = Mensal, B = Bimestral e T = Trimestral.

tracking esteve inferior 40% (ver figura 2.5), que traz as volatilidades anualizadas das carteiras com composição percentual fixa e com quantidade de ações mantida fixa, onde fica claro que a estratégia *index tracking* com carteiras cointegradas apresentaram padrão de volatilidade similar ao Ibovespa. A partir de agosto de 2008, quando a crise nos mercados financeiros atingiu seu estágio mais agudo, os níveis de volatilidade se elevaram consideravelmente, passando de 100% por alguns períodos. Mesmo nesse momento de maior volatilidade, a estratégia se comportou de maneira bastante aderente ao índice referência, apresentando padrão de volatilidade muito próximo ao benchmark mesmo nos períodos de maior instabilidade dos mercados.

No que diz respeito à volatilidade não parece haver muita diferença quando se compara as carteiras com diferentes periodicidades de rebalanceamento. Também, levar em conta os custos de transação não altera em nada seu padrão de comportamento, dado que os custos são relativamente estáveis e muito pequenos quando comparados com os retornos diários da estratégia.

Para completar a caracterização da estratégia *index tracking* foram estimados os betas das carteiras através do filtro de Kalman¹³, onde foi utilizado o algoritmo do suavizador de Kalman para obter estimativas suavizadas dos betas, apresentadas na figura 2.6. Os betas suavizados permitem verificar eventuais descolamentos estruturais das carteiras em relação ao índice de referência e quão aderentes as carteiras são ao benchmark. Também são apresentados os coeficientes de correlação obtidos por suavização exponencial da estratégia *index tracking* em relação ao Ibovespa, figura 2.7. Para o período como um todo, os coeficientes de correlação EWMA ($\lambda = 0.94$) ficaram próximos de 0.941, para todas as carteiras, con-

¹³Para detalhes sobre estimação por filtro de Kalman ver Durbin and Koopman (2001); Anderson and Moore (1979a); Simon (2006); Hamilton (1994).

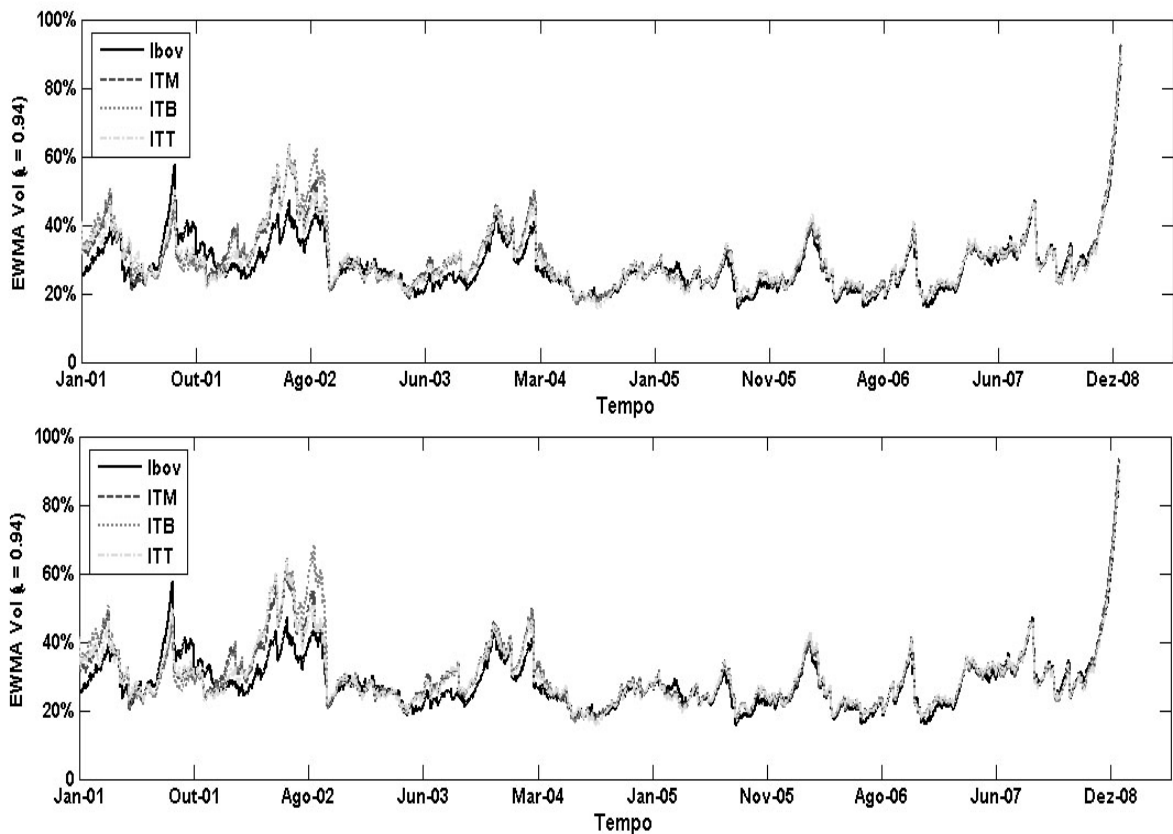


Figura 5: Volatilidade das Carteiras Index Tracking (EWMA, com $\lambda = 0.94$).

Fonte: Elaboração própria.

Nota: No primeiros gráfico carteiras com pesos e no segundo quantidades fixas, considerando rebalanceamentos mensal, bimestrale e trimestral nos dois casos.

forme tabela 2.3. Já os β 's estimados para todo o período ficaram muito próximos a um, corroborando com a análise dos resultados anteriores.

Mais uma vez os resultados corroboram com o uso da metodologia de cointegração para otimização de carteiras no âmbito da estratégia *index tracking*. Os β 's das carteiras se comportaram de maneira relativamente estável na maior parte do período analisado. Já os coeficientes de correlação, apesar de apresentarem alguma oscilação, principalmente no período inicial, quase sempre estiveram em patamares superiores a 0.90.

Nota-se que para o período analisado, os resultados fora da amostra das carteiras *index tracking* cointegradas apresentaram elevada correlação com o Ibovespa próxima a um, para todas as periodicidades de rebalanceamento. Ao analisar o índice de Sharpe percebe-se que independente da periodicidade de rebalanceamento, todas as carteiras apresentam níveis muito próximos ao benchmark, o que já era esperado, tendo em vista a semelhança nos padrões de retorno e volatilidade. O índice de Sharpe foi calculado como a média do excesso de retorno

Tabela 3: Coeficiente de Correlação EWMA ($\lambda = 0.94$) e β 's das Carteiras *Index Tracking*

	Pesos Fixos			Quantidades Fixas		
	ITM	ITB	ITT	ITM	ITB	ITT
EWMA Correl	0.941	0.940	0.940	0.940	0.940	0.939
β 's	0.998	0.992	0.993	0.996	0.993	0.993

Fonte: Elaboração própria.

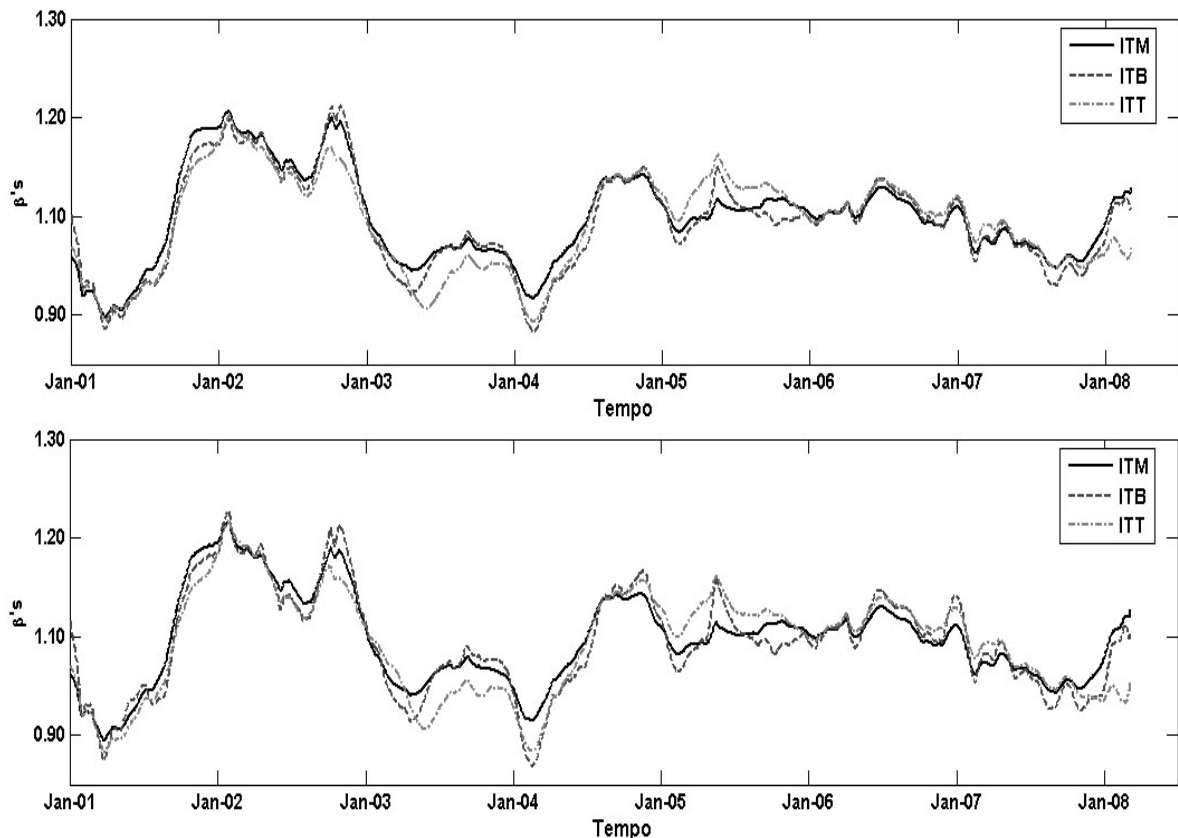


Figura 6: *Betas suavizados das Carteiras Index Tracking.*

Fonte: Elaboração própria.

Nota: No primeiros gráfico carteiras pesos fixos e no segundo quantidades fixas, considerando rebalanceamentos mensal, bimestral e trimestral nos dois casos.

anual da estratégia de investimento sobre a taxa livre de risco¹⁴, dividido pelo seu respectivo desvio padrão anualizado. As tabelas trazem também o índice razão de informação, que nada mais é que a média dos retornos anuais da estratégia dividido pelo seu desvio padrão.

¹⁴A taxa livre de risco considerada no cálculo do índice de Sharpe foi o CDI diário.

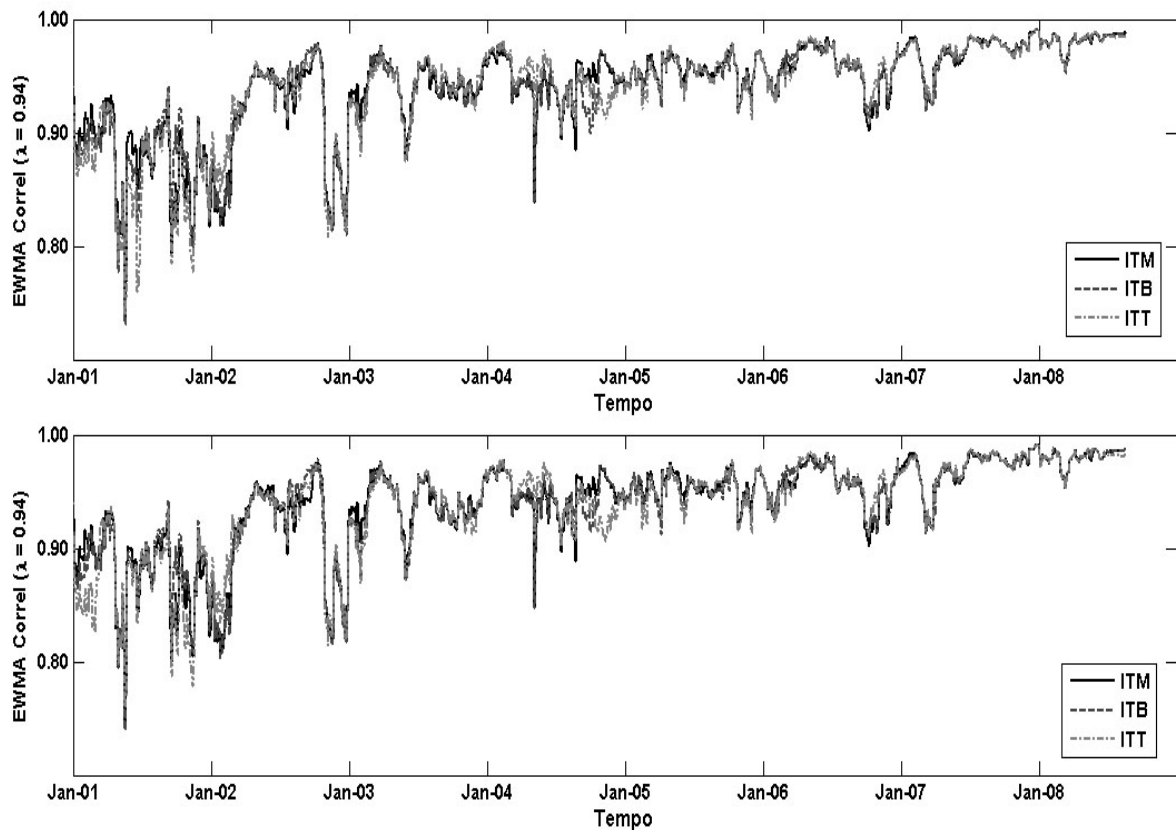


Figura 7: EWMA correlações para as carteiras index tracking ($\lambda = 0.94$).

Fonte: Elaboração própria.

Nota: No primeiros gráfico carteiras com pesos fixos e no segundo quantidades fixas, considerando rebalanceamentos mensal, bimestrale e trimestral nos dois casos.

Estratégia Long-Short

A estratégia *long-short* surge como uma extrapolação natural da estratégia *index tracking*. Conforme destacado anteriormente, a estratégia consiste de uma carteira comprada (*long*) financiada por uma carteira vendida (*short*). A composição líquida da carteira é obtida a partir da estimação das equações 2.2 e 2.3, onde o Ibovespa *plus*, benchmark para a carteira *long* estimada, é construído adicionando-se $\delta\%$ ¹⁵ ao retorno anual do Ibovespa distribuído diariamente. De maneira similar é construído o benchmark *minus*, subtraindo-se do Ibovespa um retorno anual de $\delta\%$ distribuídos diariamente. Todas as equações estimadas passaram nos testes de cointegração tanto de Jonhansen quanto Engle-Granger, porém, a composição das carteiras é determinada pela equação de cointegração de Engle-Granger. Depois de estimados as carteiras *long* e *short* e normalizados seus coeficientes de forma que somem um, obtém-se

¹⁵A determinação do parâmetro δ , que define o percentual de retorno a ser adicionado na construção dos benchmarks das carteiras *long* e *short* foi determinado via calibragem para o período dentro da amostra. Os melhores resultados foram obtidos com $\delta = 10\%$.

a carteira líquida somando os coeficientes da carteira *long* com os coeficientes da carteira *short* com o sinal invertido (todos os coeficientes da carteira *short* são multiplicados por -1, uma vez que essa é a carteira de ações que deve ser vendida para financiar a compra das ações que compõem a carteira *long*).

Assim como no caso da estratégia *index tracking*, as carteiras cointegradas com os benchmarks *plus* e *minus* foram obtidas com 10 ações, em média. Os resultados acumulados da estratégia são apresentados na figura 2.8, contemplando as alternativas de rebalanceamento mensal, bimestral e trimestral. Também foi considerada a possibilidade da composição da carteira ser mantida com a mesma quantidade de ações durante toda a vigência da mesma. Essa alternativa visa dar mais realidade à análise e redução dos custos de transação inerentes aos ajustes diários necessários para manter a composição percentual da carteira constante. Os resultados acumulados já estão descontados dos custos de transação, que neste caso foram considerados de 0.20% entre corretagem e emolumentos (considerando a compra e a venda de cada uma das carteiras) e os aluguéis das ações que compõem a carteira *short* foram considerados de 2.00%a.a.

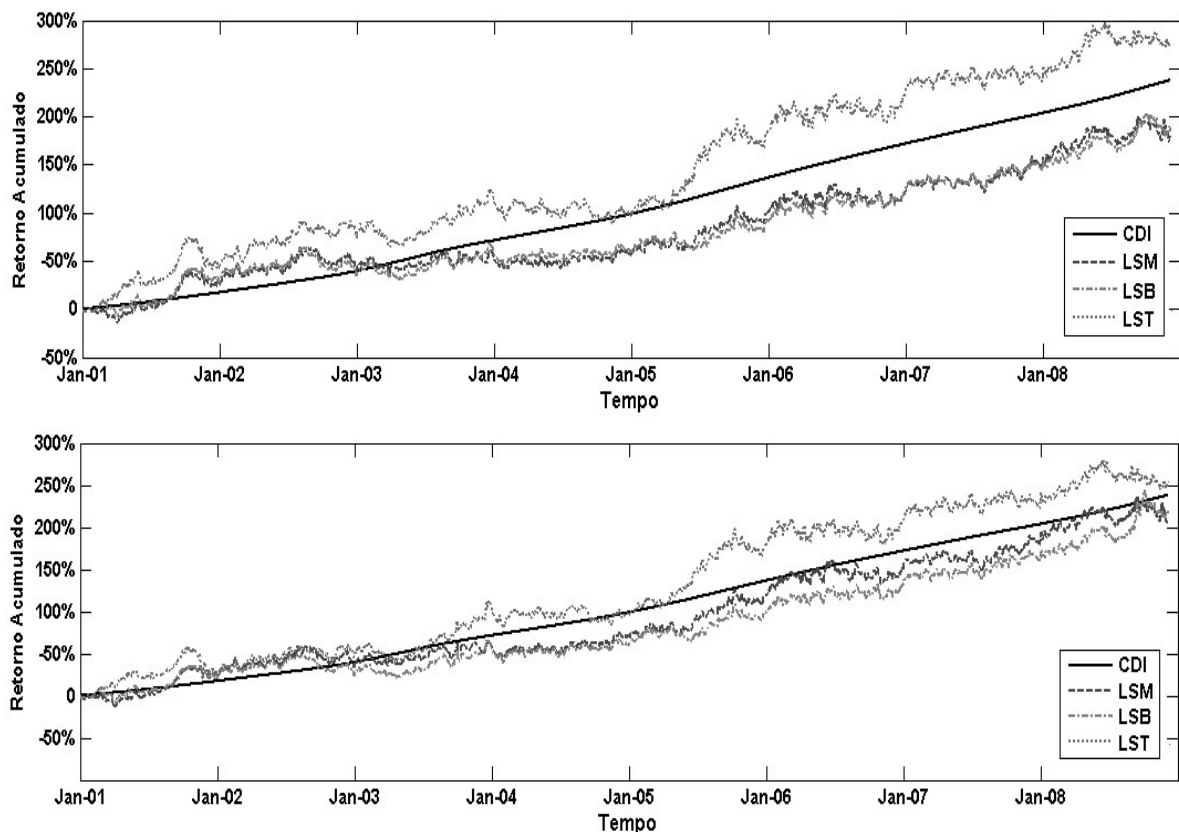


Figura 8: Resultados Acumulados da Estratégia Long Short e CDI (Jan-01 a Dez-08).

Fonte: Elaboração própria.

Nota: No primeiro painel estão as carteiras que consideram pesos fixos e no segundo quantidades fixas.

Tabela 4: Estatísticas da Estratégia *Long-Short* e CDI (Jan-01 a Dez-08).

	CDI	LSWFM	LSWFB	LSWFT	LSQFM	LSQFB	LSQFT
Retorno Médio Anual	20.83	19.19	17.98	22.75	20.71	19.5	21.53
Ret Acumulado	338.17	281.51	284.71	380.94	313.06	314.96	356.17
Vol Anual	-	17.89	17.87	17.45	18.29	18.21	17.34
Corr Ibov	-	-0.060	-0.055	-0.062	-0.055	-0.045	-0.028
Ind Sharpe	-	0.807	0.837	1.022	0.878	0.899	0.959
Assimetria	-	0.022	-0.100	-0.080	0.027	-0.090	-0.15
Curtose	-	3.542	3.633	3.776	3.690	3.792	4.625

Fonte: Elaboração própria.

Nota: Resultados já descontados dos custos de transação. Retorno acumulado e volatilidade anualizada em %. WF = pesos fixos, QF = quantidades fixas, M = Mensal, B = Bimestral e T = Trimestral.

A figura 2.8 traz também o CDI acumulado para o período, que apesar de não ser um benchmark mais adequado para a estratégia é usado como referência para a maioria dos fundos *long-short* do mercado brasileiro. Como a estratégia consiste em comprar uma carteira, a parte *long*, financiada pela venda de outra carteira, a parte *short*, os custos operacionais envolvidos são bastante relevantes para a estratégia. Além dos custos operacionais envolverem os custos das duas carteiras, as ações vendidas precisam ser alugadas no mercado. Os custos de aluguel de alguns ativos são apresentados nas tabelas A.2 e A.3 do apêndice A.1. Pode-se observar, que nos dois casos considerados o critério de rebalanceamento trimestral acumulou retorno levemente superior, parte desse diferencial é explicado pelos custos de transação, já que no caso do rebalanceamento trimestral a estratégia é operada apenas quatro vezes ao ano, enquanto que com rebalanceamento mensal são doze operações.

Vale destacar que a estratégia *long-short* é praticamente uma estratégia que envolve caixa zero¹⁶. Assim, poderia se considerar, para efeitos de apuração da rentabilidade, que o volume de recursos disponível é aplicado à taxa do CDI, o que potencializa os resultados da estratégia. Essa possibilidade não é considerada aqui. Na tabela 2.4 apresentamos estatísticas relativas à estratégia *long-short*, todos os itens foram calculados com retornos já descontados os custos de transação. Em todas as alternativas consideradas a estratégia *long-short* apresentou índice de Sharpe positivo, também apresentou volatilidade anualizada inferior a 20%.

Na tabela 2.5 mostramos as rentabilidades anuais e índices de Sharpe da estratégia *long-short* para os anos de 2001 a 2008. Em praticamente todos os anos os resultados foram positivos, mesmo nos anos de 2001, marcado pelo evento de 11 de setembro que abalou os mercados financeiros e 2008, ano da crise do subprime, que afetou fortemente os mercados financeiros mundiais, inclusive o brasileiro. Exceto no ano de 2004, a estratégia apresentou

¹⁶Só não é caixa zero devido às margens de garantia exigidas pela BM&F, que equivalem a 25% do valor alocado na carteira *short*.

índice de Sharpe próximo a um em todos os anos analisados. Para as carteiras com pesos mantidos fixos, a estratégia com rebalanceamento mensal superou as demais na maioria dos anos. Já no caso das carteiras com quantidades fixas não há uma alternativa de rebalanceamento que sobressaia e nesse caso, quando levamos em conta os menores custos de transação envolvidos, o rebalanceamento trimestral acaba se destacando, como pode ser visto nos resultados para o período como um todo.

Tabela 5: Rentabilidade Anual e Índice de Sharpe da Estratégia *Long-Short* (Jan-01 a Dez-08).

	Rentabilidade Anual						Índices de Sharpe					
	Pesos Fixos			Quantidades Fixas			Pesos Fixos			Quantidades Fixas		
	LSM	LSB	LST	LSM	LSB	LST	LSM	LSB	LST	LSM	LSB	LST
2001	32.83	35.21	48.96	27.80	23.95	27.70	1.43	1.50	1.84	1.24	1.09	1.13
2002	22.54	14.01	31.77	20.60	12.58	27.49	1.12	0.78	1.54	1.05	0.76	1.45
2003	12.25	18.13	23.82	22.36	29.56	42.06	0.73	1.02	1.30	1.16	1.44	1.95
2004	2.34	-0.68	-9.15	7.62	1.43	-5.26	0.22	0.04	-0.49	0.53	0.17	-0.24
2005	28.22	19.27	42.05	28.75	22.15	40.60	1.46	1.10	2.00	1.50	1.25	2.06
2006	15.11	19.62	16.86	20.76	21.41	14.13	0.81	1.00	0.90	1.06	1.11	0.77
2007	24.14	20.79	13.79	23.32	22.85	14.84	1.72	1.50	1.08	1.66	1.54	1.10
2008	16.11	17.52	13.89	14.43	22.08	10.71	1.11	1.22	1.07	0.97	1.33	0.78

Fonte: Elaboração própria.

Parte muito relevante na análise da estratégia *long-short* é a observação da correlação com o índice de mercado de referência. Por construção, a estratégia busca neutralidade ao mercado, uma vez que tanto para a carteira comprada quanto para a carteira vendida espera-se apresentar isoladamente elevada correlação com o índice de referência. Tanto para os dados anuais, quanto para período completo, a estratégia apresentou baixa correlação com o Ibovespa, assumindo valores entre -0.25 e 0.25 na maior parte do período analisado, conforme pode-se observar na figura 2.9, que traz o coeficiente de correlação EWMA ($\lambda = 0.94$) para todas as carteiras. Assim, a estratégia cumpriu satisfatoriamente o objetivo de neutralidade ao mercado.

Outro ponto a ser destacado é que mesmo nos períodos de maior volatilidade dos mercados, mês de setembro de 2001 e a partir do mês de agosto de 2008, os resultados da estratégia não se deterioraram, evidenciando uma das principais características que se busca em estratégias *long-short*, que é serem neutras aos movimentos do mercado.

Uma das razões para os níveis dos índices de Sharpe obtidos para a estratégia está no padrão de volatilidade apresentado por todos as carteiras. Para todas as alternativas analisadas, a volatilidade anualizada ficou abaixo de 20%. Já as volatilidades condicionais¹⁷, apresentadas na figura 2.10, exibem padrão de comportamento bastante estável e em níveis

¹⁷A volatilidade condicional foi estimada usando média móvel com suavização exponencial (EWMA), com $\lambda = 0.94$.

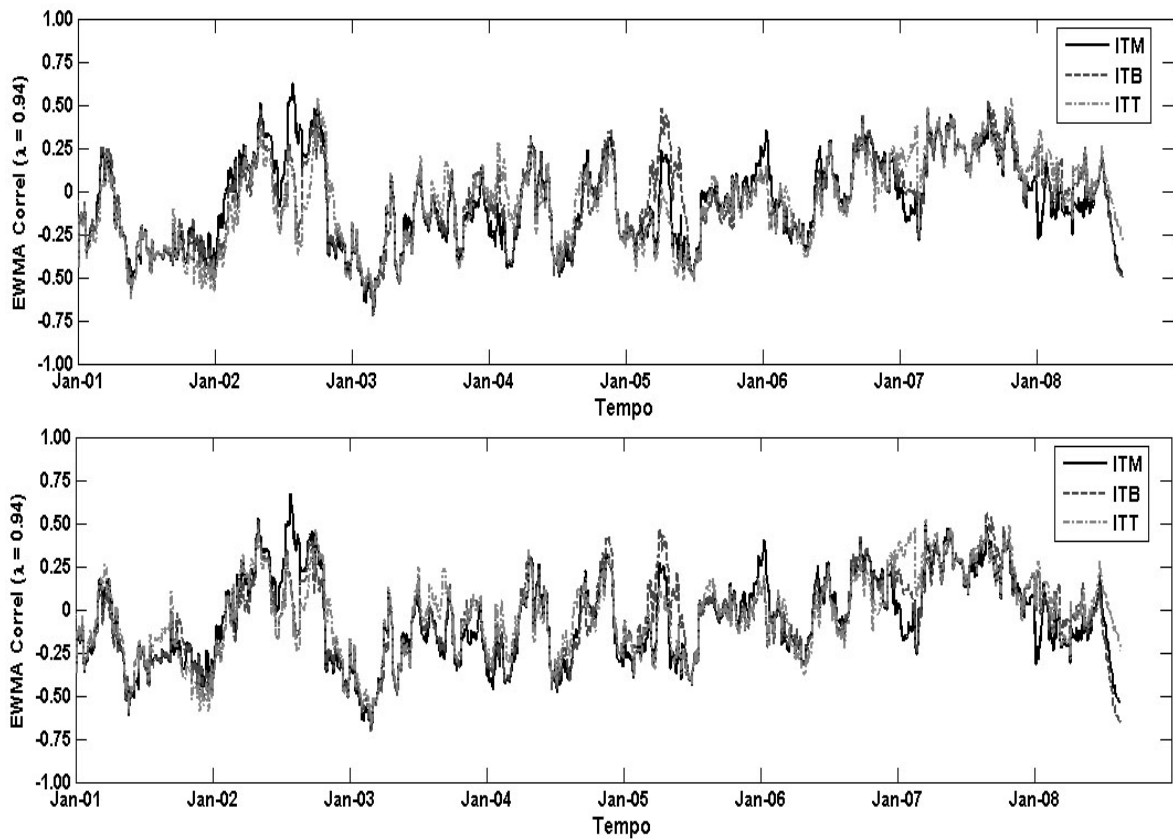


Figura 9: EWMA correlações para Estratégia Long-Short ($\lambda = 0.94$).

Fonte: Elaboração própria.

Nota: No primeiros gráfico carteiras com pesos fixos e no segundo quantidades fixas, considerando rebalanceamentos mensal, bimestrale e trimestral nos dois casos.

relativamente baixos, permanecendo próximo a 20% para maior parte do período analisado. Os períodos de maior volatilidade são verificados exatamente no último quadrimestre de 2001 e a partir do mês de agosto de 2008, quando inicia o período mais agudo da crise no mercado financeiro.

Mesmo apresentando níveis mais elevados de volatilidade para esse período, em nenhum momento a volatilidade anualizada superou 50%, enquanto o Ibovespa apresentou níveis de volatilidade superiores a 100% em termos anualizados.

As estratégias neutras ao mercado *long-short*, como envolvem a negociação de duas carteiras, são auto financiadas, uma vez que a carteira comprada é totalmente financiada pela carteira vendida. Assim, a estratégia envolve desembolso de recursos próximo a zero, sendo o desembolso no momento da abertura da operação constituído pelos custos de transação e margens de garantia. Dessa forma é muito comum, na análise desse tipo de estratégia, considerar que o volume de recursos alocado, geralmente com um nível de alavancagem de 2:1, é remunerado a alguma taxa livre de risco, que no caso brasileiro seria a taxa de juros do CDI.

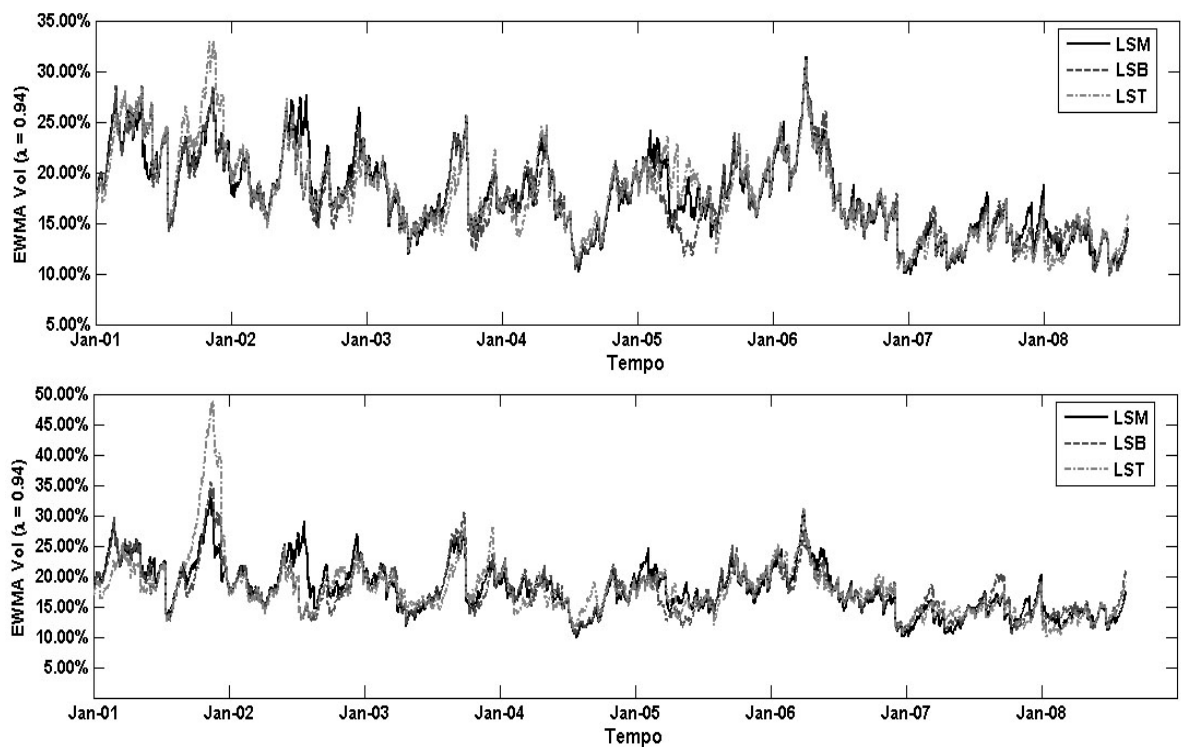


Figura 10: EWMA Volatilidade da Estratégia Long-Short ($\lambda = 0.94$).

Fonte: Elaboração própria.

Nota: No primeiro gráfico carteiras com pesos fixos e no segundo quantidades fixas, considerando rebalanceamentos mensal, bimestral e trimestral nos dois casos.

Os resultados apresentados neste artigo não contemplam essa possibilidade.

A observação dos resultados apresentados evidencia que, no caso da estratégia *long-short*, as três alternativas de rebalanceamento apresentaram resultados bem parecidos, sendo que quando se considera os custos de transação a estratégia com rebalanceamento trimestral supera as demais nos resultados acumulados. Uma vez que as alternativas analisadas apresentaram o mesmo padrão de volatilidade, o fator de decisão fica por conta dos custos operacionais. Neste caso, a estratégia *long-short* com rebalanceamento trimestral parece mais atraente do ponto de vista da gestão recursos, tanto no caso em que foram consideradas carteiras com percentuais fixas, como quantidades fixas.

2.6 Conclusões

A principal motivação deste artigo foi demonstrar os benefícios do uso da metodologia de cointegração, que se baseia em relações de longo prazo entre os preços dos ativos, para otimização de carteiras com dados do mercado brasileiro no contexto da estratégia neutra ao mercado *long-short*, que é uma extrapolação da clássica estratégia *index tracking*, também analisada. Estratégias tradicionais, baseadas em correlação, não asseguram estacionariedade do *tracking error* e exigem, portanto, frequentes rebalanceamentos para que a estratégia se mantenha em linha com o benchmark. As estratégias baseadas em cointegração são baseadas em critérios de reversão à média para o benchmark, uma vez que o *tracking error* é estacionário por construção. Conforme foi demonstrado, esse objetivo pode ser alcançado com carteiras compostas com um número pequeno de ações, dez em média, e com baixos custos de rebalanceamento, tendo em vista a estabilidade das carteiras.

A capacidade de replicar um benchmark em termos de retorno e volatilidade oferecida pela cointegração, com características de reversão à média do *tracking error*, estabilidade na composição das carteiras e melhor uso das informações contidas nos preços dos ativos, possibilita seu uso na implementação de uma variedade de estratégias, como *index tracking*, estratégias neutras ao mercado *long-short*, estratégias envolvendo pares de ações, *pair-trading*, que não foram analisadas neste artigo, entre outras.

Os resultados obtidos neste estudo são bastante animadores, tendo em vista a capacidade da estratégia *index tracking* baseada em cointegração, de replicar o benchmark em termos de retorno e volatilidade. Da mesma forma, a construção de estratégias *long-short* com carteiras cointegradas gerou resultados superiores ao benchmark e com baixo nível de volatilidade. Mesmo contendo algumas hipóteses simplificadoras, como o uso apenas das 50 ações mais líquidas do índice Bovespa, a não adoção de um critério de encerramento das posições diante de um dado nível de ganho ou perda, a metodologia de cointegração se mostrou eficaz, encorajando seu uso por parte de gestores de *hedge funds*, fundos de pensão e participantes do mercado em geral. Os resultados aqui encontrados, em termos de retorno, neutralidade e aderência ao índice de referência, se mostraram superiores aos obtidos em Alexander e

Dimitriu (1999) e Dunis e Ho (2002), que também empregam a metodologia de cointegração na otimização de carteiras. Uma das possíveis razões é que aqui o processo de otimização das carteiras envolveu simulações com todas as combinações possíveis dentro do universo de ações considerado, sendo a carteira ótima definida em função dos melhores critérios de cointegração apresentados. Vale destacar também que o parâmetro δ , que define o diferencial de retorno empregado na construção dos benchmarks para as carteiras empregadas na estratégia *long-short*, foi definido através de um processo de calibragem, em que foi escolhido o valor que apresentou os melhores resultados para os testes dentro da amostra.

A definição de critérios para encerramento das estratégias, quando estas atingem determinados níveis de ganho ou perda antes da data do rebalanceamento, podem potencializar consideravelmente os resultados, e são de fácil implementação. Outra possibilidade, para o caso das estratégias *long-short*, que não envolvem desembolso de recursos, é o aumento do grau de alavancagem. Nos resultados aqui apresentados foi considerado apenas um nível de 2:1. Um nível de alavancagem de 4:2, por exemplo, possibilitaria uma melhora considerável nos resultados em relação ao benchmark, e é bastante razoável, principalmente se for levado em conta o baixo nível de volatilidade apresentado pela estratégia no período analisado.

De maneira geral, a principal conclusão é que a otimização de carteiras através da metodologia de cointegração para implementação de estratégias *index tracking* e estratégias neutras ao mercado, *long-short*, gera resultados significantes em termos de retorno e volatilidade. Além disso, dado que o *tracking error* é estacionário, a otimização de carteiras por este método evita os frequentes rebalanceamentos como ocorre com os métodos tradicionais de otimização de carteiras. Outro ponto positivo é que as carteiras resultantes não são concentradas, evitando o risco inerente à concentração dos investimentos em poucos ativos. Assim, pode-se afirmar que a otimização de carteiras por meio de cointegração são capazes de gerar valor econômico para o investidor. Mesmo em períodos de maior instabilidade, como no último quadrimestre de 2001 e a partir do mês de agosto de 2008, a estratégia *index tracking* se manteve próxima ao benchmark. Já a estratégia *long-short*, que tem como principal objetivo gerar retorno sem exposição a riscos direcionais se mostrou neutra ao mercado, mesmo nesses períodos, com leve aumento no padrão de volatilidade, mas mantendo o mesmo padrão de retornos. Assim, a crescente indústria de *hedge funds*, fundos multimercados e gestores quantitativos, podem fazer uso dessa ferramenta, além da análise de correlação, no processo de otimização de carteiras.

3 Estimação e Previsão Eficientes da Curva Juros para Dados do Brasil

Modelar a estrutura a termo da taxa de juros é extremamente importante para macroeconomistas e participantes do mercado financeiro em geral. Neste artigo é empregada a formulação de Diebold-Li para ajustar e fazer previsões da estrutura a termo da taxa de juros brasileira. São empregados dados diários referentes às taxas dos contratos de DI Futuro negociados na BM&F que apresentaram maior liquidez para o período de Janeiro de 2006 a Fevereiro de 2009. Diferentemente da maior parte da literatura sobre curva de juros para dados brasileiros, em que o modelo de Diebold-Li é estimado pelo método de dois passos, neste trabalho o modelo é colocado no formato de estado espaço, e os parâmetros são estimados simultaneamente, de forma eficiente, pelo Filtro de Kalman. Os resultados obtidos tanto para o ajuste, mas principalmente no que diz respeito à previsão, mostram que a estimação do modelo através do Filtro de Kalman é a mais adequada, gerando melhores previsões para todas as maturidades quando é considerado horizontes de previsão de um mês, três meses e seis meses.

3.1 Introdução

Compreender a evolução da estrutura a termo da taxa de juros é do interesse de macroeconomistas, economistas financeiros e gestores de renda fixa, o que gerou grande avanço na modelagem teórica, bem como na estimação deste tipo de processo nas últimas décadas. Os principais modelos desenvolvidos nesse período podem ser divididos em três classes: os modelos de não arbitragem; os modelos de equilíbrio; e modelos estatísticos ou paramétricos. Os modelos de não arbitragem têm foco no perfeito ajuste da estrutura a termo em um dado ponto do tempo, assegurando que não existam possibilidades de arbitragem, o que é importante para o apreamento de derivativos. Exemplos de modelos desta classe são Hull and White (1990), Heath et al. (1992). Os modelos de equilíbrio têm foco na modelagem da taxa instantânea, utilizando tipicamente modelos afins, depois disso, as taxas de outros vencimentos podem ser derivadas sob várias suposições acerca do prêmio de risco. Modelos desta classe foram desenvolvidos por Vasicek (1977), Cox et al. (1985), e em Duffie and Kan (1996).

A classe de modelos estatísticos ou paramétricos é composta por modelos de componentes principais, modelos de fatores ou de variáveis latentes, bem como por modelos de interpolação. De acordo com Matzner-Lober and Villa (2004), grande parte da intuição a respeito da dinâmica da rentabilidade de títulos e bônus provém de modelos dessa classe, como em Robert and Scheinkman (1991) e em Pearson and Sun (1994). Dentre os modelos de fatores, o modelo de Nelson and Siegel (1987) e suas variantes, são os mais populares entre gestores de renda fixa e bancos centrais. A atratividade dos modelos de fatores do tipo Nelson Siegel

se deve à sua parcimonialidade e boa performance empírica. Modelos deste tipo conseguem capturar a maior parte da evolução da estrutura a termo da taxa de juros através do uso de apenas três fatores. Modelos com maior número de fatores foram utilizados por Svensson (1994), Almeida and Vicente (2008), Laurini and Hotta (2008) entre outros.

Modelos de interpolação foram desenvolvidos, por exemplo, por McCulloch (1971, 1975), que faz uma interpolação da função de desconto e não da taxa de juros ou dos preços dos ativos diretamente; e por Vasicek and Fong (1982), que ajustaram splines exponenciais para a curva de desconto, proporcionando ajustes mais suaves para os trechos mais longos da curva.

Diebold and Li (2006) argumentam que, apesar dos grandes desenvolvimentos na modelagem teórica da estrutura a termo dos juros, pouca atenção foi dada à previsão da estrutura a termo. Os modelos de não arbitragem são focados no ajuste para um dado ponto do tempo e têm muito pouco a dizer sobre dinâmica ou previsão fora da amostra. Já os modelos de equilíbrio possuem algumas implicações dinâmicas dado um certo prêmio de risco, o que possibilita algum tipo de conclusão a respeito de previsões fora da amostra. Entretanto, ainda de acordo com Diebold e Li, a maioria dos trabalhos dedicados a modelos de equilíbrio focaram na performance dentro da amostra. Exceções são Duffee (2002), que mostra que os modelos livre de arbitragem têm baixa performance em previsões fora da amostra; e Egorov et al. (2010) que mostram que modelos afins com volatilidade estocástica são capazes de prever a distribuição condicional conjunta da rentabilidade dos bônus.

Ter boas previsões da taxa de juros é crucial para cálculo do valor do mercado de uma carteira de ativos, na avaliação de derivativos de renda fixa, e na construção de estratégias de investimento, bem como para o desenvolvimento da política monetária.

Em outra linha de pesquisa, Ang and Piazzesi (2003), Hördahl et al. (2006), e Wu (2001), analisaram modelos com variáveis macroeconômicas e mostraram que estas variáveis contribuem para uma melhor previsão da dinâmica da curva de juros. Diebold et al. (2006) (a partir de agora DRA), usaram um modelo com fatores latentes da curva de juros e também incorporaram variáveis macroeconômicas. Diferentemente de modelos anteriores, que consideravam uma relação unidirecional das variáveis macro para a curva de juros, ou da curva de juros para as variáveis macro, DRA analisaram a possibilidade de relação bidirecional, e mostraram que a inclusão das variáveis macroeconômicas melhoram a previsibilidade do modelo principalmente para horizontes de seis meses e um ano à frente, para as maturidades intermediárias da curva de juros estudada.

Vicente and Tabak (2007) comparam o modelo afim gaussiano com o modelo Diebold e Li para dados brasileiros e chegam à conclusão de que o modelo Diebold e Li é ligeiramente superior no que diz respeito à previsão da curva de juros. Almeida et al. (2007) obtêm resultados de previsão superiores aos do modelo Diebold e Li usando uma versão dinâmica do modelo de Svensson. Vargas (2007) utiliza dados brasileiros de contratos de DI futuro para replicar os resultados de Diebold Li. Laurini and Hotta (2008), por sua vez, estimam um modelo de Svensson (1994) estendido, onde os parâmetros de decaimento variam no tempo e

volatilidade estocástica é adicionada à equação de medida do sistema de estado espaço.

Neste artigo, utiliza-se o modelo de três fatores da estrutura a termo proposto por Nelson and Siegel (1987), porém reinterpretando os fatores como nível, inclinação e curvatura da curva de juros assim como Diebold e Li, para fazer previsões fora da amostra. Para estimar os modelos e proceder ao exercício de previsão é empregada a abordagem de estado espaço introduzida nesse contexto por DRA, que permite ajustar simultaneamente a curva de juros em cada ponto do tempo e estimar a dinâmica dos fatores subjacentes por máxima verossimilhança. Este procedimento evita a seleção *a priori* do parâmetro de decaimento, e possibilita a extração de estimativas suavizadas dos fatores, que são posteriormente usadas no exercício de previsão. A base de dados do estudo consiste de séries diárias de taxas de juros à vista dos contratos de DI Futuro negociados na BM&F, evitando o uso de taxa de SWAP que, muitas vezes, não refletem taxas de negócios efetivamente executados.

Nosso artigo avança em relação à literatura em dois pontos: nós estimamos o modelo da curva de juros de Diebold e Li eficientemente em uma única etapa através do filtro de Kalman, evitando a seleção *a priori* do parâmetro de decaimento λ ; e nós empregamos uma base de dados composta apenas por dados de DI-futuro para maturidades com alta liquidez. Além desta introdução este artigo está organizado como segue. A seção 2 apresenta a estrutura do modelo de Diebold-Li para a curva de juros e sua versão no formato de espaços de estados. Na seção 3 são apresentados os dados empregados na estimação, e além da análise do ajuste do modelo é realizado um exercício de previsão com o objetivo de verificar a capacidade do modelo de gerar boas previsões fora da amostra. Na seção 4 são apresentadas algumas conclusões e propostas de trabalhos futuros.

3.2 Modelo de Fatores da Curva de Juros

Nesta seção será apresentado o modelo de fatores para a estrutura a termo da taxa de juros. A versão aqui apresentada segue o modelo de três fatores proposto por Diebold and Li (2006), que é na verdade uma reinterpretação da curva de juros de Nelson and Siegel (1987), onde os três fatores da curva de Nelson e Siegel são interpretados como fatores de nível, inclinação e curvatura.

3.3 Função de Desconto, Curva Forward e Curva de Juros

Antes de descrever a estrutura do modelo será apresentada a definição de curva de desconto, curva *forward* e curva de juros, bem como suas inter-relações. A estrutura a termo das taxas de juros é representada por um conjunto de pontos no espaço taxa de juros spot versus prazo. Cada ponto corresponde a uma taxa de juros $y_i(\tau)$ associada a uma maturidade τ , obtida com base em algum título negociado no mercado (Varga, 2008).

Em qualquer ponto do tempo t , existirá uma coleção de títulos sem cupom (*zero-coupon bonds*) que diferem uns dos outros apenas pela maturidade. Porém, em um dado momento do tempo pode não existir título disponível para todas as maturidades desejadas, já que no mercado financeiro não são negociados títulos para todas as maturidades possíveis.

Uma das construções mais fundamentais descrevendo a estrutura a termo da taxa de juros, geralmente a partir da qual as demais curvas são derivadas, é a função de desconto. Seja $P_t(\tau)$, o preço de um título zero-cupom no tempo t , e que paga \$1, na maturidade τ . Por suposição, todo título zero-cupom é livre de default e tem preços estritamente positivos. Assim, a função de desconto é definida por:

$$P_t(\tau) = e^{-\tau y_t(\tau)}. \quad (3.1)$$

A taxa de juros $y_t(\tau)$, à qual o título é descontado, é a taxa interna de retorno do título zero-cupom, no tempo t , e com maturidade τ , expressa por:

$$y_t(\tau) = \frac{-\ln(P_t(\tau))}{\tau}. \quad (3.2)$$

A taxa forward no tempo t , aplicada para o intervalo de tempo entre τ_1 e τ_2 , é definida por:

$$f_t(\tau_1, \tau_2) = \frac{1}{\tau_2 - \tau_1} \int_{\tau_1}^{\tau_2} y_t(x) dx. \quad (3.3)$$

O mesmo argumento se aplica para taxas forwards para k -períodos. A taxa forward pode ser interpretada como a taxa marginal de retorno para manter um título por um instante adicional. O limite da equação (3.3) quando τ_1 se aproxima de τ_2 , denotada por $f_t(\tau)$, é a taxa *forward* instantânea:

$$f_t(\tau) = \frac{-P_t(\tau)}{P_t(\tau)}. \quad (3.4)$$

A curva de taxa forward instantânea, $f_t(\tau)$, fornece a taxa de decaimento da função de desconto $P_t(\tau)$ em cada ponto τ . A curva de juros $y_t(\tau)$ é a média da taxa de decaimento para o intervalo de 0 a τ , expressa por:

$$y_t(\tau) = \frac{1}{\tau} \int_0^{\tau} f_t(x) dx. \quad (3.5)$$

A função $f_t(\tau)$, das taxas *forward*, descreve a taxa de retorno (instantânea) de um investimento mantido por um intervalo de tempo muito curto. A curva de taxas forwards instantâneas é uma construção teórica muito importante, apesar de seu valor para uma única maturidade ser de pouco interesse prático, devido ao elevado custo de transação para se fazer um contrato entre dois pontos no futuro se esses dois pontos forem muito próximos um do outro. Apenas a média de $f_t(\tau)$ para um intervalo de tempo futuro é de interesse prático.

Em qualquer ponto do tempo t , haverá um conjunto de títulos com diferentes maturidades, τ , e diferentes fluxos de pagamentos, que podem ser usados para estimar as curvas de juros, curvas de desconto e curvas forwards, que não são observáveis na prática. Existem algumas abordagens para construção de curvas de juros. McCulloch (1971, 1975) e Vasicek and Fong (1982) constroem curvas de juros a partir de curvas de desconto suaves estimadas, convertendo-as para taxas nas maturidades relevantes. A metodologia proposta por McCulloch (1971, 1975) faz uma interpolação da função de desconto com cubic splines. A vantagem dessa metodologia é que os parâmetros aparecem linearmente no modelo de estimação. Uma desvantagem dessa metodologia é gerar curvas erráticas para maturidades mais longas, isto é, a curva de desconto ajustada diverge para maturidades mais longas, ao invés de convergir para zero. A abordagem de Vasicek e Fong (1977) sugere que o emprego de splines exponenciais para ajustar a função de desconto, que eliminaria o problema da divergência para maturidades mais longas.

Modelos estatísticos não lineares também foram empregados para estimar a estrutura a termo da taxa de juros, ao invés da função de desconto, como em Nelson and Siegel (1987), Svensson (1994), entre outros. Esses modelos se mostraram bastante úteis para análise e precificação de títulos de renda fixa, com destaque para o trabalho de Nelson and Siegel (1987), que recebeu uma nova interpretação em Diebold and Li (2006), onde os fatores de curto, médio e longo prazos passaram a ser interpretados como fatores de inclinação, curvatura e nível.

Fama and Bliss (1987) propuseram uma metodologia para construção da estrutura a termo via taxas forwards estimadas para as maturidades observadas. O método consiste em seqüencialmente construir as taxas forwards necessárias para precificar sucessivamente títulos com maturidades mais longas, chamadas de taxas forwards não-suavizadas de Fama e Bliss. A taxa de juros a termo resultante deste procedimento é uma função com saltos (descontínua) em relação ao vencimento do título que está sendo utilizado.

3.4 O Modelo da Curva de Juros de Diebold e Li

O problema clássico da estrutura a termo requer a estimação de uma curva de juros suave a partir dos preços dos títulos observados. Em anos recentes o método tem sido computar as taxas *forwards* implícitas para precificar sucessivamente títulos com maturidades mais longas nas maturidades observadas, chamadas de taxas *forwards* não suavizadas. Então uma curva de taxas forwards suaves é obtida ajustando uma forma funcional paramétrica através das taxas não suavizadas. Uma das formas funcionais paramétricas mais amplamente usadas na estimação da curva de juros foi proposta por Nelson and Siegel (1987), que desenvolveram um modelo flexível o suficiente para representar curvas com diversos formatos. Neste modelo os parâmetros são associados à taxa de juros de longo prazo, médio prazo e curto prazo. Essencialmente essa forma descreve a curva de juros em termos de três fatores, os quais são

interpretados como nível, inclinação e curvatura, e outro fator que representa uma escala de tempo.

Nelson and Siegel (1987) propõem ajustar a curva *forward* em um dado momento através de uma classe matemática de funções de aproximação. A forma funcional proposta é uma função de Laguerre, à qual consiste do produto entre um polinômico e um termo de decaimento exponencial. A curva *forward* aproximada resultante é assumida como solução de uma equação diferencial de segunda ordem com raízes iguais para a taxa de juros à vista:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t}e^{-\lambda_t\tau} + \beta_{3,t}\lambda_t e^{-\lambda_t\tau}. \quad (3.6)$$

Os parâmetros $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ são determinados pelas condições iniciais e λ_t é uma constante associada com a equação. Recentemente Diebold and Li (2006) reinterpretaram o modelo exponencial proposto por Nelson and Siegel (1987), considerando uma forma paramétrica para a evolução da estrutura a termo ao longo do tempo, em que os coeficientes são tratados como nível, inclinação e curvatura. A curva de juros correspondente é:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} \right) + \beta_{3,t} \left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau} \right). \quad (3.7)$$

A curva de juros de Nelson e Siegel também corresponde à função de desconto, assumindo valor 1 na maturidade zero, e aproximando de zero quando a maturidade converge para o infinito.

A forma da curva de juros é determinada pelos três fatores e pelos factor loadings associados a eles. O parâmetro λ , mantido fixo em Diebold and Li (2006), governa a taxa de decaimento exponencial, pequenos (grandes) valores de λ estão associados a um decaimento suave (rápido), e ajustam melhor às maturidades longas (curtas). O peso do primeiro componente é sempre 1 (constante), e é interpretado como nível da curva de juros, que influencia igualmente as taxas de curto e longo prazo. O peso do segundo componente $\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} \right)$ começa em 1 e converge para zero monotonicamente e rapidamente, sendo interpretado como inclinação. Esse fator influencia muito as taxas de juros curto prazo. Já o peso do terceiro componente, $\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau} \right)$, é uma função côncava, assumindo valor zero para a maturidades zero, cresce, depois converge monotonicamente para zero nas maturidades mais longas. Assim esse fator está associado às taxas de juros de médio prazo, e tratado como curvatura da curva de juros.

Uma vez que o peso do primeiro componente é o único que é igual a 1 quando a maturidade se aproxima do infinito, $\beta_{1,t}$ é associado à taxa de juros de longo prazo. A inclinação da curva de juros geralmente é definida como $y_t(\infty) - y_t(0)$, nesse caso a inclinação converge para $\beta_{2,t}$. A curvatura é definida como $2y_t(\tau^*) - y_t(\infty) - y_t(0)$, onde τ^* representa uma maturidade de médio prazo. Note que a curvatura é aproximadamente $-\beta_{3,t}$ ¹.

Dado que em cada ponto do tempo, são observados títulos com várias maturidades, tem-se

¹Diebold and Li (2006) definem a inclinação como $(y_t(120) - y_t(3)) = -0.78\beta_{2,t} + 0.06\beta_{3,t}$, e a curvatura como $(2y_t(24) - y_t(120) - y_t(3)) = 0.000053\beta_{2,t} + 0.37\beta_{3,t}$.

um conjunto de taxas de juros com maturidades $(\tau_1, \tau_2, \dots, \tau_N)$, para todo t . Portanto Diebold and Li (2006) propõem fixar λ em um valor, *a priori*, e estimar a equação (3.7) por meio de mínimos quadrados ordinários para cada t , de onde são obtidas séries de tempo dos fatores $\beta_{j,t}$. Os $\beta_{j,t}$ são obtidos estimando a equação a seguir para cada t :

$$y_t(\tau_i) = \beta_{1,t} + \beta_{2,t} \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} \right) + \beta_{3,t} \left(\frac{1 - e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i} \right) + \varepsilon_{i,t}, \quad (3.8)$$

Onde os erros $(\varepsilon_{1,t}, \varepsilon_{2,t}, \dots, \varepsilon_{N,t})$ são, por suposição, independentes, com média zero e variância constante σ_t^2 para um dado tempo t .

Em geral existem várias especificações podem ser empregadas para ajustar os dados. Entretanto, grande parte da literatura existente emprega basicamente duas especificações para ajustar o modelo. Em um dos casos assume-se que as três variáveis de estado seguem um processo autorregressivo de primeira independentes, empregado, por exemplo, em Diebold and Li (2006). No outro caso, os três fatores não observados do modelo, colocado em formato estado espaço, são modelados por um vetor autorregressivo de primeira ordem, VAR (1), como em Diebold et al. (2006) e Koopman et al. (2010). Tanto em Diebold and Li (2006), quanto em Diebold, Diebold et al. (2006), os factor loadings dependem de um único parâmetro de decaimento, e para possibilitar a estimação dos fatores latentes variantes no tempo em um formato linear, os factor loadings são mantidos constantes ao longo do tempo para cada maturidade.

No caso em que a equação de transição, que governa a dinâmica do vetor de estados, é definida por um VAR(1):

$$\begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} = \begin{pmatrix} \mu_{\beta_1} \\ \mu_{\beta_2} \\ \mu_{\beta_3} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t-1} \\ \beta_{2,t-1} \\ \beta_{3,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix}. \quad (3.9)$$

No caso em que o modelo é ajustado por um processo autorregressivo em primeira ordem a matriz A , com elementos $a_{i,j}$ acima, é diagonal. A equação de medida, a qual associa as taxas de juros das N maturidades aos três componentes não observados, é dada por:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1 - e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1 - e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1 - e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1 - e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1 - e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1 - e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix} \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{N,t} \end{pmatrix}. \quad (3.10)$$

O sistema, composto pela equação de transição e equação de medida, pode ser escrito

usando notação matricial:

$$\beta_{t+1} = \mu + \Phi\beta_t + \eta_t, \quad (3.11)$$

$$y_t = \Lambda(\lambda)\beta_t + \varepsilon_t, \quad (3.12)$$

Onde $\Lambda(\lambda)$ é uma matriz $N \times 3$ de pesos dos fatores, sendo que seus elementos são variantes no tempo somente se o parâmetro de decaimento for variável.

A equação de medida (3.12) define o vetor de taxas de juros $N \times 1$ para as N maturidades, como a soma dos fatores latentes multiplicados pelos seus pesos, com um vetor de erros distribuídos normalmente e independente entre as maturidades. O vetor β_t de dimensão 3×1 representa os fatores e Φ é a matriz com os parâmetros do VAR dos estados.

Se o objetivo for apenas ajustar a estrutura a termo da curva de juros, a equação de medida é suficiente. Entretanto, para fazer previsões da estrutura a termo é necessário modelar também a dinâmica dos fatores. Seguindo Diebold et al. (2006) e Koopman et al. (2010), a dinâmica dos fatores é especificada como um processo autorregressivo de primeira ordem. Por fim, assume-se que os erros da equação de medida e equação de estado são ortogonais entre si e em relação ao vetor de estados iniciais, β_0 , e são normalmente distribuídos:

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim \mathcal{N} \left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \Sigma_\eta & \mathbf{0} \\ \mathbf{0} & \Sigma_\varepsilon \end{pmatrix} \right].$$

Além disso, assume-se que os erros da equação transição e erros de medidas são ortogonais ao estado inicial:

$$\begin{aligned} E[\beta_0 \eta'] &= \mathbf{0} \\ E[\beta_0 \varepsilon'] &= \mathbf{0}. \end{aligned}$$

A matriz de covariância dos erros de medida ε_t constituem uma matriz diagonal ($N \times N$), Σ_ε . A suposição de que a matriz Σ_ε é diagonal implica que os desvios das taxas de juros para diferentes maturidades são não-correlacionados. Esta suposição facilita a estimação do modelo, reduzindo o número de parâmetros, e é bastante comum na literatura. Por outro lado, a suposição de que a matriz Σ_η é irrestrita permite que choques nos três fatores sejam correlacionados.

3.5 Estimação e previsão usando o formato espaços de estados

Com o modelo no formato estado espaço, duas abordagens podem ser empregadas para estimar os fatores latentes e os parâmetros. A abordagem inicial proposta por Diebold and Li (2006) é baseada em dois estágios e, portanto, é ineficiente, não levando em conta a incerteza inerente às estimativas do primeiro estágio no estágio subsequente. No primeiro passo estima-se a equação de medida usando dados cross-section, em que são obtidos os estimadores para os parâmetros para cada momento do tempo. Assumindo que o parâmetro de decaimento é constante, a equação de medida torna-se linear e pode ser estimada por mínimos quadrados. No segundo passo a dinâmica temporal dos parâmetros é especificada e estimada como um processo AR(1) ou VAR(1).

Diebold et al. (2006) mostraram que é possível estimar este modelo por máxima verossimilhança em apenas um estágio utilizando o filtro de Kalman, obtendo-se então estimativas eficientes dos parâmetros e estimativas suavizadas dos fatores não observáveis. Esta é a abordagem adotada não só em Diebold et al. (2006), mas também em Koopman et al. (2010), entre outros. O procedimento usa o filtro de Kalman para a construção da função de verossimilhança, que é então maximizada para a obtenção de estimativas dos parâmetros. Nós consideramos o modelo de Nelson-Siegel 3.11 e 3.12 como um modelo em espaços de estados linear Gaussiano. O vetor de estados não observados β_t pode ser estimado condicional às observações passadas e correntes, y_1, \dots, y_t via filtro de Kalman. Define $\beta_{t|t-1}$ como o valor esperado de β_t condicional a y_1, \dots, y_{t-1} com matriz de erro quadrático médio (MSE) $P_{t|t}$. Para dados valores de $\beta_{t|t-1}$ e $P_{t|t-1}$, quando a informação y_{t-1} é disponível, o erro de previsão pode ser calculado $v_t = y_t - \Lambda(\lambda)\beta_{t|t-1}$. Assim, após observar y_t , podem ser obtidas inferências mais acuradas de β_t and $P_{t|t}$:

$$\beta_{t|t} = \mu + P_{t|t-1} + P_{t|t-1}\Lambda(\lambda)'F_t^{-1}v_t,$$

$$P_{t|t} = \beta_{t|t-1} - P_{t|t-1}\Lambda(\lambda)'F_t^{-1}\Lambda(\lambda)P_{t|t-1},$$

onde $F_t = \Lambda(\lambda)P_{t|t-1}\Lambda(\lambda)' + \Sigma_\epsilon$ é a matriz de covariância do erro de previsão. O estimador do vetor de estados para o próximo período $t + 1$, condicional a y_1, \dots, y_t , é dado pelo passo da previsão:

$$\beta_{t+1|t} = \mu + \Phi\beta_{t|t},$$

$$P_{t+1|t} = \Phi P_{t|t} \Phi' + \Sigma_\eta.$$

Para uma dada série de tempo y_t , os cálculos do filtro de Kalman são realizados recur-

sivamente para $t = 1, \dots, T$ com valores iniciais $\beta_{1|0}$ e $P_{1|0}$. Os parâmetros na matriz de coeficientes do VAR, matrizes de covariância Σ_η e Σ_ε juntos com μ e λ são tratados como coeficientes desconhecidos, os quais são colocados juntos no vetor ψ . A estimação de ψ é baseada em maximização numérica da função de verossimilhança que é construída via decomposição do erro de predição e dada por:

$$l(\psi) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^T \log |F_t| - \frac{1}{2} \sum_{t=1}^T v_t' \log F_t^{-1} v_t. \quad (3.13)$$

Como um resultado, $l(\psi)$ pode ser avaliada pelo filtro de Kalman para um dado valor de ψ (para detalhes sobre estimação através de filtro de Kalman ver Durbin and Koopman, 2001; Anderson and Moore, 1979a; Simon, 2006).

O estimador de máxima verossimilhança obtido dessa forma é preferível à estimação em duas etapas, uma vez que a estimação dos parâmetros na segunda etapa não considera a incerteza sobre os valores dos parâmetros estimados na primeira etapa, gerando estimadores não eficientes. A estimação conjunta das equações de medida e estado, por outro lado, não sofre desse problema e produz estimativas eficientes dos parâmetros. Outra vantagem da estimação por verossimilhança é a estimação conjunta do parâmetro de decaimento, que na estimação em dois passos precisa ser calibrado de acordo com alguma medida. Almeida et al. (2007) mostram que diferentes regras para a calibração do parâmetro de decaimento geram diferentes resultados para a previsão fora da amostra da estrutura a termo da taxa de juros, evidenciando que o método de estimação em dois estágios carece de robustez. Além disso, o suavizador de Kalman permite a obtenção de estimativas suavizadas das variáveis latentes, que consideram toda a amostra para inferir sobre a série de tempo dos fatores, que é então usada no exercício de previsão.

3.6 Dados e Análise dos Resultados

O contrato futuro de depósito interbancário (DI Futuro) com maturidade τ é um contrato futuro do qual o ativo básico é taxa de juros,² apurada em base diária, capitalizada entre o momento da negociação t , e a maturidade τ . O valor do contrato é dado pelo seu valor no vencimento, R\$100,000.00 descontado pela taxa de juros acumulada, negociada entre o vendedor e o comprador do contrato.

Ao comprar um contrato de DI Futuro ao preço ID^* no tempo t e mantê-lo até a maturidade τ , o ganho ou perda é dado por:

²A taxa DI é a taxa média diária dos depósitos interbancários (emprestar/tomar emprestado), calculada pela Câmara de Custódia e Liquidação (CETIP) para todos os dias úteis. A taxa DI, publicada em bases diárias, é expressa em termos compostos anualmente, com base em 252 dia úteis.

$$100.000 \left(\frac{\prod_{i=1}^{\zeta(t,\tau)} (1 + y_i)^{\frac{1}{252}}}{(1 + DI^*)^{\frac{\zeta(t,\tau)}{252}}} - 1 \right),$$

Onde y_i denota a taxa DI, $(i - 1)$ dias após o dia da negociação. A função $\zeta(t, \tau)$ representa o número de dias entre t e τ .

O contrato de DI é muito similar ao *zero-cupom Bond*, exceto pelo fato de pagar os ajustes de margem diariamente. A cada dia o fluxo de caixa é a diferença entre o preço de ajuste (liquidação) do dia corrente e o preço de ajuste do dia anterior, corrigido pela taxa DI do dia anterior.

Os contratos de DI Futuro são negociados na BM&F, que determina o número de vencimentos com contratos autorizados. Geralmente, para cada dia, existem em torno de 20 vencimentos com contratos autorizados, mas nem todos apresentam liquidez. Normalmente por volta de 10 vencimentos apresentam contratos com maior liquidez. Existem contratos com vencimentos mensais para os meses que iniciam cada trimestre, janeiro, abril, julho e outubro. Além disso, existem contratos com vencimentos para os quatro meses subseqüentes ao mês corrente. A data de vencimento é o primeiro dia útil do mês de vencimento do contrato.

3.7 Dados

Os dados usados nesse artigo consistem de observações diárias das taxas de dos contratos de DI Futuro, preços de fechamento. Na prática, não são observados diariamente contratos com todas as maturidades utilizadas. Assim, a partir das taxas observadas diariamente para as maturidades disponíveis, os dados foram convertidos para maturidades as fixas de 1, 2, 3, 4, 6, 9, 12, 15, 18, 24, 27, 29, 31 e 33 meses, por meio de interpolações pelo método exponencial splines. Os dados foram observados para o período de Janeiro de 2006 a Fevereiro de 2008, e representam os contratos de DI mais líquidos negociados para o período analisado.

São utilizados apenas os dados de ajuste dos contratos de DI Futuro, não incluindo taxas SWAP. Conforme critério de apuração da BM&F o dado de fechamento das taxas de SWAP PRÉ DI são obtidos a partir dos dados do ajuste dos contratos de DI Futuro negociado na BM&F, não correspondendo a dados de negócios efetivamente executados sob a modalidade SWAP. Logo, como os dados de SWAP são obtidos através do próprio dado do DI Futuro ou através de uma interpolação deste, consideramos que ao utilizar apenas o próprio dado de DI Futuro o modelo estará livre de distorções decorrentes da utilização das taxas divulgadas de SWAP como se fossem informações de negócios efetivamente executados. Assim, na interpolação para obtenção das taxas das maturidades fixas utilizados no modelo, serão utilizados como fonte de informação os dados de ajuste dos contratos de DI Futuro, que refletem as taxas de negócios efetivamente executados, evitando-se nesse caso uma interpolação a partir

de dados que já são resultado de uma interpolação anterior.

A curva de juros para o período estudado contempla vários formatos, com várias mudanças de inclinação e curvatura, assumindo formatos ascendentes e invertidos por várias vezes nesse período. A figura 3.1 apresenta o gráfico em três dimensões da curva estudada.

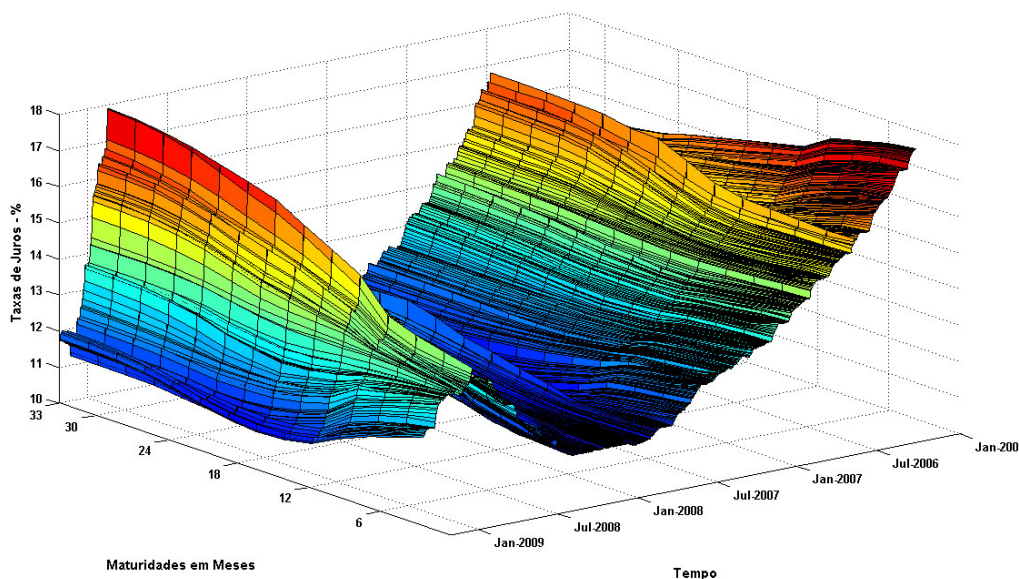


Figura 11: *Estrutura a Termo da Taxa de Juros (Jan 2006 to Feb 2009).*

Fonte: Elaboração própria.

Nota-se uma grande quantidade de variações temporais no nível da curva. O período analisado foi marcado por várias mudanças na condução da política monetária brasileira. Essas alterações na condução da política monetária se refletem nas taxas de juros praticadas no mercado de títulos públicos e privados, fazendo com que a estrutura a termo da taxa de juros brasileira assumira uma variedade de formatos nesse período. Por vários momentos a estrutura a termo da taxa de juros apresenta alterações não só no padrão de curvatura, mas também de inclinação. Assim, o período analisado parece bastante apropriado para verificar a capacidade preditiva do modelo de fatores de Nelson and Siegel (1987), estendido por Diebold and Li (2006).

3.8 Avaliação Empírica do Modelo para Dados do DI-Futuro

Na seção 2 o modelo de Diebold and Li (2006) foi apresentado no formato estado espaço, com um VAR(1) para a equação de transição, que modela a dinâmica dos fatores, e uma equação de medida linear relacionando as taxas de juros observadas ao vetor de estados. Os parâmetros foram estimados simultaneamente por máxima verossimilhança através do Filtro de Kalman, que além de ser o estimador eficiente, evita a problemática de como calibrar o parâmetro de decaimento. As taxas de juros empregadas consistem de dados diários das taxas

de DI-futuro para período de janeiro de 2006 a fevereiro de 2009, perfazendo um total de 772 observações diárias para cada uma das 14 maturidades, sendo que as últimas 252 observações (um ano comercial) foram utilizadas para análise fora da amostra.

A minimização do negativo do logaritmo da função de verossimilhança foi obtida usando o método de quasi-Newton com atualizações do inverso da matriz hessiana através do método BFGS. Mais especificamente, foi usado o algoritmo csminwel desenvolvido por Christopher Sims para ser robusto à certas patologias comuns à funções de verossimilhança como descontinuidades no hiperplano, o qual está disponível em <http://sims.princeton.edu/yftp/optimize/>. O algoritmo foi configurado para encerrar as iterações quando não for mais possível aumentar o valor da função em pelo menos $1.0e-05$.

Diferentemente da estimação em dois passos, na estimação por Filtro de Kalman, os parâmetros são estimados em uma única etapa, sendo o parâmetro λ , que governa a velocidade de decaimento dos factor-loadings do nível e curvatura, estimado juntamente com os demais, e não determinado a priori. Os valores iniciais dos parâmetros para inicialização do Filtro de Kalman foram obtidos a partir da estimação dos parâmetros de Diebold and Li (2006) pelo método de dois passos. A figura 3.2 abaixo traz os pesos dos fatores para o nível, inclinação e curvatura, obtidos a partir do parâmetro λ estimado.

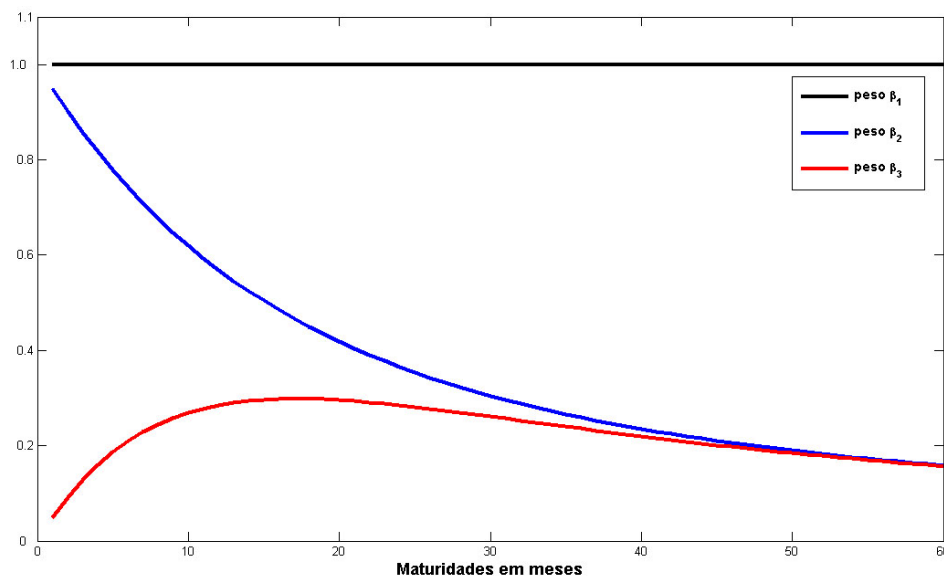


Figura 12 Pesos dos Fatores ($\hat{\lambda} = 0.1047$).

Fonte: Elaboração própria.

Com um λ estimado de 0.1047, o factor-loading relativo à curvatura assume valor máximo para maturidades entre 13 e 18 meses.

O principal argumento em favor do modelo com três fatores de Diebold and Li (2006) é a capacidade de gerar boas previsões. Mesmo não sendo o melhor modelo quando o objetivo é o ajuste da estrutura a termo da taxa de juros, a formulação proposta por Diebold-Li é capaz

de replicar os vários formatos assumidos pela curvas de juros. A figura 3.3 traz os dados reais da curva de juros para alguns dias e a curva ajustada pelos parâmetros do modelo de fatores estimado.

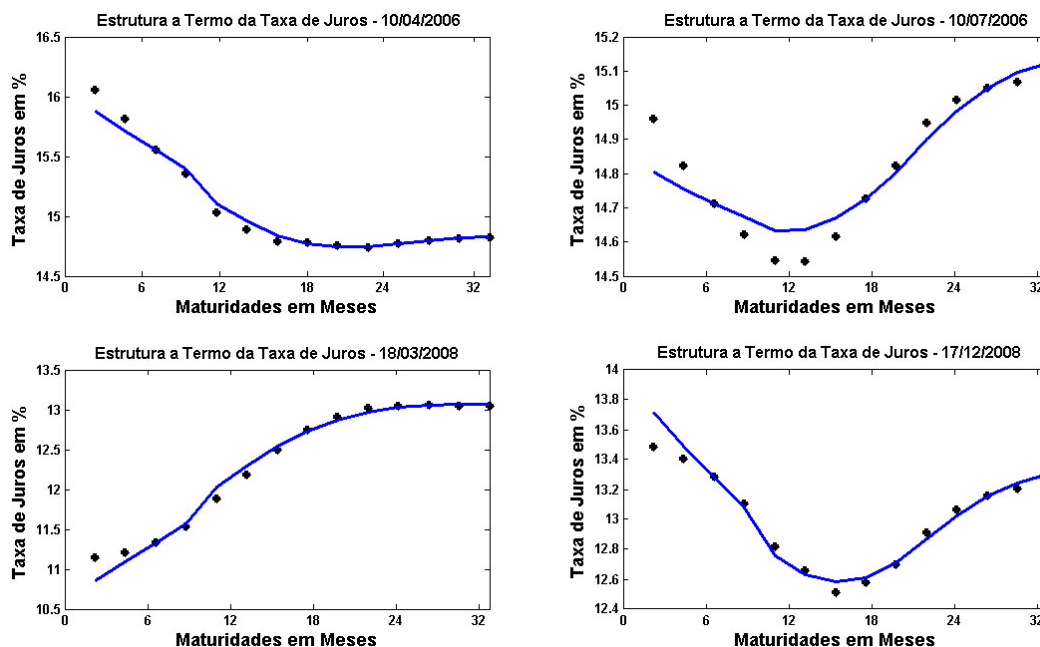


Figura 13: Ajuste da Curva de Juros para Diferentes Pontos do Tempo.

Fonte: Elaboração própria.

Nota-se que o modelo estimado com três fatores ajusta-se bem a uma variedade de formatos da curva de juros: positivamente inclinada, negativamente inclinada e com distintas formas de curvaturas. A figura 4 plota os resíduos diários da curva de juros obtidos com o modelo ajustado. Pode-se observar que os resíduos não apresentam comportamento sistemático, e são de baixa magnitude, evidenciando a capacidade do modelo de replicar os padrões apresentados pela curva de juros no período. A observação do gráfico permite notar também que os resíduos apresentam maior volatilidade para as maturidades mais curtas. Tal situação é considerada um fato estilizado quando se trata de curvas de juros, maturidades mais curtas são mais voláteis que as taxas para maturidades mais longas. Uma das possíveis explicações é que as maturidades mais curtas estão mais suscetíveis aos movimentos da taxa básica de juros - taxa SELIC - que é instrumento de política monetária.

O nível estimado da curva de juros, β_1 , tem média de 13,72, estatisticamente significativa e exibindo elevada persistência. Pode-se observar que o nível da curva de juros apresentou comportamento mais volátil no período a partir de agosto de 2008, quando a crise dos mercados financeiros refletiu fortemente nos ativos negociados no mercado brasileiro. Foi também nesse período que o nível da curva de juros assumiu seu maior valor para o período analisado, 19,47. Para o período em questão nota-se brusca mudança de comportamento também para os fatores latentes de inclinação e curvatura. Fica claro nas figuras 5.1 a 5.3 que não só o

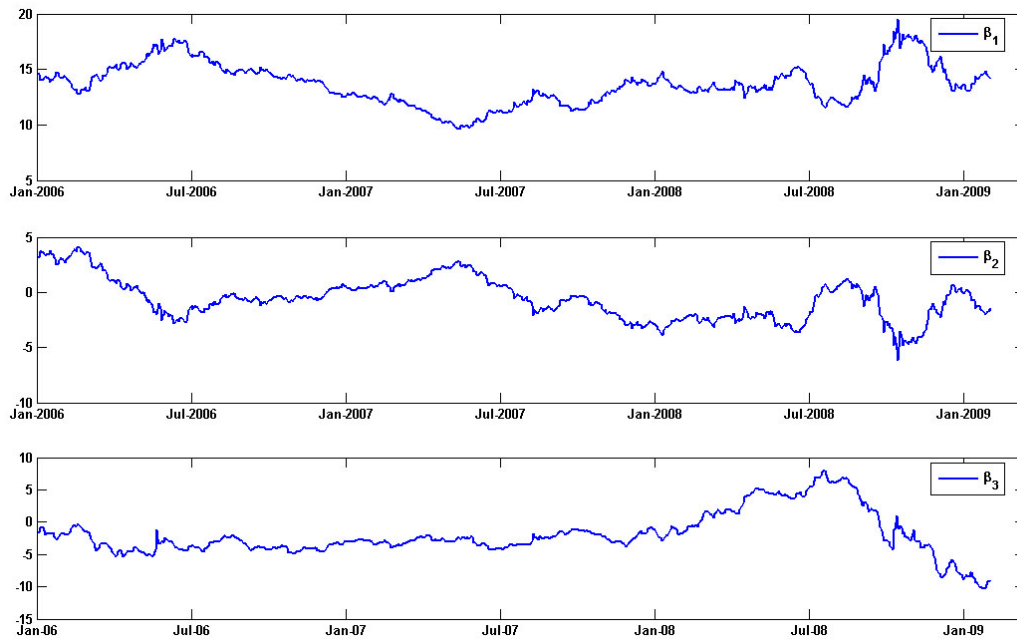


Figura 14: *Estimativas Suavisadas dos Fatores.*

Fonte: Elaboração própria.

nível, mas também a inclinação e curvatura apresentam flutuações ao longo do período, e as autocorrelações revelam alta persistência desses dois fatores.

A avaliação da capacidade preditiva do modelo é realizada dividindo a amostra em duas partes. Uma parte é usada para fazer a estimação do modelo, compreendendo 520 observações, com dados de Janeiro de 2006 a Janeiro de 2008. A segunda parte é usada para avaliar performance das previsões geradas pelo modelo, com dados de Fevereiro de 2008 a Fevereiro de 2009, perfazendo 252 observações. São realizadas previsões para horizontes de um dia, um mês, três meses e seis meses a frente com base num esquema *rolling window* com janelas crescentes. Para completar o exercício de previsão são incluídas previsões obtidas via *random walk* e com o modelo de Diebold-Li estimado pelo método de dois passos (RW e 2S)³.

As tabelas 3.1 e 3.2 apresentam os parâmetros estimados do VAR e a matriz de covariância dos fatores estimados. As covariâncias entre dois fatores são estatisticamente significantes para todos os pares, evidenciando que o VAR é a estrutura mais adequada para capturar a dinâmica dos fatores.

A análise dos parâmetros do VAR indica alta persistência na dinâmica dos três fatores latentes. Os efeitos cruzados estatisticamente significantes para a dinâmica dos fatores são observados de $\beta_{1,t-1}$ em $\beta_{2,t}$, e $\beta_{2,t-1}$ em $\beta_{3,t}$, reforçando a importância da especificação dos fatores como um VAR.

³Na estimação por dois passos o valor do λ foi de 0.1182, o qual minimiza o RMSE. Este valor maximiza o peso do fator curvatura para maturidade média entre 13 e 15 meses.

Tabela 6: Parâmetros do VAR Estimados

	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$	μ
$\beta_{1,t}$	0.991*** (0.002)	0.003* (0.002)	0.001 (0.002)	13.72*** (1.01)
$\beta_{2,t}$	0.008*** (0.003)	0.993*** (0.002)	0.002 (0.001)	-.72 (1.35)
$\beta_{3,t}$	-0.003 (.002)	-.0012*** (.003)	1.000*** (0.002)	1.46 (1.66)

Fonte: Elaboração própria.

Nota: Os erros padrões entre parênteses são calculados como a raiz quadrada dos elementos da diagonal do inverso da matriz Hessiana. *, **, e *** indica significância estatística aos níveis de 10%, 5% e 1% respectivamente.

Tabela 7: Matriz de Covariância do VAR Estimado

	$\eta_{1,t}$	$\eta_{2,t}$	$\eta_{3,t}$
$\eta_{1,t}$	0.25*** (0.001)	-0.07*** (0.001)	-0.04*** (0.002)
$\eta_{2,t}$		0.26*** (0.001)	0.03 (0.002)
$\eta_{3,t}$			0.40*** (0.003)

Fonte: Elaboração própria.

Nota: Os erros padrões entre parênteses são calculados como a raiz quadrada dos elementos da diagonal do inverso da matriz Hessiana. *, **, e *** indica significância estatística aos níveis de 10%, 5% e 1% respectivamente.

A abordagem para fazer previsão da curva de juros com o modelo de Diebold-Li consiste em fazer previsões dos fatores, e então usar os fatores previstos para ajustar previsões da curva de juros. As previsões no tempo t , para $t + h$, da taxa de juros com maturidade τ , são dadas por:

$$\hat{y}_{t+h|t}(\tau) = \hat{\beta}_{1,t+h|t} + \hat{\beta}_{2,t+h|t} \left(\frac{1 - e^{-\hat{\lambda}\tau}}{\hat{\lambda}\tau} \right) + \hat{\beta}_{3,t+h|t} \left(\frac{1 - e^{-\hat{\lambda}\tau}}{\hat{\lambda}\tau} - e^{-\hat{\lambda}\tau} \right).$$

As previsões dos fatores são obtidas usando os parâmetros estimados da equação de transição (3.4):

$$\hat{\beta}_{t+h|t} = \hat{\mu} + \hat{A} \cdot \hat{\beta}_t.$$

As tabelas 3.3 e 3.4 apresentam os RMSE para as previsões fora da amostra realizadas com o modelo estimado através do Filtro de Kalman (FK), para horizontes de um dia, um mês, três meses e seis meses à frente. São apresentados também os RMSE para os mesmos horizontes obtidos através do *Random Walk* e pelo modelo de Diebold-Li estimado pelo método de dois passos para comparação dos resultados.

Para horizonte de previsão um dia à frente o modelo estimado através do Filtro de Kalman supera o *random walk* apenas para as maturidades médias de 3, 15, 23 e 26 meses. Para as demais maturidades, embora sejam melhores que as previsões obtidas quando o modelo é estimado pelo método de dois passos, são piores que o *random walk*. Vale destacar que o pior desempenho é observado para a maturidade mais curta, que conforme mencionado anteriormente é mais suscetível a movimentos na taxa Selic. Entretanto a qualidade das previsões melhora substancialmente quando o horizonte é ampliado. Para previsões um mês à frente o modelo estimado através do Filtro de Kalman supera os concorrentes em todas as maturidades, excluindo a mais curta. Para as maturidades intermediárias, entre 3 e 24 meses, as previsões obtidas por FK apresentam RMSE em média 15 pontos base menores que RW e 2S.

Quando são analisadas as previsões para os horizontes 3 e 6 meses nota-se que o FK supera consistentemente seus competidores. Para previsões 3 meses à frente o FK supera o RW e 2S em todas as maturidades, apresentando RMSE em média 35 pontos base inferior àqueles obtidos por 2S.

Já para as previsões com horizonte de 6 meses o FK supera o RW em todas as maturidades, mas perde para 2S para as maturidades mais longas. A menor liquidez de contratos com maturidades mais longas pode levar a piora na qualidade das previsões, conforme Bali et al. (2006), liquidez é um fator relevante para a previsibilidade das taxas de juros.

Para confirmar se as diferenças entre as previsões fora da amostra obtidas através do filtro de Kalman são estatisticamente significantes, nós aplicamos o teste de Diebold Mariano (DM) para comparar as previsões (see Diebold and Mariano, 1995; McCracken, 2007). Nós

Tabela 8: RMSE para Previsões Fora da Amostra (Fev 2008 a Fev 2009)

Maturidade	Um Dia a Frente			Um Mês a Frente		
	RW	2S	FK	RW	2S	FK
1	4.13	10.81	26.06	38.55	48.30	47.67
2	4.57	5.96	11.61	41.25	51.41	37.42
3	5.44	10.74	5.10	45.33	56.01	32.90
4	6.94	14.02	9.71	52.54	62.94	36.23
7	10.49	16.72	16.02	70.18	78.20	51.58
9	13.01	17.36	16.59	80.78	88.53	63.01
12	15.76	18.39	16.29	91.76	99.35	75.86
15	18.02	20.18	17.92	98.02	106.21	84.61
19	19.18	21.15	19.32	105.00	110.52	90.86
23	20.65	21.79	20.36	111.21	113.80	95.91
26	21.53	22.74	21.22	117.03	117.72	101.20
29	22.20	24.59	22.06	121.28	121.86	105.86
31	22.66	26.36	22.99	123.64	124.93	109.09
32	22.84	27.70	23.73	124.79	126.89	111.01

Fonte: Elaboração própria.

Nota: RMSE expresso em pontos base, e maturidades em meses (RW = Modelo Random Walk, 2S = Estimação por Dois Passos, KF = Estimação por Filtro de Kalman).

Tabela 9: RMSE para Previsões Fora da Amostra (Fev 2008 a Fev 2009)

Maturidade	Três Meses a Frente			Seis Meses a Frente		
	RW	2S	FK	RW	2S	FK
1	88.44	100.69	71.49	141.17	133.81	96.20
2	94.41	106.67	63.03	144.59	133.30	91.99
3	103.89	112.69	62.97	151.28	132.77	95.78
4	117.66	121.39	71.31	161.65	134.04	108.59
7	146.13	140.99	97.11	189.94	144.11	144.43
9	161.48	153.07	114.38	205.41	150.81	166.10
12	176.11	164.42	131.42	222.08	158.63	186.12
15	184.61	170.51	140.98	230.55	162.08	196.50
19	186.01	172.97	144.58	229.31	161.33	197.64
23	187.03	173.43	144.46	223.25	157.71	192.37
26	189.83	173.79	144.18	219.10	154.77	186.64
29	192.88	175.16	144.63	214.69	153.46	182.12
31	194.89	176.22	145.20	212.30	152.78	179.17
32	196.01	177.05	145.72	211.16	152.52	177.48

Fonte: Elaboração própria.

Nota: RMSE expresso em pontos base, e maturidades em meses (RW = Modelo Random Walk, 2S = Estimação por Dois Passos, KF = Estimação por Filtro de Kalman).

comparamos aos pares as previsões geradas através do filtro de Kalman, método de dois passos e *radom walk*. Seja $\{d_i\}_{i=1}^n$ uma função das diferenças dos quadrados dos erros de previsão produzidas por dois modelos. Nós escrevemos d_i como:

$$d_i = (\hat{y}_{t+h,i}(\tau) - y_{t+h}(\tau))^2 - (\hat{y}_{t+h,j}(\tau) - y_{t+h}(\tau))^2, \quad (3.14)$$

onde $i = FK$ ou $2S$ e $j = 2S$ ou RW . As variáveis $\hat{y}_{t+h,i}$ são previsões h -passos à frentes no tempo t de FK, 2S e RW, respectivamente. Diebold & Mariano propõem um teste para checar se a diferença das médias das perdas $\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i$ é estatisticamente diferentes de zero, o qual é dado por:

$$DM = \frac{\bar{d}}{\sqrt{\frac{\hat{\delta}}{n}}} \xrightarrow{d} N(0,1), \quad (3.15)$$

onde $\hat{\delta}$ é a estimativa da matriz de covariância de longo prazo de d_i . Nós empregamos as estimativas de Newey and West (1987) para $\hat{\delta}$, à qual permite controlar para correlação serial nos erros de previsão. As tabelas 3.5 and 3.6 apresentam as estatísticas de Diebold-Mariano para uma função de perda quadrática. Valores negativos indicam a superioridade do primeiro método do par. Valores absolutos grandes indicam elevada probabilidade de rejeitar a hipótese nula (diferença entre as médias dos erros quadráticos é negligenciável). valores absolutos maiores que 1.96 indicam rejeição da hipótese nula a 95% de confiança. As estatísticas de Diebold-Mariano corroboram com a maioria das conclusões acima, tais como, bons resultados das previsões obtidas através do filtro de Kalman para horizontes de um mês e três meses, superando o método de 2S para horizonte de um dia e apresentando performance levemente inferior para previsões seis meses à frente.

Uma possível alternativa para melhorar a previsibilidade do modelo pode ser obtida levando-se em conta a presença de heterocedasticidade condicional dos fatores, que pode ser tratada através da inclusão de um componente GARCH para os fatores latentes.

3.9 Conclusão

No presente trabalho o modelo de Diebold e LI, normalmente estimado para dados brasileiros através do método ineficiente de dois passos, foi colocado no formato espaço de estados e estimado eficientemente por máxima verossimilhança usando o filtro de Kalman. A estimação por máxima verossimilhança permite a estimação conjunta de todos os parâmetros do modelo, evitando a seleção a priori do parâmetro de decaimento. Estimativas suavizadas dos parâmetros, que consideram toda a informação da amostra para inferir sobre a série de tempo dos fatores, foram obtidas através do suavizador de Kalman e usadas para previsão fora da amostra. Os resultados mostram que o modelo estimado por máxima verossimilhança produz previsões fora da amostra de melhor qualidade que o modelo estimado em duas etapas para

Tabela 10: Teste de Diebold-Mariano para previsões fora da amostra

Maturidade (Meses)	Um dia à Frente			Um dia mês Frente		
	FK x 2S	FK x RW	2S x RW	FK x 2S	FK x RW	2S x RW
1	2.66	2.94	4.05	0.58	1.99	4.58
2	3.30	4.22	4.39	-0.15	-1.59	0.90
3	-3.50	-0.58	3.46	-0.90	-3.18	-0.09
4	1.00	3.28	3.38	-1.51	-3.68	-0.11
7	1.05	2.96	5.27	-2.24	-3.17	-0.16
9	-1.35	3.31	6.15	-2.46	-2.52	-0.03
12	-7.63	1.45	8.22	-2.29	-1.63	0.21
15	-5.83	-0.66	6.20	-2.19	-1.73	0.27
19	-5.23	0.61	4.95	-1.56	-1.82	-0.21
23	-5.04	-0.79	4.68	-1.41	-2.10	-1.01
26	-4.27	-1.25	3.99	-1.36	-2.20	-1.36
29	-3.17	-0.60	2.77	-1.32	-2.13	-1.48
31	-2.30	1.13	2.19	-1.31	-2.13	-1.52
32	-1.63	2.22	2.07	-1.23	-2.13	-1.56

Fonte: Elaboração própria.

Nota: Estatísticas de Diebold-Mariano para um dia e um mês à frente, comparando os métodos aos pares. Valores negativos indicam superioridade das previsões obtidas com o primeiro método do par.

Tabela 11: Teste de Diebold-Mariano para previsões fora da amostra

Maturidade (Meses)	Três Meses à Frente			Seis Meses à Frente		
	FK x 2S	FK x RW	2S x RW	FK x 2S	FL x RW	2S x RW
1	-1.95	-1.30	-0.93	-0.18	-1.55	-1.17
2	-8.35	-1.89	-1.04	-0.17	-17.91	-7.18
3	-1.31	-2.24	-1.24	-0.09	-23.38	-5.30
4	-1.14	-2.76	-1.28	0.09	-9.88	-7.51
7	-1.03	-3.53	-1.40	2.44	-9.16	-2.15
9	-1.23	-3.45	-1.39	0.78	-21.02	-1.43
12	-6.85	-3.25	-1.38	0.73	-12.93	-1.58
15	-5.99	-3.32	-1.51	0.60	-32.44	-1.81
19	-2.16	-3.52	-1.83	0.52	-1.77	-1.94
23	-0.77	-3.07	-2.23	0.42	-1.68	-1.99
26	-0.78	-2.71	-2.50	0.35	-1.64	-1.92
29	-0.82	-2.45	-2.48	0.86	-1.62	-1.89
31	-0.82	-2.37	-2.31	9.67	-1.62	-1.90
32	-0.84	-2.34	-2.18	4.91	-1.64	-1.92

Fonte: Elaboração própria.

Nota: Estatísticas de Diebold-Mariano para um dia e um mês à frente, comparando os métodos aos pares. Valores negativos indicam superioridade das previsões obtidas com o primeiro método do par.

todos os diferentes horizontes de previsão. Além disso, as previsões baseadas no modelo estimado por máxima verossimilhança são superiores às do modelo de passeio aleatório para todas as maturidades quando são considerados horizontes de um mês, três meses e seis.

Os modelos de fatores da curva de juros são amplamente usados por bancos centrais de vários países e por muitos participantes do mercado para ajustar e prever a estrutura a termo da taxa de juros. Os resultados obtidos mostram a flexibilidade e capacidade do modelo para ajustar a uma ampla variedade de formatos da curva de juros, e que a estimação por filtro de Kalman é superior a estimação por dois passos. Uma possível sugestão para investigações futuras é a estimação do modelo usando quatro fatores, como proposto por Cochrane and Piazzesi (2005), o qual inclui uma curvatura adicional que melhora a previsibilidade em mercados com curvas de juros mais voláteis, como ocorre em mercados emergentes.

4 Inferência Bayesiana Aplicada ao Modelo Dinâmico de Nelson-Siegel com Volatilidade Estocástica nos Fatores

Neste artigo nós propomos estimar o modelo dinâmico da estrutura a termo da curva de juros de Nelson e Siegel (1987) considerando duas especificações alternativas. Na primeira, nós consideramos os pesos dos fatores como variantes no tempo e tratamos a heterocedasticidade condicional via um modelo volatilidade estocástica com fatores comuns. No segundo caso, consideramos um modelo onde os fatores latentes seguem individualmente processos autoregressivos com volatilidade estocástica. Os assim chamados fatores de volatilidade buscam capturar a incerteza ao longo do tempo associada ao nível, inclinação e curvatura da curva de juros. A estimação é realizada através de métodos de inferência bayesiana, por Markov Chain Monte Carlo. Os resultados mostram que os fatores de volatilidade são altamente persistentes, dando suporte ao fato estilizado de que os choques na volatilidade das taxas de juros são altamente persistentes, e também indicam que o uso de estruturas de volatilidade estocástica levam a melhores ajustes dentro da amostra para a curva de juros observada.

4.1 Introdução

Ajustar e prever a estrutura a termo das taxas de juros é de extrema relevância para formuladores de política econômica. Em finanças, as informações reveladas pela curva de juros são importantes para a precificação de títulos, derivativos da taxa de juros, gestão de carteiras e alocação de ativos. Na macroeconomia, a curva de juros traz importantes informações sobre o estado da economia e dos negócios. Assim, muitos pesquisadores tem se detido sobre o tema, fazendo florescer uma extensa literatura tratando do ajuste e previsão da estrutura a termo¹. Apesar de existir um grande número de trabalhos relacionados ao tema, poucos se dedicaram a ajustar e prever a curva de juros levando em conta a heterocedasticidade condicional presente na estrutura a termo da curva de juros.

Este problema é de grande importância já que usualmente nestes modelos o pressuposto de volatilidade constante nas taxas de juros tem implicações práticas importantes sobre as políticas de gerenciamento de risco, que podem ser demasiadamente simples ignorando o risco de uma estrutura de volatilidade que se altera no tempo. Operações de hedge e arbitragem de taxas de juros também podem ser influenciadas pela presença de volatilidade variante no tempo, já que nestas operações é necessária incorporar compensações para o preço de mercado de risco de taxas de juros. Outra implicação importante é que na presença de volatilidades condicionais os intervalos de confiança dos ajustes e previsões derivadas destes modelos serão calculados de forma incorreta em amostras finitas.

¹Veja por exemplo Filipovic (2009) para uma revisão sobre modelagem de taxas de juros.

Para analisar as consequências da presença de estruturas de volatilidade variantes no tempo, é importante recuperar as características das principais classes de modelos utilizados na modelagem da estrutura a termo de taxas de juros. Estes modelos podem ser divididos em três classes principais: os modelos de não arbitragem; os modelos de equilíbrio; e os modelos estatísticos.

Os modelos de não arbitragem para a estrutura a termo de taxas de juros são obtidos através da imposição de condições de consistência entre as curvas de juros de diversas maturidades que impeçam a existência sistemática de oportunidades de arbitragem, e são baseados na existência de uma medida martingale equivalente. Exemplos de modelos desta classe são os modelos de Hull and White (1990), e o framework geral que define estes modelos está colocado na contribuição fundamental de Heath et al. (1992). Em relação ao problema de volatilidades variantes no tempo, note que a estrutura definida em Heath et al. (1992) permite a possibilidade de estruturas de volatilidade estocástica.

No entanto, a presença de estruturas de volatilidade estocástica na classe de modelos HJM leva a dois problemas importantes. O primeiro problema é dado forma como estes modelos são normalmente utilizados na precificação de instrumentos derivativos. Nesta utilização, os parâmetros do modelo HJM são obtidos pela calibração perfeita da curva de juros observada em um determinado dia através de preços de instrumentos de mercado. Embora esta forma permita a precificação livre de arbitragem de instrumentos derivativos, ela não permite a realização de previsões para a curva de juros, já que este ajuste é de natureza cross-section e não permite recuperar diretamente a dinâmica observada na estrutura a termo da taxas de juros. O segundo ponto importante é que usualmente a imposição de estruturas gerais de volatilidade estocástica em modelos de não-arbitragem para a estrutura a termo de taxas de juros leva a dinâmicas não Markovianas para estas curvas de juros, o que dificulta o uso de métodos econométricos para a estimação de parâmetros, como por exemplo o uso do filtro de Kalman.

A segunda classe de modelos, os modelos de equilíbrio, são obtidos através da imposição de condições de equilíbrio entre os rendimentos das diversas maturidades na curva de juros. Estes modelos são normalmente baseados na modelagem da taxa a termo instantânea, utilizando tipicamente a estrutura de modelos modelos afins, em que as taxas de outros vencimentos podem ser derivadas assumindo que o prêmio de risco é dado por uma função afim. Modelos desta classe foram desenvolvidos em Vasicek (1977), Cox et al. (1985), Duffie and Kan (1996) ou em Dai and Singleton (2002). Estes modelos, embora tenha a vantagem de ser analiticamente tratáveis e permitirem a obtenção de fórmulas fechadas para precificação de títulos e derivativos, tem um problema fundamental no ajuste e previsões para a curva de juros. A literatura aponta que estes modelos são normalmente deficientes nestes aspectos, e as previsões geradas pelos modelos de equilíbrio não superam as previsões do modelo de passeio aleatório, como mostrado, por exemplo, em Duffie and Kan (1996). Em relação ao problema da presença de volatilidades variantes no tempo, a imposição de estruturas de volatilidade estocástica nesta classe de modelos leva a problemas como a não existência de formulações

afins ou então a necessidade de modelos não-homogeneos no tempo, gerando a necessidade do uso de métodos de calibração, e que da mesma forma que na classe de modelos de não-arbitragem, impede o ajuste dinâmico da curva de juros e a elaboração de previsões. Uma discussão geral destes problemas pode ser encontrada em Brigo and Mercurio (2006).

A terceira principal classe de modelos é dada pelos modelos estatísticos. Estes modelos são obtidos como representações puramente estatísticas da evolução da estrutura a termo de taxas de juros, e assim são obtidos sem a imposição direta de condições de não-arbitragem ou equilíbrio. Esta classe é composta principalmente por modelos de componentes principais, modelos de fatores ou de variáveis latentes, bem como por modelos de interpolação. De acordo com Matzner-Lober and Villa (2004), grande parte da intuição a respeito da dinâmica da rentabilidade de títulos e bônus provém de modelos dessa classe, como em Robert and Scheinkman (1991) e em Pearson and Sun (1994).

Dentre os modelos estatísticos, o modelo de Nelson-Siegel (Nelson and Siegel (1987)) e suas variantes, são os mais populares entre gestores de renda fixa e bancos centrais. A atratividade dos modelos de fatores do tipo Nelson Siegel se deve à sua parcimonia e boa performance empírica. Modelos deste tipo conseguem capturar a maior parte da evolução da estrutura a termo da taxa de juros através do uso de apenas três fatores. Várias extensões do modelo de Nelson-Siegel foram propostas (ver Almeida et al. (2007), Laurini and Hotta (2008), Bjork and Christensen (1999), Rudebusch and Wu (2008), entre outros).

Os trabalhos de Diebold and Li (2006) e Diebold et al. (2006) voltaram atenção para o modelo de Nelson-Siegel e reinterpretaram o modelo dinâmico de Nelson-Siegel como um modelo estatístico de três fatores para descrever a curva de juros ao longo do tempo. Os três fatores são interpretados como nível, inclinação e curvatura, de forma consistente a interpretação obtida em Robert and Scheinkman (1991), em que a interpretação tem algum sentido econômico. A principal contribuição é que Diebold e Li e Diebold, Rudebusch e Aruoba mostram que as previsões geradas superam muitos outros modelos, incluindo modelos de séries de tempo, tais como modelos de vetores autoregressivos (VAR) e modelos dinâmicos de correção de erro (ECM). Diebold et al. (2006) estendem a estrutura incluindo fatores macroeconômicos não latentes, como inflação e hiato do produto. Além disso, o modelo de Nelson-Siegel é colocado no formato de espaço de estados, onde os três fatores são tratados como componentes não observados e são modelados por um vetor autoregressivo.

A estimação dos parâmetros, em Diebold and Li (2006) e Diebold et al. (2006), parte de duas suposições simplificadoras. Os pesos dos fatores do modelo de Nelson-Siegel dependem de um único parâmetro de peso (decaimento), e para viabilizar a estimação dos fatores latentes variantes no tempo em um arcabouço linear, os pesos dos fatores são mantidos constantes ao longo do tempo para cada maturidade. Em Diebold and Li (2006), o parâmetro de decaimento é determinado exogenamente, sendo mantido fixo em toda a amostra, conseqüentemente, os pesos dos fatores não variam com o tempo. A segunda hipótese, também assumida em Diebold et al. (2006), Yu and Zivot (2008), Almeida et al. (2007), Caldeira et al. (2010) diz

respeito à volatilidade, que é mantida constante para todas as maturidades e ao longo de todo o período amostral.

Estes pressupostos estão relacionados a maior dificuldade de estimação na presença de volatilidades e fatores de decaimento (peso) variantes no tempo. Ao abandonar estes pressupostos, a estimação pelos métodos usuais de mínimos quadrados ordinários não pode ser mais utilizada, e também as técnicas de estimação simultânea como a verossimilhança obtida através da decomposição do erro de previsão através do filtro de Kalman utilizadas em Caldeira et al. (2010) já não podem ser utilizadas pelo fato de que com a representação com volatilidades e decaimentos variantes no tempo não é possível obter uma representação em espaço de estados linear e gaussiana.

Outro problema está relacionado com a alta dimensionalidade da estrutura a termo, que é tornada mais problemática ainda com o uso de volatilidades estocástica multivariadas. Neste problema, se a estrutura de volatilidade é adotada para todos os fatores latentes, o processo de volatilidade se torna um processo multivariado, o que torna os problemas de estimação bastante mais complicados, já que modelos multivariados de volatilidade estocástica apresentam problemas de inferência não triviais, como pode ser visto por exemplo em Chib et al. (2009). Algumas formas de tratar a questão foram propostas por Koopman et al. (2010), Laurini and Hotta (2008) e Haustsch and Ou (2009). Em Koopman et al. (2010), é adotada uma estrutura de volatilidade fatorial para as variáveis latentes, baseada em uma especificação baseada na classe de modelos GARCH. Esta representação, no entanto, apresenta algumas limitações - para evitar os problemas de identificação e estimação em modelos GARCH multivariados, a estrutura de volatilidades é baseada em uma estrutura basicamente univariada, já que nesta especificação a volatilidade dos demais fatores latentes é obtida como um deslocamento linear de apenas um fator para a volatilidade. Também é importante notar que a forma de estimação utilizada é baseada em uma aproximação para a verossimilhança do processo, que pode não ser ter propriedades de otimalidade estatística. Outra dificuldade inerente a este modelo é a incompatibilidade de modelos GARCH com representações em tempo contínuo, o que limita a interpretação destes modelos em termos de processos de precificação em finanças.

Em Laurini and Hotta (2008) e Haustsch and Ou (2009), o processo de volatilidade estocástica é baseado nos modelos conhecidos como modelos log-normais de volatilidade estocástica, introduzidos por Taylor (1986). No modelo de Laurini and Hotta (2008), a especificação log-normal de volatilidade estocástica é assumida apenas para os erros de medida da curva de juros, enquanto que a estrutura de fatores latentes é assumida homocedástica. Na especificação proposta por Haustsch and Ou (2009), é assumido uma estrutura de volatilidade estocástica log-normal para cada fator latente, assim obtendo uma representação mais geral para o processo de volatilidade estocástica. Note que, nestas duas abordagens, a estimação é baseada em métodos Bayesianos utilizando Markov Chain Monte Carlo, devido às dificuldades inerentes a estimação de modelos de volatilidade estocástica, em que não é possível obter fórmulas fechadas para a avaliação da função de verossimilhança.

Neste trabalho, apresentamos uma generalização das estruturas de volatilidade apresentadas em Koopman et al. (2010), Laurini and Hotta (2008) e Hautsch and Ou (2009). Utilizando dados dos contratos de DI-Futuro negociados na BM&F, nós estimaremos o modelo dinâmico de Nelson-Siegel com os pesos dos fatores e volatilidade variando no tempo. Nós consideramos duas especificações para modelar a volatilidade condicional. Na primeira, análoga à utilizada em Koopman et al. (2010), o grau de risco inerente à estrutura a termo é capturado via um componente comum de volatilidade. Na segunda alternativa, propomos modelar diretamente a volatilidade estocástica de cada fator latente do modelo de Nelson-Siegel. Modelar a volatilidade diretamente nos fatores, além de ser uma abordagem mais parcimoniosa, permite capturar a incerteza associada ao nível, inclinação e curvatura da curva de juros. Assim, a chamada volatilidade do nível reflete a volatilidade relacionada ao nível geral dos juros, enquanto a volatilidade da inclinação captura a incerteza inerente ao spread entre as taxas de longo e curto prazos. Já a volatilidade de curvatura, está associada ao risco de mudanças na curvatura da estrutura a termo. Nesta especificação, também é possível modelar o erro de medida da curva de juros como um processo com volatilidade estocástica, na forma utilizada em Laurini and Hotta (2008).

Ao contrário das especificações utilizadas em Koopman et al. (2010) e Hautsch and Ou (2009), nas duas especificações propostas o parâmetro que determina o peso dos fatores, λ_t , é também considerado como um processo latente que varia estocasticamente ao longo do tempo, sendo tratado como um quarto fator latente. Esta adição é importante para a modelagem de curvas de juros de países emergentes, caracterizadas pelas frequentes mudanças nos formatos e maturidades máximas nas curvas de juros, como discutido em Laurini and Hotta (2008). A dinâmica dos fatores latentes de nível, inclinação e curvatura, juntamente com o parâmetro que determina o peso dos fatores é modelada conjuntamente como um processo autoregressivo.

A modelagem da heterocedasticidade condicional tem por objetivo melhorar a qualidade do ajuste e previsão do modelo, incorporando os períodos de mudança na volatilidade, e também capturar o fato estilizado da presença de diferentes estruturas de volatilidade para cada maturidade da estrutura a termo, e também incorporar melhor fatos estilizados de curvas de juros dos mercados emergentes, que geralmente apresentam maior volatilidade do que curvas de juros de países desenvolvidos, caracterizadas por uma maior estabilidade no tempo e menores volatilidades.

Este artigo está estruturado da seguinte forma - na seção 2 revisamos a estrutura básica da classe de modelos de Nelson-Siegel; na seção 3 apresentamos as duas formulações estudadas de modelos com volatilidade estocástica; a seção 4 apresenta a metodologia de estimação Bayesiana utilizando Markov Chain Monte Carlo utilizada na estimação. A seção 5 apresenta a estrutura a termo de taxas de juros utilizada no estudo, de Futuros de DI, os resultados das estimações e as comparações entre modelos, e na seção 6 temos as conclusões finais.

4.2 Modelo de fatores de Nelson e Siegel para estrutura a termo da taxa de juros

Em um dado ponto do tempo, t , a curva de juros, denotada aqui por $y_t(\tau)$, é uma função representando a taxa de juros como uma função das maturidades τ . O modelo exponencial da curva de juros proposto por Nelson and Siegel (1987) e reinterpretado por Diebold and Li (2006) considera uma forma paramétrica para a evolução da estrutura a termo da taxa de juros ao longo do tempo, em que os coeficientes são tratados como nível, inclinação e curvatura. A curva de juros correspondente é:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \beta_{3,t} \left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right). \quad (4.1)$$

A forma da curva de juros é determinada pelos três parâmetros ($\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$) e pelos pesos associados a eles. O parâmetro λ_t , tratado como fixo em Diebold and Li (2006), governa a taxa de decaimento exponencial, pequenos (grandes) valores de λ_t estão associados a um decaimento suave (rápido), e ajustam melhor as maturidades mais longas (curtas). O peso do primeiro componente é 1 (constante) para todas as maturidades, assim $\beta_{1,t}$ representa o fator de longo prazo, e é interpretado como nível da curva de juros, que influencia igualmente as taxas de curto e longo prazo. O peso do segundo componente, $\left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right)$ começa em 1 e converge para zero monotonicamente e rapidamente, sendo $\beta_{2,t}$ interpretado como inclinação da curva de juros, ou fator de curto prazo, dado que este fator influencia muito as taxas de juros curto prazo. O peso do terceiro componente, $\left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau} \right)$, é uma função côncava, que assume valor zero para a maturidade zero, cresce, depois converge monotonicamente para zero nas maturidades mais longas. Assim, $\beta_{3,t}$ é interpretado como curvatura, ou fator de médio prazo, pois está associado às taxas de juros de médio prazo.

Em cada ponto do tempo existem taxas de juros $y_t(\tau)$ para diversas maturidades τ . Então, a equação pode ser estimada em cada ponto do tempo, de onde se obtém séries de tempo dos parâmetros, β_t . Assim, se o comportamento dos fatores latentes β_t puder ser modelado e previsto, pode-se obter previsões da curva de juros a partir deles. Diebold and Li (2006), Almeida et al. (2007), entre outros, adotam um procedimento de dois passos para estimar o modelo. No primeiro passo, são obtidas séries de tempo dos parâmetros através de regressões por mínimos quadrados em cada t , conseqüentemente as três séries de tempo dos β_t 's são modeladas por um processo autoregressivo de primeira ordem.

Uma alternativa consiste em colocar o modelo no formato de espaços de estados e estimar todos os parâmetros simultaneamente. O sistema, composto pela equação de transição e equação de medida, é representado da seguinte forma:

$$y_t(\tau) = \Lambda(\lambda_t) \beta_t + \varepsilon_t. \quad (4.2)$$

Onde $\Lambda(\lambda_t)$ é uma matriz $(N \times 3)$ de pesos dos fatores, que será variável no tempo apenas se o parâmetro de decaimento λ for variável.

$$\beta_t = (I - \Phi)\mu + \Phi\beta_{t-1} + \eta_t. \quad (4.3)$$

A equação de medida (4.2) define o vetor de taxas de juros $(T \times N)$ para as N maturidades, como a soma dos fatores multiplicados pelos seus pesos, com um vetor de erros distribuídos Gaussianamente e independente entre as maturidades. O vetor β_t de dimensão (3×1) representa os fatores, e Φ é a matriz de coeficientes do VAR que modela a dinâmica dos estados (fatores latentes) ao longo do tempo. Para estimar o modelo em formato linear, através de Filtro de Kalman, por exemplo, os pesos dos fatores precisam ser mantidos constantes ao longo do tempo para cada maturidade, fazendo $\lambda_t = \lambda$.

4.3 Modelos de Nelson-Siegel Generalizados

Os modelos introduzidos neste estudo são baseados no uso simultâneo de duas modificações sobre os modelos da classe Nelson-Siegel estudados na literatura. A primeira modificação é permitir que os parâmetros de decaimento λ_t sejam variantes no tempo, e a segunda modificação se refere a introdução de duas formas de modelagem da volatilidade estocástica, a primeira no uso de uma estrutura fatorial de volatilidade estocástica e a segunda o uso de processos separados para a volatilidade de cada fator latente. Note que nos modelos estudados também é permitido que a equação de medida da curva de juros também siga um processo de volatilidade estocástica. Estas modificações são detalhadas a seguir.

Ao se tratar o parâmetro λ_t como fixo ao longo do tempo, a maturidade para a qual a curvatura é máxima se mantém a mesma para todo o período, e a velocidade de decaimento do fator $\beta_{2,t}$ também se mantém inalterada. Aqui, assumimos que o parâmetro λ_t é variante no tempo ele passa a ser tratado como um quarto fator latente. Assim, o novo vetor de estados, β_t , agora com quatro parâmetros, incluindo $\lambda_t = \beta_{4,t}$, é modelado por um VAR. A nova equação de medida resultante desta alteração é não linear no parâmetro $\beta_{4,t}$.

$$y_t = \Lambda(\beta_{4,t}) (\beta_{1,t}, \beta_{2,t}, \beta_{3,t})' \quad (4.4)$$

Adotamos duas formulações na modelagem da volatilidade estocástica nesta classe de modelos. A primeira é analoga a formulação adotada por Koopman et al. (2010) com a especificação de um modelo com uma especificação de um componente comum de volatilidade, que é modelada por um processo GARCH. Nesta formulação a volatilidade dos fatores é um processo basicamente univariado, já que existe apenas um fator de volatilidade e a volatilidade dos demais fatores é dada por um deslocamento linear do fator comum. Nesse caso, o modelo é linearizado e estimado através de filtro de Kalman estendido, baseados em simplificações da verossimilhança do processo. Nós estimamos uma especificação semelhante, onde o pro-

cesso GARCH é substituído por um processo de volatilidade estocástica log-normal, sendo que a formulação utilizada é a de um modelo autoregressivo para o componente comum não observado de volatilidade estocástica, com:

$$\boldsymbol{\varepsilon}_t = \Gamma \boldsymbol{\varepsilon}_t^* + \boldsymbol{\varepsilon}_t^+, \quad t = 1, \dots, T \quad (4.5)$$

Onde Γ é um vetor de pesos ($N \times 1$), $\boldsymbol{\varepsilon}_t^*$ é um ruído escalar, e $\boldsymbol{\varepsilon}_t^+$ é um vetor de ruídos ($N \times 1$). Os ruídos são mutuamente independentes e suas distribuições são dadas por:

$$\boldsymbol{\varepsilon}_t^* \sim NID(0, \sigma_t^2), \quad \boldsymbol{\varepsilon}_t^+ \sim NID(0, \Sigma_\varepsilon^t), \quad t = 1, \dots, T. \quad (4.6)$$

Onde a variância σ_t^2 segue um processo ARSV(1), dado por:

$$\log(\sigma_t^2) = \mu_\sigma + \phi \log(\sigma_{t-1}^2) + \nu_t, \quad t = 1, \dots, T. \quad (4.7)$$

Em que o ruído da equação da volatilidade, ν_t , é um ruído branco gaussiano com variância σ_ν^2 .

A escolha pela formulação utilizando o modelo log-normal de volatilidade estocástica pode ser justificada por diversos fatores. Existem evidências empíricas indicando que modelos capturam melhor fatos estilizados de séries financeiras e possuem desempenho preditivo superior a outras classes de modelos de volatilidade (e.g. Koopman et al. (2005)). Estes modelos também possuem formulações naturais em tempo contínuo, o que não acontece com os modelos GARCH usuais, e estes componentes adicionais de volatilidade estocástica podem ser formulados como fatores latentes adicionais na formulação de espaço de estados, e tratados pela mesma metodologia de estimação Bayesiana utilizando Markov Chain Monte Carlo.

$$y_t = Z_t(\boldsymbol{\alpha}_t) + \boldsymbol{\varepsilon}_t^+, \quad \boldsymbol{\varepsilon}_t^+ \sim NID(0, \Sigma_\varepsilon^+), \quad t = 1, \dots, T. \quad (4.8)$$

com $\boldsymbol{\alpha}_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t}, \beta_{4,t}, \boldsymbol{\varepsilon}_t^*)' = (\beta_t', \boldsymbol{\varepsilon}_t^*)'$ e onde $Z_t(\boldsymbol{\alpha}_t)$ é um vetor ($N \times 1$) de funções

$$Z_t(\boldsymbol{\alpha}_t) = \Lambda(\lambda_t(\beta_{1,t}, \beta_{2,t}, \beta_{3,t})' + \Gamma_\varepsilon \boldsymbol{\varepsilon}_t^*), \quad t = 1, \dots, T. \quad (4.9)$$

com $\lambda_t = \beta_{4,t}$.

A dinâmica dos fatores latentes em β_t é modelada por um processo autoregressivo de primeira ordem, VAR(1). A equação de transição dos estados é definida por:

$$\boldsymbol{\alpha}_{t+1} = c + \begin{bmatrix} \Phi & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{\alpha}_t + \begin{pmatrix} \eta_t \\ \boldsymbol{\varepsilon}_{t+1}^* \end{pmatrix}, \quad (4.10)$$

$$\begin{pmatrix} \eta_t \\ \varepsilon_{t+1}^* \end{pmatrix} \sim NID \left(0, \begin{bmatrix} \Sigma_\eta & 0 \\ 0 & h_{t+1} \end{bmatrix} \right),$$

para $t = 1, \dots, T$ e $c = [\mu'(I - \Phi)', 0]'$.

A segunda formulação analisada neste trabalho consiste em tratar a heterocedasticidade condicional diretamente nos fatores latentes da curva de juros. Nesta especificação, é modelada a volatilidade estocástica diretamente nos fatores de nível, inclinação e curvatura, permitindo capturar a incerteza ao longo do tempo inerente a cada um dos fatores individualmente. A equação de medida mantém a mesma do caso anterior, $y_t(\tau) = \Lambda(\lambda_t)\beta_t + \varepsilon_t$, com o parâmetro λ_t sendo tratado como variante no tempo, logo a matriz de pesos dos fatores $\Lambda(\lambda_t)$ também é variante no tempo. O vetor de erros ($N \times 1$) definido como:

$$\varepsilon_t := (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt}) \sim^{i.i.d} N(0, \Sigma) \quad (4.11)$$

com

$$\Sigma = \text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_N^2\}. \quad (4.12)$$

Neste caso, tanto os fatores (estados não observáveis) como as volatilidades estocásticas associadas a eles são tratados como fatores latentes, sendo que a dinâmica dos fatores latentes é modelada por um vetor autoregressivo de primeira ordem, e para a heterocedasticidade condicional a formulação utilizada é a de um modelo autoregressivo para o componente não observado de volatilidade estocástica de cada fator, ARSV(1).

Como antes, assume-se que o vetor η_t (3×1), da equação de transição dos estados (equação 4.3) é independente de ε_t , e

$$\eta_t \sim^{i.i.d} N(0, H_t). \quad (4.13)$$

A matriz de covariância H_t é especificada em termos de um processo de volatilidade estocástica com a seguinte forma:

$$\text{vec}(\ln H_t) = \mu_h + \Phi_h \text{vec}(\ln H_{t-1}) + \xi_t \quad (4.14)$$

onde μ é um vetor de parâmetros (6×1) e Φ_h é uma matriz de parâmetros (6×6). Assume-se que o termo de erro ξ_t é independente de η_t e ε_t , e é normalmente distribuído com matriz de covariância Σ_h capturando as covariâncias das covariâncias,

$$\xi_t \sim^{i.i.d} N(0, \Sigma_h). \quad (4.15)$$

Para diminuir o esforço computacional, uma alternativa é restringir as matrizes H_t e Φ_h para uma especificação diagonal. Assim, a estrutura dos fatores latentes pode ser expressa por:

$$\begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} = \begin{pmatrix} \mu_L \\ \mu_S \\ \mu_C \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t-1} \\ \beta_{2,t-1} \\ \beta_{3,t-1} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix}, \quad (4.16)$$

onde $\eta_t \sim^{i.i.d} N(0, H_t)$, com

$$\text{diag}(\ln H_t) = \begin{pmatrix} \ln(h_t^{\beta_1}) \\ \ln(h_t^{\beta_2}) \\ \ln(h_t^{\beta_3}) \end{pmatrix} = \begin{pmatrix} \mu_h^{\beta_1} \\ \mu_h^{\beta_2} \\ \mu_h^{\beta_3} \end{pmatrix} + \begin{pmatrix} \phi_h^{\beta_1} & 0 & 0 \\ 0 & \phi_h^{\beta_2} & 0 \\ 0 & 0 & \phi_h^{\beta_3} \end{pmatrix} \begin{pmatrix} \ln(h_{t-1}^{\beta_1}) \\ \ln(h_{t-1}^{\beta_2}) \\ \ln(h_{t-1}^{\beta_3}) \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix}. \quad (4.17)$$

em que $h_t^{\beta_1}$, $h_t^{\beta_2}$ e $h_t^{\beta_3}$ são os fatores de volatilidade, que capturam a incerteza ao longo do tempo relativa ao nível, inclinação e curvatura da curva de juros. O componente de nível da volatilidade, $h_t^{\beta_1}$, é a variância (variante no tempo) comum para todas as taxas de juros. Já o componente $h_t^{\beta_2}$ está associado à inclinação da curva de juros, refletindo a volatilidade do prêmio de risco das taxas de juros para as maturidades mais longas em relação às curtas. O fator $h_t^{\beta_3}$ está associado à incerteza relativa às taxas de juros para maturidades intermediárias, ou à curvatura da curva de juros.

4.4 Estimação Bayesiana baseado em MCMC

As especificações do modelo pelas equações (4.2, 4.3 e 4.7) ou (4.2, 4.3 e 4.17) constitui um sistema em formato de espaços de estados não linear, em que a equação 4.2 é a equação de medida, expressando as taxas de juros observadas $y_t(\tau)$ como função dos fatores latentes, $\beta_{t,j}$, do parâmetro de peso dos fatores λ_t e das maturidades τ . A dinâmica dos fatores latentes é determinada pela equação 4.3, que descreve um processo regressivo de primeira ordem, com matriz de parâmetros dada por Φ .

No primeiro caso, é considerado um componente agregado de volatilidade estocástica para os erros da equação de medida, descrito pela equação 4.7, que segue a especificação proposta por Koopman et al. (2010), porém, neste caso a formulação empregada é a de um ARSV(1) para o componente não observado do log volatilidade estocástica. No segundo caso, em que é modelada diretamente a variância dos fatores, a matriz de covariância é especificada em termos de um processo de volatilidade estocástica dado pela equação 4.17.

Note que nas duas especificações o parâmetro que determina o peso dos fatores λ_t é tratado como variante no tempo e é inserido vetor β_t . Assim, o modelo não pode ser estimado por métodos de estimação como filtro de Kalman, que se aplicam apenas a modelos que são lineares no vetor de estados. Uma alternativa é linearizar localmente a equação de transição dos estados e empregar o algoritmo do filtro de Kalman estendido (para detalhes ver Anderson

and Moore (1979b)). Como é de particular interesse manter a especificação, a forma de re-estimar a estimação simultânea do modelo é através de métodos de inferência bayesiana usando Markov Chain Monte Carlo. Para justificativas adicionais do uso de métodos bayesianos em modelos de estrutura a termo pode ser consultado o trabalho de Chib et al. (2009), que aponta as vantagens destes métodos em relação aos problemas de dimensionalidade, identificação e inferência em modelos multifatores para a estrutura a termo.

Ao empregar métodos de inferência bayesiana, o objetivo é encontrar a distribuição posterior dos parâmetros de interesse condicionados a amostra observada, denotada por $p(\Theta|y)$. Para encontrar a distribuição dos parâmetros condicionados a amostra, usa-se a relação:

$$p(\Theta|y) = p(\Theta, y) / p(\Theta) = p(y|\Theta)p(\Theta) / p(y) \quad (4.18)$$

onde $p(y|\Theta)$ é a verossimilhança do modelo, $p(\Theta)$ denota a distribuição a priori assumida para o parâmetro e $p(y)$ é a distribuição marginal da amostra, que precisa ser conhecida até uma constante de integração. Através do Lema de Bayes podemos escrever:

$$p(\Theta|y) = p(\Theta, y) / p(\Theta) = p(y|\Theta)p(\Theta) / c \quad (4.19)$$

assim temos que a posteriori é proporcional ao produto da verossimilhança pela distribuição a priori:

$$p(\Theta|y) \propto p(y|\Theta)p(\Theta). \quad (4.20)$$

Quando o modelo a ser estimado pode ser colocado em formato de espaços de estados, uma forma conveniente de tratar o problema é através da metodologia de estimadores de Bayes Hierárquicos. Esta formulação é especialmente útil em modelos no formato espaços de estados já que a especificação hierárquica permite estimar a distribuição dos hiperparâmetros relacionados aos fatores latentes utilizando os próprios dados, especificando uma dinâmica de evolução para os fatores latentes.

O modelo diagonal, conforme especificado acima, corresponde a um modelo hierárquico latente em três níveis com seis processos latentes. Seja Θ a coleção de parâmetros do modelo, e seja $F_t := (\beta_{1t}, \beta_{2t}, \beta_{3t})$ e $V_t := (h_t^{\beta_1}, h_t^{\beta_2}, h_t^{\beta_3})$. Então, a função de verossimilhança do modelo é dada por:

$$p(\Theta|y) = \int_{F_1} \int_{F_2} \dots \int_{F_T} p(Y|\Theta, F_1, F_2, \dots, F_T) p(F_1, F_2, \dots, F_T|\Theta) dF_1 dF_2 \dots dF_T, \quad (4.21)$$

onde $p(Y|\Theta, F_1, F_2, \dots, F_T)$ denota a densidade (condicional) das observações Y dados os parâmetros Θ , os fatores latentes e a estrutura imposta como dado pelas equações (4.2 e 4.12). Além disso, $p(F_1, F_2, \dots, F_T|\Theta)$ denota a densidade conjunta (condicional) dos fatores latentes, dados os parâmetros do modelo Θ e é determinada pela equação 4.3. Como os fa-

tores são não observáveis, precisam ser integrados, resultando em uma integral de dimensão $(3.T)$. Claramente, $p(F_1, F_2, \dots, F_T | \Theta)$ depende de um novo conjunto de componentes desconhecidos representados pelos fatores de volatilidade V_1, \dots, V_T . Sendo computada como:

$$p(F_1, F_2, \dots, F_T | \Theta) = \int_{V_1} \int_{V_2} \dots \int_{V_T} p(F_1, F_2, \dots, F_T | \Theta, V_1, V_2, \dots, V_T) \times p(V_1, \dots, V_T | \Theta) dV_1 dV_2 \dots dV_T, \quad (4.22)$$

onde $(V_1, \dots, V_T | \Theta)$ denota a densidade conjunta dos componentes da volatilidade como definido em 4.14. A função de verossimilhança acima não pode ser calculada analiticamente em forma fechada e requer técnicas de aproximação numérica. Assim, o modelo pode ser estimado usando inferência baseada em Markov Chain Monte Carlo (MCMC), que aproximam esta integral como o resultado da simulação de cadeias de Markov. Consequentemente, considere $\Omega := \{\Theta, F_1, \dots, F_T, V_1, \dots, V_T\}$ como um vetor aleatório, cuja distribuição posterior $p(\Omega | Y)$ pode ser arranjada como:

$$\begin{aligned} p(\Omega | Y) &= p(F_1, F_2, \dots, F_T, V_1, V_2, \dots, V_T, \Theta | Y) \\ &\propto p(Y | F_1, F_2, \dots, F_T, V_1, V_2, \dots, V_T, \Theta) \\ &\times p(F_1, F_2, \dots, F_T | V_1, V_2, \dots, V_T, \Theta) \\ &\times p(V_1, V_2, \dots, V_T, \Theta) \\ &\times p(\Theta) \end{aligned}$$

A estrutura de prioris utilizada é dada por uma distribuição normal multivariada com matriz de precisão dada por uma distribuição Wishart-inversa para os parâmetros do vetor autoregressivo de fatores latentes; os parâmetros de decaimento seguem distribuições log-normais na formulação hierárquica com parâmetros de precisão dados por uma distribuição gamma inversa, os fatores latentes de nível, inclinação e curvatura prioris gaussianas hierárquicas, também utilizando gammas inversas como parâmetros de precisão. Os parâmetros de intercepto e persistência das equações de volatilidade são especificados pelo par normal-gamma inversa e no caso do modelo de fatores na volatilidade dada pela equação 4.5 a matriz de cargas também é dada por um par normal-gamma inversa. A estrutura geral para os fatores latentes da volatilidade é dado por priores log-normais hierárquicas com precisões gamma inversas. A estrutura do modelo é finalizada assumindo-se uma distribuição normal multivariada para os yields da curva de juros.

O modelo segue basicamente a mesma estrutura de priores e hiperparâmetros adotada em Hautsch and Ou (2009), com excessão dos parâmetros de decaimento e volatilidades variantes no tempo como detalhado. É importante notar que o ajuste do modelo é basicamente invariante as escolhas destes hiperparâmetros, o que pode ser compreendido pelo número elevados de observações contidos na estrutura a termo de taxas de juros no tempo. Os valores dos hiperparâmetros utilizados nas estimações podem ser obtidos com os autores, e são omitidos por questões de espaço.

Com a definição da estrutura das distribuições a priori e da verossimilhança, é possível recuperar a distribuição posterior dos parâmetros e fatores latentes. Devido ao uso de estruturas não-lineares e da uso de estruturas de prioris nãoconjugadas, não é possível obter as posteriores de forma analítica, e assim é necessário o uso de métodos numéricos. Nesse caso utilizamos algoritmos baseados no uso de métodos de Markov Chain Monte Carlo para obter amostras das distribuições posteriores, utilizando o algoritmo Gibbs na amostragem das distribuições condicionais conhecidas e o algoritmo de Metropolis-Hastings nas simulações para os casos onde as distribuições condicionais não sejam conhecidas de forma analítica. Na estrutura de prioris assumida, o algoritmo de Metropolis-Hastings é necessário para avaliar as distribuições posteriores dos parâmetros e fatores latentes associados a parâmetro de decaimento e a volatilidade estocástica, que representam os componentes não-lineares do modelo. Nos demais casos as distribuições condicionais são conhecidas e o algoritmo de Gibbs pode ser utilizado. Para mais detalhes veja por exemplo Haustsch and Ou (2009). No procedimento de Markov Chain Monte Carlo inicializamos a estimação com 20000 iterações de burn-in, e na construção das posteriores mais 20000 iterações. Para evitar problemas de não-convergência das cadeias, utilizamos somente cada décima iteração, descartando as demais iterações neste intervalo. A convergência é assegurada através do diagnóstico de Gelman-Rubin. Na sumarização dos resultados utilizamos a média a posteriori, e para a construção dos intervalos de credibilidade os percentis de 2.5% e 97.5% da distribuição posterior obtida pelo procedimento de MCMC.

4.5 Dados, Estimação e Resultados

Nesta seção, inicialmente nós descrevemos a base de dados relativa as taxas de juros a serem usadas e o método de estimação. A seguir examinamos estimações e ajuste dentro da amostra quatro especificações do modelo.

4.6 Dados

O contrato futuro de depósito interbancário (DI Futuro) com maturidade τ é um contrato futuro do qual o ativo básico é taxa de juros acumulada diariamente (DI), capitalizada entre o momento da negociação t e τ . O valor do contrato é dado pelo seu valor no vencimento, R\$ 100.000,00 descontado pela taxa de juros acumulada, negociada entre o vendedor e o comprador do contrato. Ao comprar um contrato de DI Futuro ao preço no tempo t e mantê-lo até a maturidade τ , o ganho ou perda é dado por:

$$100.000 \left(\frac{\prod_{i=1}^{\zeta(t,\tau)} (1 + y_i)^{\frac{1}{252}}}{(1 + DI^*)^{\frac{\zeta(t,\tau)}{252}}} - 1 \right),$$

Onde y_i denota a taxa DI ($i - 1$) dias após o dia da negociação. A função ζ representa o número de dias entre t e τ .

O contrato de DI é muito similar a um título zero-cupom, exceto pelo fato de pagar os ajustes de margem diariamente. A cada dia o fluxo de caixa é a diferença entre o preço de ajuste (liquidação) do dia corrente e o preço de ajuste do dia anterior, corrigido pela taxa DI do dia anterior.

Os contratos de DI Futuro são negociados na BM&F, que determina o número de vencimentos com contratos autorizados. Geralmente, para cada dia, existem em torno de 20 vencimentos com contratos autorizados, mas nem todos apresentam liquidez. Normalmente por volta de 10 vencimentos apresentam contratos com maior liquidez. Existem contratos com vencimentos mensais para os meses que iniciam cada trimestre, janeiro, abril, julho e outubro. Além disso, existem contratos com vencimentos para os quatro meses subsequentes ao mês corrente. A data de vencimento é o primeiro dia útil do mês de vencimento do contrato.

Os dados usados neste trabalho consistem de observações diárias das taxas dos contratos de DI-Futuro, preços de fechamento. Na prática, não observados diariamente contratos com todas as maturidades. Portanto, a partir das taxas observadas diariamente os dados foram convertidos em maturidades fixas de 1, 2, 3, 4, 6, 9, 12, 15, 18, 24, 27, 29, 31 e 33 meses, por meio de interpolações usando cubic splines. A base de dados é composta pelas maturidades que apresentaram maior liquidez para o período de janeiro de 2006 a fevereiro de 2009, perfazendo um total de 756 observações. A figura 4.1 mostra a evolução desta curva de juros no tempo. Trabalhar com maturidades fixas facilita a análise dos resultados e empregar dados não interpolados é mais relevante quando o objetivo é verificar condições de não-abrigo, como em Laurini and Westin (2010), por exemplo, que usando uma base de dados similar para o mesmo período concluem para a não ocorrência de oportunidades de arbitragem. Consequência, possivelmente, da elevada liquidez dos contratos de DI-Futuro no período para as maturidades empregadas. Na tabela 4.1, são apresentadas algumas estatísticas descritivas da curva de juros. Para cada maturidade são reportadas média, desvio-padrão, mínimo, máximo e alguns coeficientes de autocorrelação. Nota-se que as taxas de juros para maturidades mais curtas são mais voláteis e apresentam maior persistência, diferindo um pouco do que geralmente se observa em curvas de juros de economias com mercados mais maduros. Pode-se observar também, mesmo que levemente, a curva de juros apresenta inclinação positiva em média, característica tipicamente observada.

4.7 Análise dos Resultados

Para comparar as estimativas do modelo de Nelson-Siegel nas especificações com componente comum de volatilidade, e com fatores específicos de volatilidade nós também estimamos outras versões do modelo.

Inicialmente estimamos o modelo através do método de dois passos, com parâmetro λ

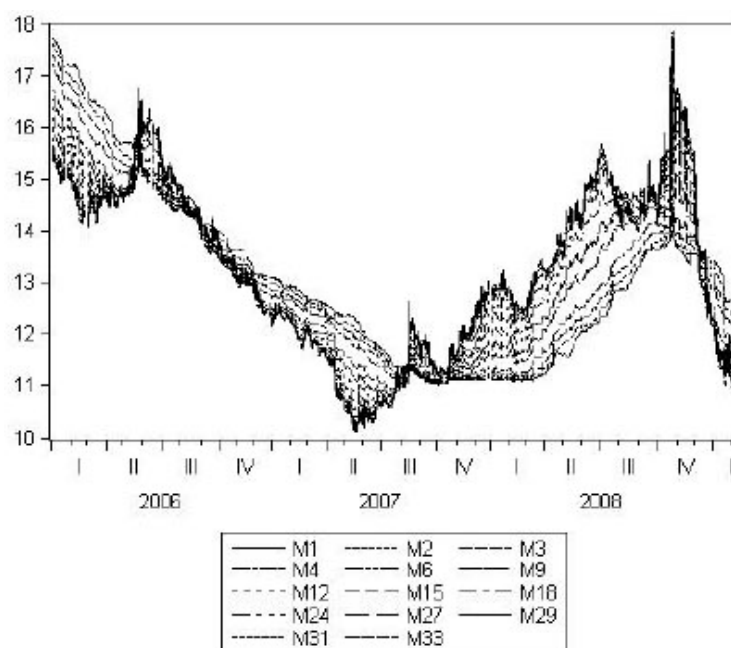


Figura15: *Base de Dados de Futuros de DI.*

Fonte: Elaboração própria.

Tabela13: Estatísticas descritivas da curva de juros (Jan-06 a Fev-09).

Maturidade (Meses)	Média	Desv. pad	Mínimo	Máximo	$\hat{\rho}(21)$	$\hat{\rho}(63)$
1	13.125	1.731	11.05	17.71	0.894	0.672
2	13.082	1.677	11.04	17.60	0.890	0.662
3	13.05	1.620	11.02	17.40	0.885	0.650
4	13.027	1.566	11.02	17.21	0.879	0.635
6	13.028	1.523	10.90	16.72	0.866	0.597
9	13.058	1.534	10.78	16.54	0.858	0.575
12	13.121	1.567	10.68	16.47	0.851	0.560
15	13.193	1.600	10.60	16.91	0.849	0.553
18	13.248	1.614	10.46	17.15	0.845	0.550
24	13.287	1.615	10.37	17.36	0.840	0.543
27	13.311	1.625	10.28	17.57	0.836	0.536
29	13.321	1.630	10.21	17.72	0.831	0.527
31	13.325	1.632	10.15	17.83	0.828	0.521
33	13.326	1.633	10.13	17.87	0.827	0.517

Fonte: Elaboração própria.

Nota: As duas últimas colunas apresentam autocorrelações para defasagens de 1 e 3 meses.

fixo e sem levar em conta heterocedasticidade condicional, como proposto originalmente por Diebold and Li (2006). O parâmetro λ , que determina a forma da curva de juros, foi escolhido

tal que minimize o RMSE para ajuste dentro da amostra ($\lambda = 0.1046$). Na tabela 4.2 são apresentados os parâmetros do VAR estimado para a dinâmica dos fatores. A dinâmica dos fatores latentes β_1 , β_2 e β_3 exibe elevada persistência, com coeficientes autoregressivos de 0.98, 0.99 e 0.99, respectivamente.

Tabela14: Parâmetros do VAR para Modelo Estimado pelo Método de Dois Passos

	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$	μ
$\beta_{1,t}$	0.984*** (0.012)	-0.002* (0.012)	0.003 (0.003)	13.189*** (0.1504)
$\beta_{2,t}$	0.015*** (0.012)	1.000*** (0.012)	0.002 (0.004)	-.211 (0.152)
$\beta_{3,t}$	-.033 (0.016)	-.045*** (0.016)	0.993*** (0.005)	0.384 (0.205)

Fonte: Elaboração própria.

Nota: Erros padrões entre parênteses. Parâmetros obtidos da estimação do vetor autoregressivo para os fatores latentes. *, **, e *** indica significância estatística aos níveis de 10%, 5% e 1% respectivamente.

Na figura 4.2 plotamos os fatores latentes estimados. Parece claro que todos fatores exibem elevada persistência. A figura 4.3 traz a variância do erro de medida associada a esta especificação

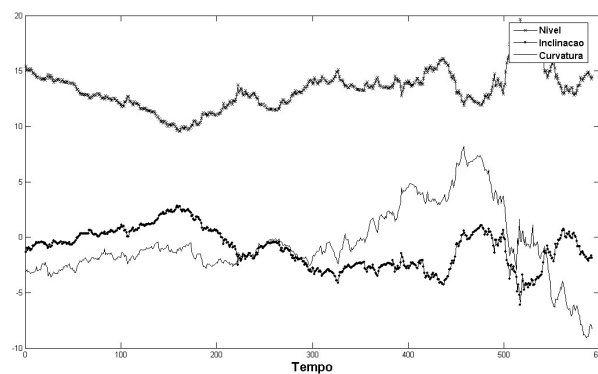


Figura16: Fatores Latentes do Modelo Estimado pelo Método de Dois Passos.

Fonte: Elaboração própria.

β_1 , β_2 e β_3 obtidos por mínimos quadrados para cada período t .

Na tabela 4.3 a seguir apresentamos estatísticas para ajuste dentro da amostra para o modelo estimado pelo método de dois passos. Para critérios de comparação da qualidade do ajuste do modelo para as especificações alternativas propostas, calculamos ME (*Mean Error*), RMSE

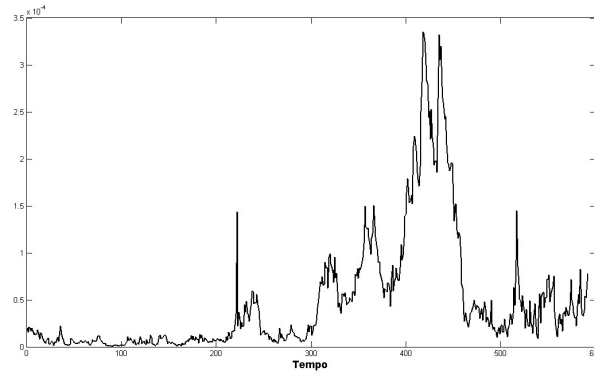


Figura17: Variância da Equação de Medida.

Fonte: Elaboração própria.

Variância dos erros de medida para modelo estimado pelo método de dois passos.

(*Root Mean Squared Error*), MAE (*Mean Absolute Error*), MPE (*Mean Percentage Error*), MAPE (*Mean Absolute Percentage Error*) e autocorrelação para uma defasagem. Nota-se que o modelo captura parte substancial da dinâmica da curva de juros, especialmente para as maturidades intermediárias, 12 e 15 meses, por exemplo. Entretanto, a elevada autorrelação nos resíduos indicam que informações importantes na média e variância estão sendo negligenciadas.

Tabela15: Medidas de Ajuste da curva de Juros *in-sample* (Jan-06 a Fev-09).

Maturidade (Meses)	ME	RMSE	MAE	MPE	MAPE	ACF1
1	0.0028	0.005	0.0039	2.323	3.143	0.9776
2	0.0023	0.0051	0.004	-1.9552	3.1894	0.9903
3	0.0036	0.0068	0.0054	-2.9965	4.2797	0.9907
4	0.0042	0.0069	0.0055	-3.35	4.2983	0.9894
6	0.004	0.0054	0.0045	-3.1044	3.4733	0.9848
9	0.0033	0.0041	0.0035	-2.5686	2.7343	0.9781
12	0.0021	0.0026	0.0023	-1.6303	1.7834	0.9673
15	0.0007	0.0018	0.0015	-0.5964	1.1796	0.975
18	0.0006	0.0021	0.0017	0.3694	1.2656	0.982
24	0.0019	0.0032	0.0025	1.3454	1.8675	0.9821
27	0.0029	0.0041	0.0034	2.0897	2.5009	0.9829
29	0.0036	0.0049	0.0041	2.6405	3.0315	0.9827
31	0.0041	0.0054	0.0046	3.0056	3.3934	0.9823
33	0.0044	0.0057	0.0049	3.219	3.5979	0.9818
TOTAL	0.0029	0.0048	0.0037	-0.0863	2.8384	0.992

Fonte: Elaboração própria.

Nota: Modelo estimado pelo método de dois passos com fixo, $\lambda = .1046$, como em Diebold e Li.

Nota-se que, excluindo-se as maturidades mais curtas, o modelo ajusta melhor para todas

as demais. Há destaque para as maturidades mais longas, acima de 18 meses, em que a melhora no ajuste é ainda mais significativa. Assim, parece ficar claro que tratar o parâmetro λ_t como fixo, apesar de facilitar os procedimentos de estimação do modelo, implica em perda na qualidade do ajuste. Esse resultado indica também que a maturidade para a qual a curvatura assume peso máximo não é constante ao longo do tempo. Assim, manter λ fixo ao longo de todo o período amostral pode ser muito restritivo, uma vez que os dados abrangem um longo período de tempo. Em particular, a maturidade na qual o fator curvatura, β_{3t} é maximizada e a velocidade de decaimento do fator de inclinação β_{2t} dependem apenas do parâmetro λ , e são mantidas fixas por consequência. Entretanto, essas características da curva de juros podem mudar ao longo do tempo.

A primeira especificação alternativa proposta para o modelo de Diebold and Li (2006) consiste em incluir um fator comum de volatilidade estocástica nos erros de medida, conforme definido pelas equações (4.2, 4.3 e 4.7), discutidas na seção 3. Neste caso, o fator de decaimento λ_t é tratado como um quarto fator latente e é estimado simultaneamente com os demais fatores. A figura 4.4 apresenta o fator comum de volatilidade estocástica estimada. Pode-se observar que a volatilidade comum é especialmente alta no segundo semestre de 2008, período que coincide com a fase mais aguda da crise financeira internacional. O componente de volatilidade estocástica estimado mostra a capacidade do modelo em capturar a heterocedasticidade condicional existente nas taxas de juros, identificando os momentos de mudanças, como inversões de formato, por exemplo, ocorridas no final de 2006 e no último trimestre de 2008, períodos caracterizados por mudanças nos rumos da política monetária.

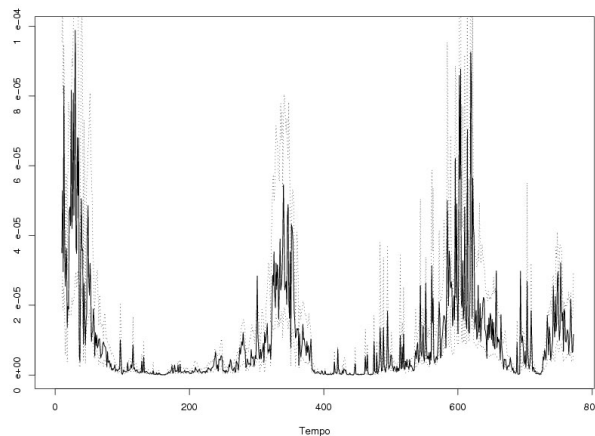


Figura18: *Volatilidade Variante no Tempo.*

Fonte: Elaboração própria.

Nota: Volatilidade do modelo com componente comum de volatilidade estocástica e pesos dos fatores variando no tempo.

A tabela 4.4 traz estatísticas dos erros de medida para o modelo com componente comum de volatilidade estocástica e λ_t variando no tempo. Quando comparado com a especificação anterior, que não considerava heterocedasticidade condicional nos erros de medida. A mel-

hora no ajuste para as maturidades mais longas reflete a presença de volatilidade condicional, principalmente nas taxas de juros para prazos mais longos.

Tabela16: Medidas de Ajuste da curva de Juros *in-sample* (Jan-06 a Fev-09).

Maturidade (Meses)	ME	RMSE	MAE	MPE	MAPE	ACF1
1	0.0019	0.0067	0.0047	1.3454	3.5182	0.2459
2	0.0018	0.0061	0.0045	1.1712	3.3533	0.7794
3	0.0014	0.0055	0.0041	0.9088	3.0897	0.8117
4	0.0009	0.0045	0.0032	0.6578	2.4134	0.771
6	0.0004	0.0026	0.0016	0.322	1.2154	0.5771
9	0.0003	0.0019	0.0014	0.2186	1.0575	0.5091
12	0.0004	0.0019	0.0014	0.2537	1.0829	0.7425
15	0.0006	0.002	0.0015	0.3479	1.0947	0.9269
18	0.0005	0.0019	0.0015	0.2985	1.1135	0.981
24	0.0002	0.0017	0.0013	0.0779	0.961	0.8621
27	0.0002	0.0017	0.0011	-0.1896	0.8147	0.6102
29	0.0006	0.002	0.0014	-0.4801	1.0768	0.4604
31	0.0009	0.0024	0.0018	-0.7042	1.3644	0.4285
33	0.0011	0.0027	0.0021	-0.8571	1.5551	0.424
TOTAL	0.0004	0.0036	0.0022	0.2408	1.6936	0.6345

Fonte: Elaboração própria.

Nota: Especificação com componente comum de volatilidade estocástica e λ_t variando no tempo.

A seguir apresentamos resultados da estimação do modelo considerando heterocedasticidade condicional nos fatores latentes de Nelson e Siegel e nos erros de medida. Nesta especificação, tanto a volatilidade dos fatores como o parâmetro de decaimento λ_t são tratados como variantes no tempo. A figura 4.5 apresenta as estimativas dos fatores, $(\beta_1, \beta_2, \beta_3)$ e λ_t , que é tratado como um quarto fator latente. Neste caso foi considerada uma especificação diagonal para a equação de transição que modela a dinâmica dos fatores, sendo a dinâmica do λ_t modelada por um AR(1).

Ao incluir um fator de volatilidade específico para cada fator latente, a persistência exibida pelos fatores latentes reduz significativamente, conforme pode ser observado na tabela 4.5, que mostra intervalos de confiança para os parâmetros estimados do modelo com componentes de volatilidade estocástica, especificado pelas equações (4.2, 4.3 e 4.17). Já o parâmetro λ_t apresenta persistência considerável, com coeficiente autoregressivo médio de 0.64.

As estimativas dos parâmetros relativos à dinâmica dos fatores de volatilidade estocástica são apresentados nas duas últimas colunas da tabela 4.5. Nós encontramos significante evidências de forte autocorrelações nos fatores de volatilidade, com coeficientes autoregressivos para dinâmica dos componentes de volatilidade para o nível, inclinação e curvatura em média

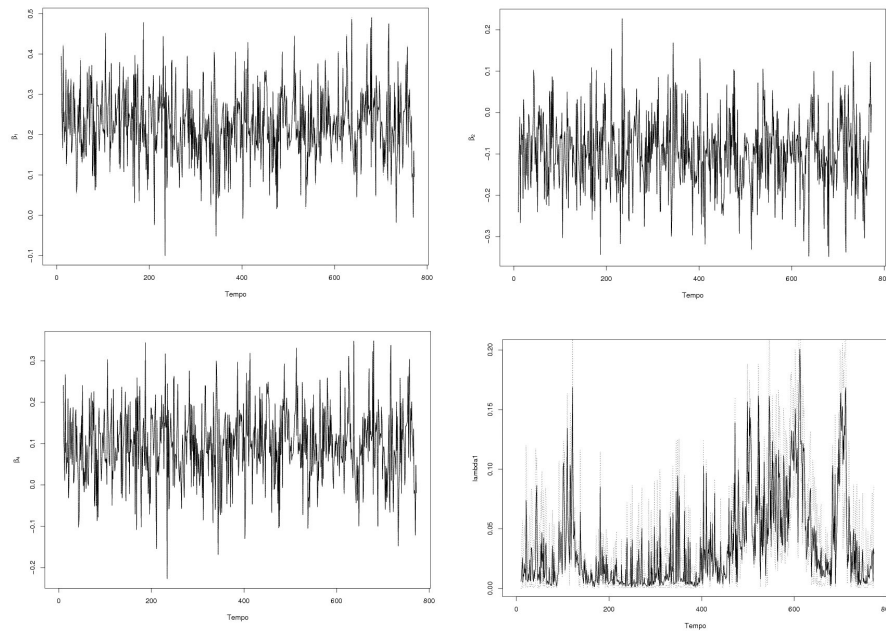


Figura 19: Estimativas dos Fatores Latentes e Fator de Decaimento - $(\beta_1, \beta_2, \beta_3, \lambda_t)$.

Fonte: Elaboração própria.

Estimativas dos fatores obtidas com especificação diagonal para matriz da equação de transição e AR(1) para o fator de peso.

Tabela.17: Intervalos de Confiança dos Parâmetros a 95% - Φ e Φ_h

	Fatores Latentes		Fatores de Volatilidade	
	μ_β	ϕ_β	μ_h^β	ϕ_h^β
$\beta_{1(.025)}$	0.1766975	0.1096750	-14.19452	0.9628975
$\beta_{1(.50)}$	0.1858774	0.15298467	-13.40192	0.9774741
$\beta_{1(.975)}$	0.1956050	0.1944075	-13.19515	0.999200
$\beta_{2(.025)}$	-0.0876405	0.07427575	-12.37183	0.9865950
$\beta_{2(.50)}$	-0.0811913	0.14051617	-10.64274	0.9952372
$\beta_{2(.975)}$	-0.0744380	0.2024050	-10.56680	0.999800
$\beta_{3(.025)}$	0.07477875	0.07177125	-11.33084	0.9845000
$\beta_{3(.50)}$	0.08151998	0.13685784	-10.67576	0.9943536
$\beta_{3(.975)}$	0.08817710	0.2026050	-10.47721	0.9998000
$\lambda_{t(.025)}$	0.0057788	0.5646200		
$\lambda_{t(.50)}$	0.0084367	0.6373952		
$\lambda_{t(.975)}$	0.0111503	0.7005025		

Fonte: Elaboração própria.

0.977, 0.995 e 0.994, respectivamente. Diferentemente dos fatores latentes, os processos de volatilidade estocástica apresentam elevada persistência. A evolução dos fatores de volatilidade é apresentada na figura 4.6. O fator nível apresenta padrão mais elevado de variações na volatilidade.

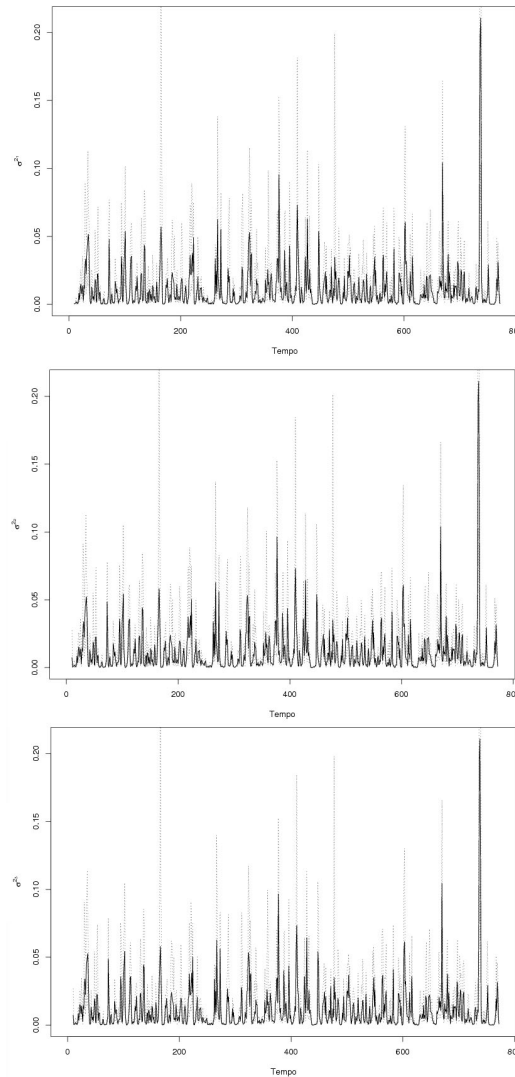


Figura 20: *Estimativas dos Fatores de Volatilidade - $(\sigma_{\beta_1}, \sigma_{\beta_2}, \sigma_{\beta_3})$.*

Fonte: Elaboração própria.

Estimativas dos fatores dos fatores de volatilidade estocástica obtidas com especificação diagonal para Φ_h^β (equação 4.17).

Intervalos de confiança para os parâmetros do processo de volatilidade estocástica dos erros de medida são apresentados na tabela 4.6.

Na tabela 4.7 apresentamos estatísticas referentes ao ajuste do modelo estimado com a especificação mais geral proposta aqui, que trata os pesos dos fatores como variantes no

Tabela.19: Intervalos de Confiança dos Parâmetros a 95% - σ_t

	$IC_{(0.025)}$	$IC_{(0.50)}$	$IC_{(0.975)}$
μ_σ	-19.71950	-18.34460	-17.33550
ϕ_σ	0.9923000	0.9967133	0.997500

Fonte: Elaboração própria.

Intervalos de confiança para os parâmetros da volatilidade estocástica dos erros de medida.

tempo, volatilidade estocástica nos erros de medida e componente de volatilidade específico para cada fator latente. Os resultados mostram que a extensão do modelo melhora significativamente a capacidade de ajuste, principalmente para as maturidades mais curtas, onde supera todas as especificações mais básicas. Quando comparado com a especificação com componente comum de volatilidade estocástica, esta versão do modelo apresenta melhor desempenho para algumas maturidades, principalmente as mais curtas, e é superada em outras.

Tabela.20: Medidas de Ajuste da curva de Juros *in-sample* (Jan-06 a Fev-09).

Maturidade (Meses)	ME	RMSE	MAE	MPE	MAPE	ACF1
1	0.0015	0.006	0.0043	1.1093	3.2918	0.1916
2	0.0015	0.0056	0.0043	1.0681	3.2743	0.744
3	0.0012	0.0053	0.004	0.8417	3.1328	0.7804
4	0.0008	0.0045	0.0032	0.6052	2.5136	0.7372
6	0.0003	0.0027	0.0017	0.2841	1.3604	0.5638
9	0.0002	0.0021	0.0015	0.1882	1.1713	0.5268
12	0.0003	0.0019	0.0014	0.2301	1.1249	0.7345
15	0.0005	0.002	0.0014	0.328	1.0749	0.9217
18	0.0004	0.0019	0.0014	0.2811	1.1197	0.9803
24	0.0001	0.0017	0.0013	0.0619	1.0613	0.8484
27	0.0002	0.0019	0.0012	-0.2019	0.9408	0.6298
29	0.0006	0.0023	0.0015	-0.4878	1.1688	0.5076
31	0.0009	0.0027	0.0019	-0.7082	1.4701	0.4713
33	0.0011	0.0029	0.0021	-0.8576	1.6628	0.4579
TOTAL	0.0003	0.0034	0.0022	0.1959	1.7405	0.6135

Fonte: Elaboração própria.

Nota: Especificação com volatilidade estocástica nos fatores e λ_t variando no tempo.

Portanto, podemos concluir que as extensões do modelo contribuem significativamente para melhorar a capacidade de ajuste do modelo de Nelson e Siegel na formulação proposta por Diebold and Li (2006). A extensão que trata o parâmetro λ_t como variante no tempo e componentes de volatilidade estocástica para cada fator foi a que apresentou os ganhos mais significantes, encorajando seu uso por participantes do mercado tanto na modelagem de risco

e precificação de derivativos de renda fixa, ou como instrumento auxiliar para formuladores de políticas monetárias.

4.8 Conclusão

Neste artigo implementamos duas extensões para o modelo da estrutura a termo de Nelson e Siegel na formulação proposta por Diebold and Li (2006). Usando dados de contratos DI-futuro negociados na BM&F nós estimamos o modelo de Nelson e Siegel com pesos dos fatores e volatilidade variando no tempo. Na primeira especificação seguimos a estrutura proposta por Koopman et al. (2010), impondo um componente comum de volatilidade estocástica para os erros de medida que é modelado por um processo de volatilidade estocástica. No segundo caso nós ampliamos a formulação apresentada por Hautsch and Ou (2009) e propomos modelar diretamente a volatilidade estocástica dos fatores latentes de Nelson e Siegel, tratando os pesos dos fatores como variantes no tempo. Esta estrutura permite capturar a volatilidade da curva de juros de uma forma flexível e ainda parcimoniosa. O modelo é representado no formato de espaços de estados, com os fatores na equação de transição apresentando volatilidade estocástica, e os próprios componentes de volatilidade estocástica seguindo um processo latente dinâmico. Uma vez que o modelo é não linear, nós propomos o uso de inferência bayesiana, com base em MCMC, para estimar o modelo de Nelson e Siegel com volatilidade estocástica.

Os resultados de nossa estimação fornecem evidências para distintas volatilidades estocásticas variantes no tempo para os fatores da curva de juros. Em particular, o fator nível é o que mais contribui para a volatilidade da curva de juros. Analisando a qualidade do ajuste obtida com o modelo, nós mostramos que a inclusão de volatilidade estocástica ao modelo melhora seu poder explicativo.

Na especificação com fator comum de volatilidade, proposta por Koopman et al. (2010), este fator pode ser interpretado como a volatilidade de um título base de um portfólio, na forma de um modelo GARCH como em Engle et al. (1990) e Engle and Ng (1993). Tal especificação não permite diferentes volatilidades nas taxas de juros individuais. Já com a especificação proposta aqui, o fator de volatilidade associado ao nível da curva de juros $h_t^{\beta_1}$, é um componente de variância comum no sentido do fator de volatilidade conjunta em Koopman et al. (2010) e pode ser visto como uma proxy implícita para a volatilidade de um título do mercado como usado em Engle et al. (1990). O fator de volatilidade $h_t^{\beta_2}$ captura as variações ao longo do tempo no spread entre as taxas para maturidades longas e curtas, e é interpretado como a volatilidade da inclinação. Similarmente, $h_t^{\beta_3}$ é a volatilidade de um portfólio de títulos com maturidades intermediárias, e assim captura a incerteza associada com a curvatura da curva de juros.

Aplicando o modelo para dados dos contratos de DI-futuro negociados na BM&F nós achamos fortes evidências para volatilidade variante no tempo nas taxas de juros, com os fa-

tores de volatilidade apresentado elevada persistência. Ao comparar os resultados do ajuste da curva de juros para as especificações consideradas, fica claro que a inclusão da volatilidade estocástica no nível, inclinação e curvatura, ou via componente comum, melhora significativamente o ajuste quando comparado com as especificações básicas. A versão original do modelo proposta por Diebold and Li (2006), ajusta melhor para as maturidades mais curtas (1 e 2 meses). Já as especificações do modelo que incluem componente comum de volatilidade, ou que trata individualmente a volatilidade dos fatores latentes mostram-se superiores às demais, apresentando melhores critérios de ajuste para todas as maturidades analisadas. Sendo que para as maturidades mais curtas o modelo com componentes específicos de volatilidade apresenta melhor ajuste, com a situação se invertendo para as maturidades mais longas. Além disso, modelar a volatilidade estocástica específica de cada fator pode propiciar uma informação importante no cálculo de risco de carteiras de título públicos e derivativos. Alternativas a serem abordadas em trabalhos futuros pode ser a inclusão de fatores adicionais para curvatura e inclinação, como proposto por Laurini and Westin (2010), tornando o modelo mais flexível, e também podem ser consideradas outras especificações para estrutura da volatilidade, considerando uma estrutura multivariada ou mudanças de regime.

5 Conclusão

A proposta desta tese foi desenvolver estratégias de alocação de recursos baseadas em métodos quantitativos a serem empregadas num contexto de *hedge funds*. Além disso, propomos modelar e prever a estrutura a termo através do modelo de (Nelson and Siegel, 1987), considerando duas situações distintas. Uma em que modelo é colocado no formato espaços de estados e é estimado através do filtro de Kalman em uma única etapa. Já na segunda, propomos duas especificações para a estrutura de volatilidade estocástica do modelo, generalizando os trabalhos de Koopman et al. (2010) e Hautsch and Ou (2009).

Assim, a tese é composta de três ensaios. No primeiro deles empregamos a metodologia de cointegração no desenvolvimento de uma estratégia *long-short*, neutra ao mercado, que buscam obter retorno sem risco direcional ao mercado assumindo posição comprada em uma carteira de ações e posição vendida em outra simultaneamente. Empregamos os testes de cointegração de Johansen (1991) e Engle and Granger (1987) para identificar carteiras de ações que compartilham uma relação de equilíbrio de longo prazo com um benchmark para serem empregadas na estratégia. São empregados dados do mercado brasileiro para o período de janeiro de 2000 a dezembro de 2009, os resultados encontrados são animadores quanto ao uso do método de cointegração para otimização de carteiras no âmbito das estratégias analisadas.

A capacidade de replicar um benchmark em termos de retorno e volatilidade oferecida pela cointegração, com características de reversão à média do *tracking error*, estabilidade na composição das carteiras e melhor uso das informações contidas nos preços dos ativos, possibilita seu uso na implementação de uma variedade de estratégias, como *index tracking*, estratégias neutras ao mercado *long-short*, estratégias envolvendo pares de ações, *pair-trading*, que não foram analisadas neste artigo, entre outras.

Os resultados obtidos neste estudo são bastante animadores, tendo em vista a capacidade da estratégia *index tracking* baseada em cointegração, de replicar o benchmark em termos de retorno e volatilidade. Da mesma forma, a construção de estratégias *long-short* com carteiras cointegradas gerou resultados superiores ao benchmark e com baixo nível de volatilidade. Mesmo contendo algumas hipóteses simplificadoras, como o uso apenas das 50 ações mais líquidas do índice Bovespa, a não adoção de um critério de encerramento das posições diante de um dado nível de ganho ou perda, a metodologia de cointegração se mostrou eficaz, encorajando seu uso por parte de gestores de *hedge funds*, fundos de pensão e participantes do mercado em geral. Os resultados aqui encontrados, em termos de retorno, neutralidade e aderência ao índice de referência, se mostraram superiores aos obtidos em Alexander e Dimitriu (1999) e Dunis e Ho (2002), que também empregam a metodologia de cointegração na otimização de carteiras. Uma das possíveis razões é que aqui o processo de otimização das carteiras envolveu simulações com todas as combinações possíveis dentro do universo de

ações considerado, sendo a carteira ótima definida em função dos melhores critérios de integração apresentados. Vale destacar também que o parâmetro δ , que define o diferencial de retorno empregado na construção dos benchmarks para as carteiras empregadas na estratégia *long-short*, foi definido através de um processo de calibragem, em que foi escolhido o valor que apresentou os melhores resultados para os testes dentro da amostra.

No segundo ensaio, o modelo de Diebold and Li (2006), normalmente estimado para dados brasileiros através do método ineficiente de dois passos, foi colocado no formato espaço de estados e estimado eficientemente por máxima verossimilhança usando o filtro de Kalman. A estimação por máxima verossimilhança permite a estimação conjunta de todos os parâmetros do modelo, evitando a seleção a priori do parâmetro de decaimento. Estimativas suavizadas dos parâmetros, que consideram toda a informação da amostra para inferir sobre a série de tempo dos fatores, foram obtidas através do suavizador de Kalman e usadas para previsão fora da amostra. Os resultados mostram que o modelo estimado por máxima verossimilhança produz previsões fora da amostra de melhor qualidade que o modelo estimado em duas etapas para todos os diferentes horizontes de previsão. Além disso, as previsões baseadas no modelo estimado por máxima verossimilhança são superiores às do modelo de passeio aleatório para todas as maturidades quando são considerados horizontes de um mês, três meses e seis.

Os modelos de fatores da curva de juros são amplamente usados por bancos centrais de vários países e por muitos participantes do mercado para ajustar e prever a estrutura a termo da taxa de juros. Os resultados obtidos mostram a flexibilidade e capacidade do modelo para ajustar a uma ampla variedade de formatos da curva de juros, e que a estimação por filtro de Kalman é superior a estimação por dois passos. Uma possível sugestão para investigações futuras é a estimação do modelo usando quatro fatores, como proposto por Cochrane and Piazzesi (2005), o qual inclui uma curvatura adicional que melhora a previsibilidade em mercados com curvas de juros mais voláteis, como ocorre em mercados emergentes.

No terceiro ensaio implementamos duas extensões para o modelo da estrutura a termo de Nelson e Siegel na formulação proposta por Diebold and Li (2006). Usando dados de contratos DI-futuro negociados na BM&F nós estimamos o modelo de Nelson e Siegel com pesos dos fatores e volatilidade variando no tempo. Sendo que na primeira especificação seguimos a estrutura proposta por Koopman et al. (2010), impondo um componente comum de volatilidade estocástica para os erros de medida que é modelado por um processo de volatilidade estocástica. No segundo caso nós ampliamos a formulação apresentada por Hautsch and Ou (2009) e propomos modelar diretamente a volatilidade estocástica dos fatores latentes de Nelson e Siegel, tratando os pesos dos fatores como variantes no tempo. Esta estrutura permite capturar a volatilidade da curva de juros de uma forma flexível e ainda parcimoniosa. O modelo é representado no formato de espaços de estados, com os fatores na equação de transição apresentando volatilidade estocástica, e os próprios componentes de volatilidade estocástica seguindo um processo latente dinâmico. Uma vez que o modelo é não linear, nós propomos o

uso de inferência bayesiana, com base em MCMC, para estimar o modelo de Nelson e Siegel com volatilidade estocástica.

Os resultados de nossa estimação fornecem evidências para distintas volatilidades estocásticas variantes no tempo para os fatores da curva de juros. Em particular, o fator nível é o que mais contribui para a volatilidade da curva de juros. Analisando a qualidade do ajuste obtida com o modelo, nós mostramos que a inclusão de volatilidade estocástica ao modelo melhora seu poder explicativo.

Na especificação com fator comum de volatilidade, porposta por Koopman et al. (2010), este fator pode ser interpretado como a volatilidade de um título base de um portfólio, na forma de um modelo GARCH como em Engle et al. (1990) e Engle and Ng (1993). Tal especificação não permite diferentes volatilidades nas taxas de juros individuais. Já com a especificação proposta aqui, o fator de volatilidade associado ao nível da curva de juros $h_t^{\beta_1}$, é um componente de variância comum no sentido do fator de volatilidade conjunta em Koopman et al. (2010) e pode ser visto como uma proxy implícita para a volatilidade de um título do mercado como usado em Engle et al. (1990). O fator de volatilidade $h_t^{\beta_2}$ captura as variações ao longo do tempo no spread entre as taxas para maturidades longas e curtas, e é interpretado como a volatilidade da inclinação. Similarmente, $h_t^{\beta_3}$ é a volatilidade de um portfólio de títulos com maturidades intermediárias, e assim captura a incerteza associada com a curvatura da curva de juros.

Aplicando o modelo para dados dos contratos de DI-futuro negociados na BM&F nós achamos fortes evidências para volatilidade variante no tempo nas taxas de juros, com os fatores de volatilidade apresentado elevada persistência. Ao comparar os resultados do ajuste da curva de juros para as especificações consideradas, fica claro que a inclusão da volatilidade estocástica no nível, inclinação e curvatura, ou via componente comum, melhora significativamente o ajuste quando comparado com as especificações básicas. A versão original do modelo proposta por Diebold and Li (2006), ajusta melhora para as maturidades mais curtas (1 e 2 meses). Já as especificações do modelo que incluem componente comum de volatilidade, ou que trata individualmente a volatilidade dos fatores latentes mostram-se superiores às demais, apresentando melhores critérios de ajuste para todas as maturidades analisadas. Sendo que para as maturidades mais curtas o modelo com componentes específicos de volatilidade apresenta melhor ajuste, com a situação se invertendo para as maturidades mais longas. Além disso, modelar a volatilidade estocástica específica de cada fator pode propiciar uma informação importante no cálculo de risco de carteiras de título públicos e derivativos. Alternativas a serem abordadas em trabalhos futuros pode ser a inclusão de fatores adicionais para curvatura e inclinação, como proposto por Laurini and Westin (2010), tornando o modelo mais flexível, e também podem ser consideradas outras especificações para estrutura da volatilidade, considerando uma estrutura multivariada ou mudanças de regime.

Referências

- ALEXANDER, C. Optimal hedging using cointegration. *Philosophical Transactions of the Royal Society, London A.357* (1999), 2039–2058.
- ALEXANDER, C., AND DIMITRIU, A. The cointegration alpha: Enhanced index tracking and long-short equity market neutral strategies. *ISMA Discussion Papers in Finance 08* (2002).
- ALMEIDA, C., GOMES, R., LEITE, A., AND VICENTE, J. Does curvature enhance forecasting? Working Papers Series 155, Central Bank of Brazil, Research Department, Dec. 2007.
- ALMEIDA, C., AND VICENTE, J. The role of no-arbitrage on forecasting: Lessons from a parametric term structure model. *Journal of Banking & Finance* 32, 12 (December 2008), 2695–2705.
- ANDERSON, B., AND MOORE, J. *Optimal Filtering*. Englewood Cliffs: Prentice Hall, New York, 1979.
- ANDERSON, B. D. O., AND MOORE, J. B. *Optimal Filtering*. Englewood Cliffs: Prentice Hall, 1979.
- ANG, A., AND PIAZZESI, M. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics* 50, 4 (May 2003), 745–787.
- ANSON, M. J. *Handbook of Alternative Assets*. John Wiley and Sons, New York, 2006.
- BALI, T., HEIDARI, M., AND WU, L. Predictability of interest rates and interest-rate portfolios by turan bali, massoud heidari, liuren wu, 2006.
- BEST, M., AND GRAUER, R. Capital asset pricing compatible with market value weights. *Journal of Finance* 4 (1985), 85–103.
- BHAR, R. Yield curve as a cointegrated system: Evidence from australian treasury securities. *Working Paper*, 35 (1994).
- BJORK, T., AND CHRISTENSEN, B. J. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance* (1999), 323–348.
- BLACK, F., AND LITTERMAN, R. Asset allocation: Combining investors views with market equilibrium. *Fixed Income Research. Goldman, Sachs & Company* (1990).

- BLACK, F., AND LITTERMAN, R. Global portfolio optimization. *Financial Analysts Journal* 48 (1992), 28–43.
- BRIGO, D., AND MERCURIO, F. *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*. Springer, 2006.
- CALDEIRA, J. F., MOURA, G. V., AND PORTUGAL, M. S. Efficient interest rate curve estimation and forecasting in brazil. *Economia*, forthcoming (2010).
- CALDWELL, T. *Introduction: The Model for Superior Performance*. Irwin Professional Publishing, 1995.
- CAPOCCI, D. The sustainability of hedge fund performance: New insights. *Journal of Financial Management* 1, 2 (2007).
- CAPOCCI D., HÜBNER, G. L'univers des hedge funds, une perspective empirique. *Revue Bancaire et Financière*, 6 (2004), 361–369.
- CHIB, S., OMORI, Y., AND ASAI, M. Multivariate stochastic volatility. In *Handbook of Financial Time Series*, T. Andersen, R. Davis, J.-P. Kreiß, and T. Mikosch, Eds. 2009, pp. 365–400.
- CHOI, I. Durbin hausmann tests for a unit root. *Oxford Bulletin of Economic Statistics* 54 (1992), 289–304.
- COCHRANE, J. H. A critique of the application of unit root tests. *Journal of Economic Dynamics Control* 15 (1991), 275–284.
- COCHRANE, J. H., AND PIAZZESI, M. Bond risk premia. *American Economic Review* 95, 1 (March 2005), 138–160.
- COX, J. C., INGERSOLL, JONATHAN E, J., AND ROSS, S. A. A theory of the term structure of interest rates. *Econometrica* 53, 2 (March 1985), 385–407.
- DAI, Q., AND SINGLETON, K. J. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics* 63, 3 (March 2002), 415–441.
- DICKEY, D. A., AND FULLER, W. A. Distribution of the estimates for autoregressive time series with a unit root. *Journal of American Statistical Association* 74 (1979), 427–429.
- DIEBOLD, F. X., AND LI, C. Forecasting the term structure of government bond yields. *Journal of Econometrics* 130, 2 (February 2006), 337–364.
- DIEBOLD, F. X., AND MARIANO, R. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13 (1995), 253–263.

- DIEBOLD, F. X., RUDEBUSCH, G. D., AND BORAG[CARON]AN ARUOBA, S. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* 131, 1-2 (2006), 309–338.
- DUFFEE, G. R. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57, 1 (02 2002), 405–443.
- DUFFIE, D., AND KAN, R. A yield-factor model of interest rates. *Mathematical Finance* 6, 4 (1996), 379–406.
- DUNIS, C. L., AND HO, R. Cointegration portfolios of european equities for index tracking and market neutral strategies. *Journal of Asset Management* 6, 1 (2005), 33–52.
- DURBIN, J., AND KOOPMAN, S. *Time Series Analysis by State Space Methods*. Oxford, New York, 2001.
- EGOROV, A. V., LI, H., AND NG, D. A tale of two yield curves: Modeling the joint term structure of dollar and euro interest rates. *Journal of Econometrics, Forthcoming* (2010).
- ENDERS, W. *Applied Econometric Times Series*, 2 ed. John Wiley and sons.,Inc, 2002.
- ENGLE, R. F., AND GRANGER, C. W. J. Cointegration and error correction: Representation, estimation and testing. *Econometrica* 55 (1987), 251–276.
- ENGLE, R. F., AND NG, V. K. Time-varying volatility and the dynamic behavior of the term structure. *Journal of Money, Credit and Banking* 3, 25 (1993), 336–349.
- ENGLE, R. F., NG, V. K., AND ROTHSCILD, M. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics* 1-2, 45 (1990), 213–237.
- ENGLE, R. F., AND YOO, B. S. Forecasting and testing in cointegrated systems. *Journal of Econometrics* 35 (1987), 143–159.
- FAMA, E. F., AND BLISS, R. R. The information in long-maturity forward rates. *American Economic Review* 77, 4 (September 1987), 680–92.
- FILIPOVIC, D. *Term Structure Models*. Springer Finance, 2009.
- HAMILTON, J. D. *Times Series Analysis*. IE-Princeton, 1994.
- HAUSTSCH, N., AND OU, Y. Analyzing interest rate risk: Stochastic volatility in the term structure of government bond yields. CFS Working Paper 2009-03, Center for Financial Studies, Mar. 2009.

- HE, G., AND LITTERMAN, R. The intuition behind black-litterman model portfolios. *Investment Management Research*. Goldman, Sachs Company (1999).
- HEATH, D., JARROW, R., AND MORTON, A. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica* 60, 1 (January 1992), 77–105.
- HENDRY, D. F. E JUSELIUS, K. Explaining cointegration analysis: Part i. *Energy Journal* 21, 1 (2000).
- HENDRY, D. F. E JUSELIUS, K. Explaining cointegration analysis: Part ii. *Energy Journal* 22, 1 (2001), 75–20.
- HÖRDAHL, P., TRISTANI, O., AND VESTIN, D. A joint econometric model of macroeconomic and term-structure dynamics. *Journal of Econometrics* 131, 1-2 (2006), 405–444.
- HULL, J., AND WHITE, A. Valuating derivative securities using the explicit finite difference method. *The Journal of Financial and Quantitative Analysis* 25 (1990), 87–100.
- IDZOREK, T. A step-by-step guide to the black-litterman model. Zephyr Associates, Inc., 2004.
- JOHANSEN, S. Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica* 59 (1991).
- JOHANSEN, S., AND JUSELIUS, K. Maximum likelihood estimation and inference on cointegration - with application to the demand for money. *Oxford bulletin of Economic and Statistics* 52 (1990).
- JUMAH, A., AND KUNST, R. M. On mean reversion in real interest rates: An application of threshold cointegration. *Reihe Okonomie - Economic Series*, 109 (2002).
- KOOPMAN, S. J., JUNGBACKER, B., AND HOL, E. Forecasting daily variability of the s&p 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance* 12 (2005), 445–475.
- KOOPMAN, S. J., MALLEE, M. I., AND VAN DER WEL, M. Analyzing the term structure of interest rates using the dynamic nelson-siegel model with time-varying parameters. *Journal of Business and Economic Statistics forthcoming* (2010).
- LAURINI, M. P., AND HOTTA, L. K. Bayesian extensions to diebold-li term structure model. Insper Working Papers wpe 120, Insper Ibmecc São Paulo, Oct. 2008.
- LAURINI, M. P., AND WESTIN, A. P. K. Arbitragem na estrutura a termo das taxas de juros: uma abordagem bayesiana. Insper Working Papers wpe 120, Insper Ibmecc São Paulo, Jan. 2010.

- LEE, W. *Advanced Theory and Methodology of Tactical Asset Allocation*. New York: John Wiley & Sons, 2000.
- LIN, Y. X.; MCCRAE, M. G. C. Loss protection in pairs trading through minimum profit bounds: A cointegration approach. *Journal of Applied Mathematics and Decision Sciences* (2006), 1–14.
- LOOMIS, C. The jones nobody keep up with. *Fortune Magazine* (1966), 237–248.
- LUC BAUWENS, D. D., AND VANDEUREN, J.-P. Modeling interest rate with a cointegrated var-garch model. *Core Discussion Paper*, 9780 (1997).
- LUCAS, A. Strategic and tactical asset allocation and the effect of long-run equilibrium relations. *Research Memorandum*, 42 (1997), 1–32.
- MATZNER-LOBER, E., AND VILLA, C. Functional principal component analysis of the yield curve. In *21th International Conference AFFI* (2004), Association Francaise de Finance.
- MCCRACKEN, MICHAEL, W. Asymptotics for out of sample tests of granger causality. *Journal of Econometrics*, 140 (August 2007), 719–752.
- MCCULLOCH, J. H. Measuring the term structure of interest rates. *Journal of Business* 44, 1 (January 1971), 19–31.
- MCCULLOCH, J. H. The tax-adjusted yield curve. *Journal of Finance* 30, 3 (June 1975), 811–30.
- MEUCCI, A. *Risk and Asset Allocation*. Springer Finance, 2005.
- MICHAUD, R., AND MICHAUD, R. *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Allocation*. New York: John Wiley & Sons, 1998.
- NELSON, C. R., AND SIEGEL, A. F. Parsimonious modeling of yield curves. *The Journal of Business* 60, 4 (1987), 473–489.
- NEWAY, W., AND WEST, K. A simple, positive semi-definite, heteroskedastic and autocorrelation consistent covariance matrix. *Econometrica*, 55 (1987), 703–708.
- PEARSON, N. D., AND SUN, T.-S. Exploiting the conditional density in estimating the term structure: An application to the cox, ingersoll, and ross model. *Journal of Finance* 49, 4 (September 1994), 1279–1304.
- PHILLIPS, P. C., AND PERRON, P. Testing for a unit root in time series regressions. *Biometrika* 75 (1988), 335–346.

- RACHEV, S. T., HSU, J. S. J., BAGASHEVA, B. S., AND FABOZZI, F. J. *Bayesian Methods in Finance*. New York: John Wiley & Sons, 2000.
- ROBERT, L., AND SCHEINKMAN, J. Common factors affecting bond returns. *Journal of Fixed Income* 1, 1 (1991), 54–61.
- RUDEBUSCH, G. D., AND WU, T. A macro-finance model of the term structure, monetary policy and the economy. *The Economic Journal* 118, 530 (2008), 906–926.
- SCHMIDT, P., AND PHILLIPS, P. C. B. Lm tests for a unit root in the presence of deterministic trends. *Oxford Bulletin Economic and Statistic* 54 (1992), 257–288.
- SIMON, D. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley, New Jersey, 2006.
- SVENSSON, L. E. O. Estimating and interpreting forward interest rates: Sweden 1992-1994. IMF Working Papers 94/114, International Monetary Fund, Sept. 1994.
- TAYLOR, S. J. *Modelling Financial Time Series*. John Wiley & Sons, 1986.
- VASICEK, O. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5, 2 (November 1977), 177–188.
- VASICEK, O. A., AND FONG, H. G. Term structure modeling using exponential splines. *Journal of Finance* 37, 2 (May 1982), 339–48.
- VICENTE, J., AND TABAK, B. M. Forecasting bonds yields in the brazilian fixed income market. Working Papers Series 141, Central Bank of Brazil, Research Department, Aug. 2007.
- WANG, G. H. K., AND YAU, J. A time series approach to testing for market linkage: unit root and cointegration tests. *Journal of Futures Markets* 14 (1994).
- WU, T. Monetary policy and the slope factor in empirical term structure estimations. Tech. rep., 2001.
- YU, W., AND ZIVOT, E. Forecastin the term structure of treasury and corporate yields: Dynamic nelson-siegel models evaluation. Ssrn working paper, SSRN, Oct. 2008.

APÊNDICE A - Coeficientes de correlação de Spearman

Na figura A.1 são apresentados os coeficientes de correlação de *rank* de Spearman das carteiras *index tracking*, considerando uma janela móvel de 63 observações (3 meses). São consideradas as três periodicidades de rebalanceamento e carteiras com composição pesos fixos (parte superior) e quantidades fixas (parte inferior). O coeficiente de correlação de *rank* de Spearman, ou ρ de Spearman, é uma medida de correlação não paramétrica que não depende de qualquer suposição sobre a natureza da relação entre as variáveis. Não é necessária, por exemplo, a suposição de que a distribuição conjunta das variáveis segue uma normal bivariada, como ocorre no cálculo do coeficiente de correlação de Pearson. Como em muitos outros métodos estatísticos não paramétricos, o ρ de Spearman é baseado nos *ranks* dos dados, em vez do valor observado. Deste modo, este coeficiente não é sensível a assimetrias na distribuição, nem à presença de *outliers*, não exigindo, portanto, que os dados provenham de duas populações normais. O coeficiente de correlação de Spearman é calculado usando a seguinte equação:

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

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

No caso das carteiras *index tracking*, o coeficiente de correlação de Spearman se manteve sempre acima de 0.80 para todas as carteiras consideradas, e na maior parte do tempo foi superior a 0.90. Assim, parece evidente que as carteiras *index tracking* otimizadas através do método de cointegração se mantêm aderentes ao índice de referência, mesmo quando se considera uma janela de tempo menor.

Já a figura A.2 traz os coeficientes de correlação de Spearman para a estratégia *long-short*. Da mesma forma são consideradas as três periodicidades de rebalanceamento e carteiras com composição percentual fixa (parte superior) e com quantidades fixas (parte inferior). Neste caso, o que se espera é que o coeficiente de correlação se mantenha o mais próximo de zero. Tanto no caso em que a composição percentual da carteira foi mantida fixa durante sua vigência, quanto no caso onde a composição da carteira foi mantida fixa em termos de quantidade, informação mais próxima da realidade, o coeficiente se manteve em níveis satisfatórios por todo o período. Na maior parte do tempo, o coeficiente assumiu valores entre -0.20 e 0.20, o que é bastante satisfatório em termos de neutralidade. No período posterior a agosto de 2008 percebe-se um aumento na instabilidade do parâmetro, principalmente quando se considera

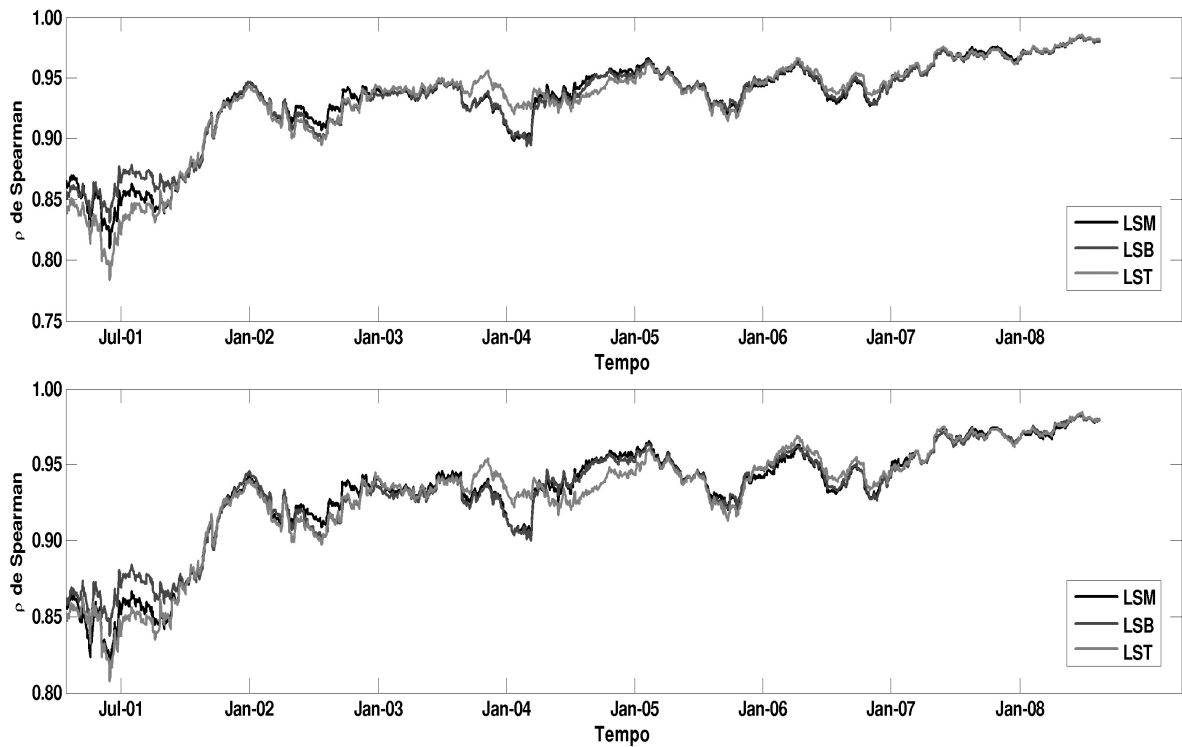


Figura A.1: Coeficiente de Correlação de Spearman - Estratégia *Index Tracking*

Fonte: Elaboração própria.

rebalanceamento bimestral. Nas outras duas periodicidades de rebalanceamento consideradas o coeficiente não assumiu valores superiores a 0.3, em termos absolutos. Assim, pode-se dizer que, mesmo quando se considera um horizonte mais curto, a estratégia *long-short* com carteiras otimizadas via cointegração, satisfazem o requisito de neutralidade ao mercado.

As tabelas A.2 e A.3 trazem os custos de aluguel de algumas ações no mercado brasileiro. As informações foram obtidas junto a CBLC e referem às taxas cobradas em determinado dia do mês de dezembro de 2008. Apesar das taxas refletirem as condições vigentes no mercado na ocasião de sua obtenção, elas refletem os custos que vigoraram no mercado ao longo do período analisado.

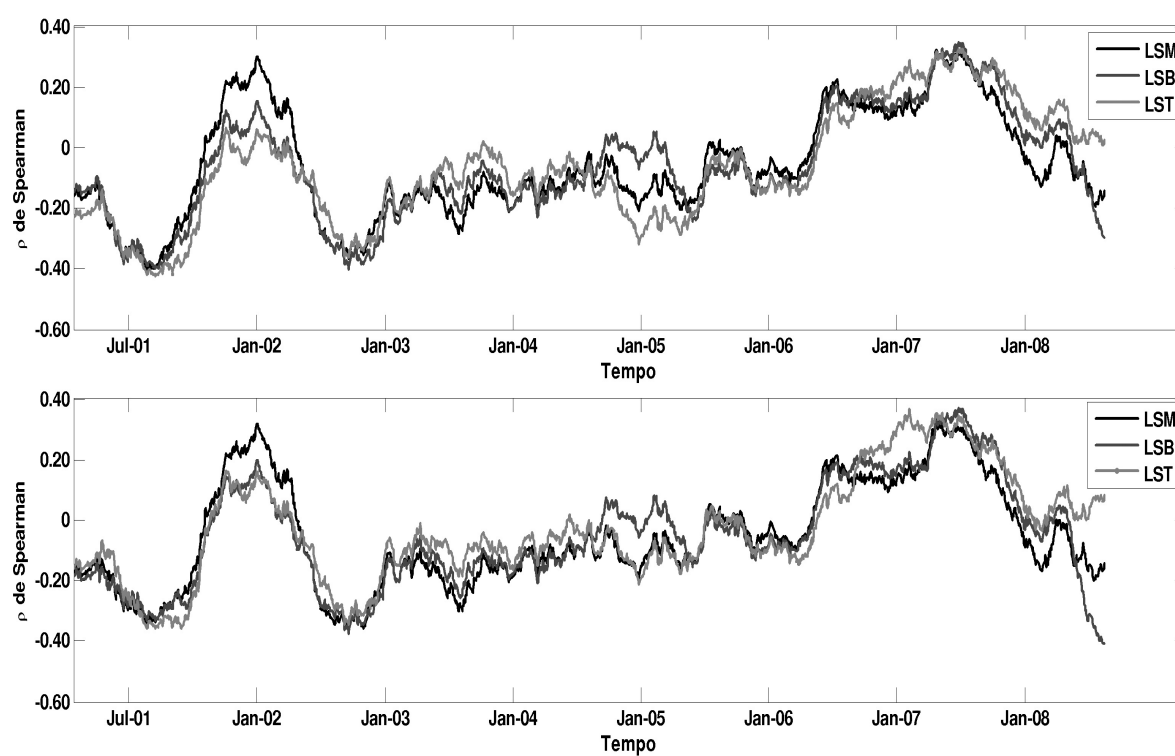


Figura A.2: Coeficiente de Correlação de Spearman - Estratégia *Long-Short*

Fonte: Elaboração própria.

Tabela A.1: Estatísticas dos Resultados Anuais das Carteiras *Index Tracking* e Ibovespa (Jan-01 a Dez-08).

Ano	Ibovespa	ITWFM	ITWFB	ITWFT	ITQFM	ITQFB	ITQFT
Retorno	-11.02	-11.59	-5.17	-8.38	-9.53	-7.64	-5.05
Corr Ibov	1.00	0.90	0.92	0.89	0.90	0.91	0.88
Curtose	4.37	4.00	3.89	3.71	3.92	3.91	3.69
Assimetria	-0.07	-0.13	-0.20	-0.09	-0.18	-0.23	-0.13
Retorno	-17.01	-29.08	-26.47	-21.26	-31.07	-30.84	-25.37
Corr Ibov	1.00	0.97	0.96	0.94	0.96	0.96	0.96
Curtose	2.89	3.71	3.55	3.22	3.78	3.67	3.29
Assimetria	-0.17	-0.35	-0.32	-0.23	-0.38	-0.30	-0.24
Retorno	97.33	104.49	104.27	92.92	106.38	107.57	97.51
Corr Ibov	1.00	0.93	0.94	0.95	0.92	0.92	0.94
Curtose	4.22	4.85	4.76	4.60	5.04	4.97	4.82
Assimetria	0.26	0.60	0.56	0.59	0.62	0.58	0.62
Retorno	17.81	20.96	23.57	14.96	22.68	24.91	15.63
Corr Ibov	1.00	0.96	0.94	0.96	0.96	0.95	0.95
Curtose	3.62	3.52	3.50	3.53	3.53	3.51	3.52
Assimetria	-0.35	-0.23	-0.23	-0.19	-0.23	-0.24	-0.19
Retorno	27.71	42.67	44.83	38.76	38.8	41.04	36.75
Corr Ibov	1.00	0.96	0.96	0.96	0.96	0.95	0.96
Curtose	2.86	2.88	2.86	2.89	2.85	2.86	2.89
Assimetria	-0.21	-0.28	-0.23	-0.25	-0.25	-0.23	-0.25
Retorno	32.93	28.27	25.89	23.02	30.84	27.87	26.63
Corr Ibov	1.00	0.96	0.96	0.96	0.96	0.96	0.96
Curtose	3.60	3.36	3.51	3.52	3.34	3.49	3.47
Assimetria	0.02	-0.03	0.00	0.00	-0.04	-0.01	-0.02
Retorno	43.65	45.6	41.36	46.61	45.28	43.25	48.4
Corr Ibov	1.00	0.98	0.98	0.98	0.98	0.98	0.98
Curtose	3.81	3.70	3.68	3.64	3.72	3.73	3.66
Assimetria	-0.57	-0.39	-0.40	-0.40	-0.41	-0.43	-0.44
Retorno	-41.22	-36.63	-34.78	-32.87	-37.71	-37.67	-33.51
Corr Ibov	1.00	0.99	0.99	0.99	0.99	0.99	0.98
Curtose	4.01	4.39	4.24	4.55	4.47	4.3	4.72
Assimetria	0.07	0.11	0.10	0.21	0.13	0.12	0.26

Fonte: Elaboração própria.

Tabela A.2: Custo do Aluguel de Acoes no Mercado Brasileiro

Ticker	Empresa	Emprestimos QTDE (CBLC)	Numr de Ações Circulação (CBLC)	Taxa % Média (aa)
VALE5	VALE	64,742,414	2,010,613,915	0.0029
ITSA4	ITAUSA	6,051,241	2,197,411,770	0.0052
CMIG4	CEMIG	12,437,687	274,033,130	0.0127
BRAP4	BRADESPAR	1,369,773	222,594,115	0.0077
CPLE6	COPEL	800,135	100,930,675	0.0151
GGBR4	GERDAU	27,615,411	621,797,235	0.0219
TNLP4	TELEMAR	4,998,467	254,747,800	0.0326
AMBV4	AMBEV	1,883,099	144,834,075	0.00545
PETR4	PETROBRAS	46,750,863	3,126,682,060	0.0033
NETC4	NET	1,203,873	197,208,850	0.01245
USIM5	USIMINAS	10,540,704	239,773,490	0.0097
CSNA3	SID NACIONAL	8,737,262	409,809,460	0.009
EMBR3	EMBRAER	3,757,544	721,469,730	0.0312
BRTO4	BRASIL TELEC	977,403	118192615	0.0131

Fonte: Elaboração própria.

Nota: Os custos de transacao e emolumentos somam 5 bps.

Tabela A.3: Custo do Aluguel de Acoes no Mercado Brasileiro

Ticker	Empresa	Emprestimos QTDE (CBLC)	Numr de Ações Circulação (CBLC)	Taxa % Média (aa)
PCAR5	P.ACUCAR-CBD	3,863,225	ND	0.02285
VCPA3	VCP	1,869,209	222,354,405	0.0711
EVEN3	EVEN	2,279,270	64,441,735	0.04115
CCPR3	CYRE COM-CCP	312,519	36,165,270	0.01
POMO4	MARCOPOLO	49,628	125,988,230	0.02005
LUPA3	LUPATECH	658,613	33,536,485	0.0879
CTAX4	CONTAX	132,278	8,776,255	0.0096
MMXM3	MMX MINER	2,515,260	106,273,600	0.0557
KLBN4	KLABIN SA	9,152,719	454,707,380	0.0422
GOAU4	GERDAU MET	1,396,027	262,602,050	0.0051
MPXE3	MPX ENERGIA	264,720	32,885,000	0.0361
NATU3	NATURA	3,151,199	112,014,425	0.0833
VIVO4	VIVO	3,446,422	120,891,820	0.0391
BRTP4	BRASIL T PAR	2,514,015	153,291,680	0.0173

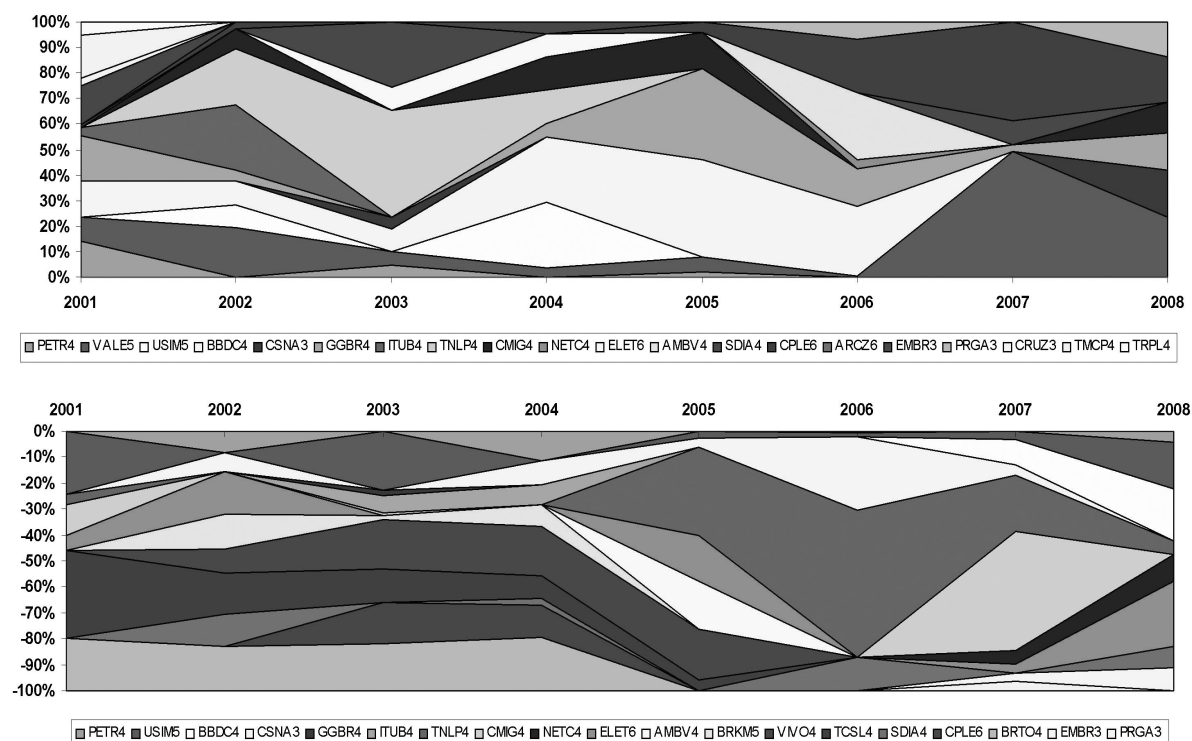
Fonte: Elaboração própria.

Nota: Os custos de transacao e emolumentos somam 5 bps.

APÊNDICE B - Composição das Carteiras

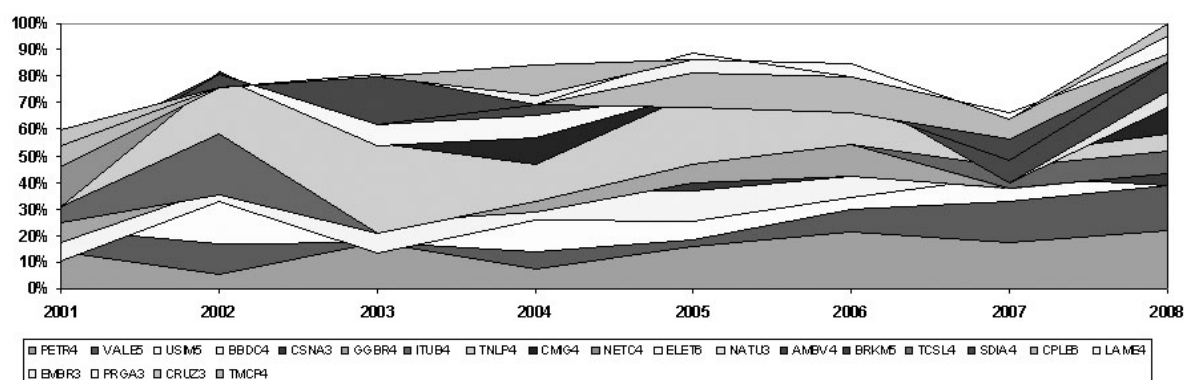
A figura A.2 traz as composições das carteiras *index-tracking* ao início de cada ano para o período de 2001 a 2008. Pode-se perceber a estabilidade da composição das carteiras ao longo do período analisado.

A composição das carteiras *long e short* para o mesmo período são apresentadas na figura B.1. Mais uma vez observa-se a estabilidade das carteiras ao longo do período analisado.



FiguraB.1: Composição das carteiras *Long e Short* no início de cada ano (Jan-01 a Dez-08)

Fonte: Elaboração própria.



FiguraB.2: Composição das carteiras *Long e Short* no início de cada ano (Jan-01 a Dez-08)

Fonte: Elaboração própria.