Universidade de São Paulo

Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Departamento de Economia

Programa de Pós-Graduação em Economia Área: Economia Aplicada

ROBERTO BALTIERI MAUAD

Análise da série do índice de Depósito Interfinanceiro: modelagem da volatilidade e apreçamento de suas opções

Orietador: Prof. Dr. Márcio Poletti Laurini

Ribeirão Preto - SP 2013

Prof. Dr. João Grandino Rodas Reitor da Universidade de São Paulo

Prof. Dr. Sigismundo Bialoskorski Neto Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

> Prof. Dr. Sérgio Kannebley Júnior Chefe do Departamento de Economia

ROBERTO BALTIERI MAUAD

Análise da série do índice de Depósito Interfinanceiro: modelagem da volatilidade e apreçamento de suas opções

Dissertação apresentada ao Programa de Pós-Graduação em Economia - Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências. Versão corrigida. A original encontra-se disponível na FEA-RP/USP

Orietador: Prof. Dr. Márcio Poletti Laurini

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRA-BALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

FICHA CATALOGRÁFICA

Mauad, Roberto Baltieri

Análise da série do índice de Depósito Interfinanceiro: modelagem da volatilidade e apreçamento de suas opções. Ribeirão Preto, 2013.

85 p.: il.; 30cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo.

Orientador: Laurini, Márcio Poletti

1. Apreçamento de opções 2. Modelo HJM Gaussiano 3. Regressão não paramétrica

FOLHA DE APROVAÇÃO

Nome: MAUAD, Roberto Baltieri	
Título: Análise da série do índice de De	epósito Interfinanceiro:
modelagem da volatilidade e apreçamen	nto de suas opções
	Dissertação apresentada ao Programa de Pós-Graduação em Economia - Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.
Aprovada em:	
BANCA	EXAMINADORA
Prof. Dr.	
Instituição:	Assinatura:
Prof. Dr	
	Assinatura:
Prof. Dr	
	Assinatura:

AGRADECIMENTOS

Escrever uma dissertação de mestrado não é uma tarefa fácil, e a ajuda e apoio das pessoas queridas auxiliam muito nessa jornada. Quero agradecer àqueles que direta ou indiretamente me apoiaram e, cada um à sua maneira, colaboraram para essa dissertação.

Primeiramente quero agradecer muito à minha família, que sempre me deu apoio, carinho e ajuda em tudo o que precisei. Ao meu pai Miguel, minha mãe Margarida e meu irmão Ricardo, muito obrigado por todas as conversas, dicas e por estarem presentes em momentos importantes da minha vida, como é esse momento da elaboração e defesa da dissertação.

A todos os professores da faculdade com quem tive aula e/ou oportunidade de conversar pessoalmente, em especial ao meu orientador Márcio Laurini, que sempre se mostrou muito solicito em ajudar e orientar. Dado seu vasto e aprofundado conhecimento da área, tivemos inúmeras conversas produtivas que me auxiliaram muito para que esse trabalho se concluísse da melhor forma possível. Muito obrigado pela excelente orientação.

A todos os grandes amigos que fiz durante o curso do mestrado. Daniel, Laura e Carol, vocês são grandes companheiros de batalha, não sei o que eu faria naquele primeiro ano repleto de matérias sem a ajuda de vocês! Risadas, finais de semana inteiros na FEA e estudo, estudo, estudo sem fim.

Bons amigos do segundo ano também me auxiliaram: Guilherme Byrro, Guilherme coquinho, João, Ivan, Wilson, Rafael Ravanelli, Grazi, e outros mais que estavam sempre na sala da pós.

Aos meus grandes amigos de fora do mestrado, pelo companheirismo, apoio e ajuda, em especial ao André, que sempre auxiliou bastante. Muito obrigado.

À FAPESP, que me auxiliou financeiramente durante a elaboração da dissertação, meus agradecimentos.

Sum'ario

Т	ista	do	T_{2}	hal	امدا
ш	asta.	ae	12	nei	as

Lista de Figuras

1	Intr	odução		p. 13
2	Rev	isão B	ibliográfica	p. 17
3	Abo	bordagem Teórica		
	3.1	Model	os de Finanças em tempo contínuo	p. 24
	3.2	Mudar	nça de medida de probabilidade	p. 27
	3.3	Model	os de curto prazo para a taxa de juros	p. 30
		3.3.1	Modelo Vasicek	p. 30
			3.3.1.1 Modelo Vasicek com dois fatores	p. 31
		3.3.2	Modelo Cox-Ingersoll-Ross	p. 35
	3.4	Heath,	Jarrow e Morton	p. 35
	3.5	Dinâm	ica do preço dos títulos e a restrição de deslocamento	p. 38
	3.6	Model	o HJM Gaussiano	p. 40
4	Abo	ordager	m teórica da estimação pelo Método de kernel	p. 42
	4.1	A Sele	ção do Parâmetro de Suavização Univariado	p. 45
		4.1.1	Seleção do parâmetro de suavização: métodos de validação cruzada	p. 46
		4.1.2	Validação Cruzada por Mínimos Quadrados	p. 46
		4.1.3	Validação Cruzada por Verossimilhanca	p. 47

	4.2	Regressão não paramétrica	p. 48
		4.2.1 Estimação de kernel constante local	p. 48
	4.3	Seleção de janela constante local	p. 50
		4.3.1 Métodos Rule-of-thumb e Plug-in	p. 50
		4.3.2 Validação cruzada de mínimos quadrados	p. 51
		4.3.3 Critério AIC_c para seleção de janela	p. 52
	4.4	Seleção de janela linear local: validação cruzada por mínimos quadrados	p. 52
	4.5	Regressão polinomial local - o caso univariado	p. 53
	4.6	O caso multivariado	p. 54
5	Apr	reçamento não paramétrico de opções	p. 56
	5.1	Simulações	p. 56
	5.2	Apreçamento de opções sobre títulos zero cupom	p. 61
	5.3	Aplicação empírica do modelo - apreçamento das opções sobre o IDI	p. 72
	5.4	Conclusões	p. 82
R	Referências		

Lista de Tabelas

1	REQM - apreçamento de opção sobre um título zero cupom - Volatilidade constante no tempo - Amostra 250	p. 66
2	REQM - apreçamento de opção sobre um título zero cupom - Volatilidade constante no tempo - Amostra 500	p. 68
3	REQM - apreçamento de opção sobre um título zero cupom - Volatilidade variante no tempo - Amostra 250	p. 69
4	REQM - apreçamento de opção sobre um título zero cupom - Volatilidade variante no tempo - Amostra 500	p. 71
5	Média do volume de negócios em outubro de 2013 - IDI e DI Futuro	p. 73
6	Apreçamento de opção de compra do IDI - Pregão 04/04/2012 - Método 3	p. 76
7	Apreçamento de opções sobre IDI - Amostra com janelas de 22 dias $$. $$.	p. 79
8	Apreçamento de opções sobre IDI - Amostra com janelas de 125 dias $$.	p. 79
9	Apreçamento de opções sobre IDI - Amostra com janelas de 250 dias $$.	p. 80
10	Aprecamento de opcões sobre IDI - Amostra com janelas de 500 dias .	p. 80

Lista de Figuras

1	Simulações com variância constante - amostra 250	p. 63
2	Simulações com variância constante - amostra 500	p. 63
3	Simulações com variância alterando no tempo - amostra 250 $\ \ldots \ \ldots$	p. 64
4	Simulações com variância alterando no tempo - amostra 500	p. 64
5	Estrutura a Termo de DI Futuros - 1999-2013	p. 75
6	Função de volatilidade pelo modelo 1 - Dia 2012-15-10 - Amostra com janelas de 500 dias	p. 77
7	Função de volatilidade pelo modelo 4 - Dia 2012-15-10 - Amostra com janelas de 500 dias	p. 77
8	Função de volatilidade pelo modelo 4 - Dia 2012-15-10 - Amostra com janelas de 22 dias	p. 78
9	Função de volatilidade pelo modelo 4 - dia2012-15-10 - Amostra com janelas de 125 dias	n 78

Resumo

MAUAD, R. B. Análise da série do índice de Depósito Interfinanceiro: modelagem da volatilidade e apreçamento de suas opções. 2013. 85 p. Dissertação (Mestrado) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2013.

Modelos bastante utilizados atualmente no apreçamento de derivativos de taxas de juros realizam, muitas vezes, premissas excessivamente restritivas com relação à volatilidade da série do ativo objeto. O método de Black and Scholes e o de Vasicek, por exemplo, consideram a variância da série como constante no tempo e entre as diferentes maturidades, suposição que pode não ser a mais adequada para todos os casos. Assim, entre as técnicas alternativas de modelagem da volatilidade que vêm sendo estudadas, destacam-se as regressões por kernel. Discutimos neste trabalho a modelagem não paramétrica por meio da referida técnica e posterior apreçamento das opções em um modelo HJM Gaussiano. Analisamos diferentes especificações possíveis para a estimação não paramétrica da função de volatilidade através de simulações de Monte Carlo para o apreçamento de opções sobre títulos zero cupom, e realizamos um estudo empírico utilizando a metodologia proposta para o apreçamento de opções sobre IDI no mercado brasileiro. Um dos principais resultados encontrados é o bom ajuste da metodologia proposta no apreçamento de opções sobre títulos zero cupom.

Palavras Chave: 1. Apreçamento de opções 2. Modelo HJM Gaussiano 3. Regressão não paramétrica

Abstract

MAUAD, R. B. Analysis of Brazilian Interbank Deposit Index series: volatility modeling and option pricing. 2013. 85 p. Dissertation (Master Degree). Graduate School of Economics, Business and Accounting, University of São Paulo, Ribeirão Preto, 2013.

Many models which have been recently used for derivatives pricing make restrictive assumptions about the volatility of the underlying object. Black-Scholes and Vasicek models, for instance, consider the volatility of the series as constant throughout time and maturity, an assumption that might not be the most appropriate for all cases. In this context, kernel regressions are important technics which have been researched recently. We discuss in this framework nonparametric modeling using the aforementioned technic and posterior option pricing using a Gaussian HJM model. We analyze different specifications for the nonparametric estimation of the volatility function using Monte Carlo simulations for the pricing of options on zero coupon bonds and conduct an empirical study using the proposed methodology for the pricing of options on the Interbank Deposit Index (IDI) in the Brazilian market. One of our main results is the good adjustment of the proposed methodology on the pricing of options on zero coupon bonds.

Key words: 1. Option pricing 2. Gaussian HJM model 3. Nonparametric regression

1 Introdução

Um problema fundamental nos procedimentos de apreçamento de ativos é a determinação do modelo teórico para o processo gerador de dados que seja mais adequado para a obtenção dos preços corretos dos derivativos. A escolha do modelo utilizado no procedimento de precificação depende fundamentalmente de duas escolhas. O primeiro fator relevante na escolha do modelo adequado é a consistência deste modelo com a dinâmica observada para o ativo objeto, e que pode ser sumarizado por meio do ajuste deste modelo com os dados observados do ativo objeto ou então para os preços de derivativos observados no mercado. O segundo ponto fundamental é a complexidade analítica do modelo. Este aspecto está relacionado à facilidade de se obter fórmulas analíticas para a precificação dos derivativos desejados e procedimentos para a calibração ou estimação dos parâmetros relevantes neste modelo.

Em geral essas duas escolhas são conflitantes - modelos mais realistas para a dinâmica histórica do ativo objeto em geral estão associados à ausência de fórmulas analíticas para a precificação de opções, ou então levam à necessidade do uso de aproximações numéricas ou o uso de métodos de Monte Carlo para a o cálculo de preços de derivativos. Este trade-off entre um ajuste adequado para o ativo de referência e a complexidade analítica e computacional do cálculo de preços de derivativos é especialmente relevante no apreçamento de opções sobre taxas de juros. Nesse caso, o ativo de referência, a estrutura a termo de taxas de juros, é um objeto multidimensional, e sua dinâmica temporal é bastante complexa. A evolução observada nos modelos propostos para a taxa de juros ilustra bem os dilemas entre complexidade na modelagem e ajuste adequado.

Os primeiros modelos para a taxa de juros apresentados na literatura são principalmente o de Vasicek e o de Cox-Ingersoll-Ross. Esses são bons instrumentos para modelar as taxas de juros de curto prazo, mas para aproximar a dinâmica de preços de derivativos, por exemplo, há outros estudos mais recentes que normalmente apresentam desempenhos superiores. Esses dois modelos são apresentados com mais detalhes na seção de metodologia desse trabalho e constatamos que, enquanto o modelo Vasicek apresenta um processo

de taxa de juros que retorna para uma média de longo prazo, o Cox-Ingersoll-Ross possui um processo que não permite taxas de juros negativas. Nesse contexto, surge uma oportunidade de aplicar outro modelo ao apreçamento de opções sobre o índice de Depósito Interfinanceiro: o modelo Heath, Jarrow and Morton (HJM).

Na realidade, de acordo com Barbachan e Ornelas (2003), o trabalho de HJM embasa uma classe de modelos, que se pautam basicamente no apreçamento por ausência de arbitragem. A ideia do modelo é partir de uma equação diferencial estocástica não para a taxa de juros de curto prazo, mas para os preços de um título pré-fixado. Assim, tornase necessário especificar as volatilidades de todas as taxas a termo instantâneas para o futuro, ou seja, definir a estrutura a termo da volatilidade. Em casos mais gerais do modelo HJM, o processo da taxa de juros em geral é não Markoviano, ou seja, o valor da taxa em dado momento não depende apenas de seu valor em um momento imediatamente anterior, mas sim de sua trajetória desde o início. Contudo, alguns casos particulares do modelo em questão podem ser gerados de modo que o processo da taxa de juros seja Markoviano. Esse é o caso dos modelos de Ho-Lee, Hull-White e Black, Derman e Toy (BDT).

Vale lembrar que esses três modelos são anteriores ao estudo de HJM, sendo que esse último realizou uma generalização de alguns estudos pré-existentes. No presente trabalho, fazemos uso do modelo HJM com características do estudo de Hull-White, em que as taxas de juros de curto prazo seguem uma distribuição normal (o modelo é gaussiano). De acordo com Gluckstern (2001), para o apreçamento de opções sobre o IDI, o modelo Hull-White apresenta um desempenho melhor que o modelo Black-Karasinski que, por sua vez, não permite taxas de juros negativas (por assumir um processo lognormal para essa variável).

O cálculo do preço de instrumentos derivativos, por sua vez, exige um trabalho de análise sobre a série do ativo objeto. Assim, é importante estudar possíveis tendências, comportamento da volatilidade, entre outros aspectos da série, dependendo do ativo a ser estudado. Especificamente, o valor das opções, quando estudadas na ausência de condições de arbitragem pelo modelo HJM, depende inteiramente das oscilações do ativo objeto. Assim, a especificação da chamada função de volatilidade, que captura a volatilidade do processo estocástico Browniano associado a cada curva forward, já é suficiente para o apreçamento do derivativo. A ideia desse resultado é que a dinâmica do drift, que é o deslocamento da variável dada uma variação no tempo, já é controlada pela ausência de condições de arbitragem que o modelo impõe. Assim, nessa classe de modelos, a curva de

juros é completamente caracterizada pela função de volatilidade e uma curva inicial de juros.

Dessa forma, desenvolvemos nesse trabalho basicamente a aplicação de um processo gerador de dados para preços de títulos zero cupom (por meio do modelo Vasicek com dois fatores), a estimação da função de volatilidade dessa série por meio de regressões de kernel (não paramétricas) e a utilização dessas funções de volatilidade como *input* para o apreçamento das opções utilizando a fórmula de Black generalizada. As referidas simulações de Monte Carlo são utilizadas no processo gerador de dados dos preços dos títulos zero cupom, e as regressões de kernel utilizadas para estimar a função de volatilidade são expressas em cinco diferentes especificações, que são comparadas entre si. O procedimento de precificação é feito por meio do modelo HJM Gaussiano, que é uma das especificações que o estudo de Heath-Jarrow-Morton permite. Assim, a distribuição das taxas de juros a termo é considerada normal.

Na parte empírica do trabalho utilizamos as mesmas especificações das regressões de kernel para estimar a função de volatilidade da série dos contratos DI do mercado brasileiro, de modo a precificar opções sobre o IDI. Os principais resultados que encontramos é que o modelo proposto apresenta um bom ajuste para encontrar o preço de opções sobre títulos zero cupom, e uma aproximação boa para os preços de mercado das opções sobre o IDI em alguns casos.

Com relação às especificações consideradas, conforme Musiela e Marek (2005), o modelo HJM Gaussiano considera que, em qualquer modelo de estrutura a termo, a volatilidade de qualquer ativo objeto seja determinista, ou seja, ela não será afetada por choques aleatórios. Uma das principais vantagens do modelo HJM Gaussiano, de acordo com Brace e Musiela (1994), é o cálculo confiável do preço de derivativos at-the-money por meio de fórmulas que podem ser derivadas matematicamente e programadas de forma razoavelmente simples. Portanto, no caso deste trabalho, a volatilidade do IDI será considerada determinista para fins de simplificação, lembrando que essa não é uma condição necessária para obter uma solução de forma fechada para o preço de uma determinada opção.

O modelo não paramétrico que utilizamos para estimar a volatilidade do IDI no Brasil, conforme mencionado, é o método de kernel multifatorial. A motivação em realizar um modelo multifatorial é a busca por mais informações, no sentido de explicar de forma mais acurada o efeito dos determinantes da volatilidade da taxa de juros sobre a variável dependente. Já a vantagem do modelo não paramétrico é evitar má especificação na

estrutura da volatilidade da taxa de juros em questão, devido à robustez deste método a problemas de especificação de uma estrutura paramétrica para a função de volatilidade, conforme exposto em Ait-Sahalia (1996).

Com relação aos dados utilizados, estes foram obtidos principalmente junto à Bolsa de Mercadorias e Futuros brasileira (BM&F), em especial informações sobre os contratos futuros de DI e preços de opções sobre o IDI. A taxa DI1 é a taxa média de depósitos interfinanceiros de um dia calculada pela Central de Custódia e de Liquidação Financeira de Títulos (Cetip), expressa em taxa efetiva anual, base 252 dias úteis. Já o objeto de negociação, ou ativo objeto da opção tratada nesse trabalho, é o já referido índice DI, que nada mais é que uma acumulação das taxas DI1, conforme definições da BM&F que são explicitadas mais adiante nesse trabalho.¹

Em termos práticos, a aplicação de um modelo com maior acuracidade para o cálculo do preço de derivativos pode ser bastante útil para fins proteção (hedge) e especulação. No sentido de calcular corretamente a quantidade de contratos a serem comprados ou vendidos para realizar um hedge adequado, é importante precificar o derivativo da melhor forma possível, de modo que o valor calculado seja significativamente próximo ao seu valor real. É nesse contexto que a aplicação do modelo proposto sobre o apreçamento de opções sobre o IDI se encaixa.

Uma operação de hedge que pode ser utilizada com as opções de compra sobre o IDI é a proteção de posições atreladas ao CDI. Uma empresa cuja dívida seja indexada a esse índice, por exemplo, pode optar por comprar opções de compra ("calls") sobre o IDI, de modo que uma elevação no contrato do índice DI permita ganhos com os derivativos que compensam a elevação no custo da dívida. O cálculo do número de opções a serem compradas depende do conhecimento do valor correto do derivativo, o que motiva a busca por um refinamento na metodologia de apreçamento dessas calls.

O texto está dividido da seguinte forma: no capítulo 2 é feita a revisão bibliográfica, que traz os trabalhos de alguns autores que contribuíram bastante para a literatura da área de pesquisa. No terceiro capítulo apresentamos a abordagem teórica, que embasa não somente as teorias que são diretamente aplicadas ao trabalho, como também aquelas que servem de referência para explicar a metodologia. Finalmente, no último capítulo, desenvolvemos o apreçamento das opções em questão, analisamos os resultados e apresentamos as principais conclusões.

 $^{^1\}mathrm{BM\&F}.$ Especificações do Contrato Futuro de Taxa Média de Depósitos Interfinanceiros de Um Dia. Disponível em http://www.bmf.com.br/bmfbovespa/pages/contratos1/Financeiros/PDF/DIfuturo.pdf. último acesso em 21/03/2013

2 Revisão Bibliográfica

Nessa parte do trabalho analisamos alguns estudos já realizados na área para embasar a metodologia e a aplicação do modelo proposto. São mencionados trabalhos que abordam metodologias similares às utilizadas aqui e conceitos importantes para o nosso modelo. Especificamente, fazemos referência a trabalhos que abordam técnicas de modelagem de taxas de juros, dentre os quais alguns consideram toda a estrutura a termo das taxas, enquanto outros não consideram. Ademais, mencionamos também trabalhos que realizam o apreçamento de derivativos com diversas técnicas, no sentido de motivar a escolha do nosso método.

Um artigo importante no contexto de modelagem de taxas de juros é o de Chan et al. (1992), em que se desenvolve a dinâmica de taxas de juros de curto prazo por meio de diversos modelos teóricos, utilizando a estimação por Método de Momentos Generalizada. Os dados do artigo são basicamente os retornos sobre Notas do Tesouro Americano, totalizando 307 observações, e um dos objetivos dos autores é comparar os modelos da dinâmica da taxa de juros, seja o não-restrito ou os modelos tradicionais com restrições de parâmetros.

Entre os resultados que os autores encontram, destaca-se o fato que, de acordo com o modelo não-restrito, há baixa evidência de reversão à média no processo de taxa de juros de curto prazo. Esse fato é importante já que grande parte da complexidade dos modelos de estrutura a termo advém dessa característica de reversão à média. Já com relação à volatilidade das taxas, o artigo mostra que os modelos são hábeis em capturar a variância que não é constante no tempo.

Com os dados sobre a dinâmica das taxas, os autores realizam também o apreçamento de opções sobre títulos de 30 anos. Fica claro que os preços calculados variam bastante entre os modelos. Assim, o artigo contribui com a importante constatação que uma boa modelagem da estrutura da taxa de juros é essencial para o bom apreçamento dos derivativos sobre essa taxa. Os autores constatam também que modelos afim para a

estrutura a termo da taxa de juros, como as versões unifatoriais de Cox-Ingersoll-Ross e Vasicek, normalmente permitem fórmulas simples para o apreçamento de títulos zero cupom e opções, mas seu ajuste para a estrutura a termo é pouco satisfatória.

Outro trabalho que faz uma boa exposição dos principais modelos de taxas de juros é o de Gluckstern (2001), em que realiza-se uma boa análise de vantagens e desvantagens de cada método. Os autores selecionam o modelo de Hull-White para precificar opções sobre o IDI após uma revisão da literatura teórica e empírica sobre os principais modelos. Assim, concluem que o modelo selecionado é superior ao empregado pelo mercado brasileiro (que é o modelo de Black).

Dentre as restrições do modelo de Black apontadas pelos autores, destaca-se que não é levada em conta toda a estrutura a termo da taxa de juros, de forma que muitas informações disponíveis no mercado são desconsideradas. Além disso, a diferença de maturidade entre os títulos de renda fixa e as ações gera uma deficiência do modelo de Black, que originalmente havia sido desenvolvido para precificar opções sobre ações, quando aplicado à renda fixa.

Os principais modelos de taxas de juros analisados por Gluckstern (2001) são: Black-Scholes e Merton, Vasicek, Cox-Ingersoll-Ross (CIR), Ho-Lee e Heath Jarrow Morton (HJM). Façamos uma breve discussão das constatações dos autores acerca desses modelos.

O modelo de Black-Scholes e Merton apresenta a possibilidade de gerar taxas de juros negativas, já que se considera a distribuição dos preços dos títulos de renda fixa como log-normal. Além dessa deficiência, utiliza-se a taxa de curto prazo como desconto para os preços dos títulos, supondo neutralidade ao risco. Porém, é incoerente pensar em volatilidade constante dos preços dos títulos quando a taxa de juros é determinista, já que esses ativos apresentam valor de face conhecido e a variância dos preços tende a diminuir com a redução do tempo até a maturidade.

O modelo Vasicek e o CIR, como são bastante importantes na literatura sobre taxas de juros, estão expostos com mais detalhes na parte da metodologia desse trabalho. Mas vale lembrar que Gluckstern (2001) constata que a principal contribuição do modelo Vasicek é basicamente a consideração que o modelo realiza acerca da arbitragem entre títulos de renda fixa, enquanto que a principal fraqueza está na suposição de um parâmetro exógeno, o preço de mercado do risco, considerado constante. Vale lembrar que essa suposição de constância no tempo quer dizer que a propensão ao risco dos investidores não se altera.

Com relação às constatações acerca do modelo CIR, pode ser mencionado que a taxa

de juros nunca pode ser negativa (pela construção do modelo) e a variância da taxa de juros aumenta com o seu valor, algo que é empiricamente observado. Contudo, tanto o CIR quanto o Vasicek deixam de explicar perfeitamente a estrutura a termo existente, de modo que surgem oportunidades de arbitragem entre a estrutura descrita pelo modelo e a observada no mercado. Dessa forma, sua aplicabilidade a problemas práticos de apreçamento é restrita.

O modelo Ho-Lee, por sua vez, considera processos de taxa de juros livres de arbitragem. Porém, como ele se baseia em árvores binomiais, uma vez que a árvore é construída e as taxas mudam, haverá divergências entre as taxas reais observadas e as taxas indicadas pela árvore.

Finalmente, o modelo de HJM, que utilizamos nesse trabalho, é um estudo da estrutura a termo considerado completo, que é especificado de forma que não haja possibilidade de arbitragem. A característica essencial desse modelo é a caracterização de condições suficientes para que haja uma única medida martingale equivalente, conceito que é explicado com mais detalhes na seção 3.2, sobre mudança de medida de probabilidade.

O estudo de Heath, Jarrow e Morton (1992), além de abordar a estrutura a termo da taxa de juros, apresenta um trabalho de apreçamento de derivativos. O trabalho mostra que a função de volatilidade é um *input* para o apreçamento de opções nesse modelo, já que os deslocamentos do processo da taxa a termo dependem somente da estrutura inicial e da volatilidade da taxa. Assim, a dinâmica da estrutura a termo e os preços dos instrumentos derivativos da taxa de juros são completamente determinados pela estrutura e pela volatilidade da taxa a termo.

Em linha com o trabalho de HJM, o artigo de Mercurio e Moraleda (2000), em que os autores propõem um modelo para a taxa de juros que possibilita uma estrutura de volatilidade em formato arqueado, é importante também para embasar nosso trabalho. Esse formato da estrutura de volatilidade foi constatado empiricamente por muitos estudos, e o modelo proposto pelos autores faz parte da classe de modelos HJM. Considerando títulos zero cupom numa economia em tempo contínuo, supõe-se que a dinâmica da taxa a termo apresente um componente determinista e um componente estocástico seguindo um processo de Wiener. Além disso, o artigo apresenta uma aplicação de outros estudos a respeito do apreçamento de uma opção de compra europeia sobre um título zero cupom, que realizamos em nosso trabalho também.

A aplicação empírica desse estudo de Mercurio e Moraleda (2000) apresenta uma comparação do modelo de Hull-White com o modelo Mercurio-Moraleda, proposto pelos

autores, para o apreçamento de opções sobre títulos zero cupom. Vale lembrar que os dados utilizados estão divididos em dois blocos: o primeiro possui informações acerca de preços de *caps* e *floors* de 3 meses expressos em dólares, enquanto o segundo apresenta dados sobre taxas de juros.

Os autores mostram que de fato o modelo proposto é mais preciso para precificar opções que o modelo de Hull-White, pois o erro de apreçamento computado foi menor. Em nosso trabalho, estamos também computando o erro de apreçamento pelo modelo proposto e comparando-o com uma forma mais usual de apreçamento de opções sobre títulos zero cupom, que é a análise de componentes principais. Assim, a diferença de nosso método em relação ao de Mercurio e Moraleda (2000) é basicamente na estimação da função de volatilidade: enquanto o nosso trabalho realiza regressões de kernel com diversos regressores diferentes, o dos autores citados realiza uma suposição paramétrica para a função de volatilidade que permite que ela assuma o formato arqueado. A principal vantagem do nosso modelo em relação a esses outros métodos que são paramétricos é que evitamos uma possível má especificação da função.

Uma referência muito importante na construção de modelos que utilizam estimação não paramétrica é o artigo de Jeffrey et al. (2004), no qual há uma generalização dessa técnica de estimação sobre modelos multifatoriais. No referido artigo, os autores apresentam a estimação não paramétrica de kernel sem realizar uma separação entre situações cross section e em séries temporais, sendo que a estimação proposta da volatilidade do ativo é bastante genérica. No teoria, funciona para qualquer caso, seja fixando uma data e variando a maturidade, seja fixando uma maturidade e variando o tempo, mas a aplicação prática da estimação proposta pelos autores é mais restrita. Assim, no nosso trabalho, estamos propondo modelos separados para lidar com essas situações, de modo que é possível aplicar o modelo à realidade brasileira e precificar as opções sobre o IDI. Esse trabalho complementa a análise proposta por Jeffrey et al. (2004) discutindo as propriedades em amostras finitas e distintas especificações do método de regressão não paramétrica para a estimação da função de volatilidade em modelos Heath-Jarrow-Morton e, especialmente, analisando as propriedades deste método no apreçamento de opções, uma análise ausente no trabalho original de Jeffrey et al. (2004).

Então, a estimação da função de volatilidade proposta neste trabalho segue a ideia apresentada em Jeffrey et al. (2004), em que a estrutura de volatilidade é estimada de forma não paramétrica a partir da curva de rendimento de um título zero cupom. Em nosso caso, a variável dependente é o retorno ao quadrado da curva do contrato DI em

função da maturidade e possíveis outros fatores que influenciam a taxa de juros. Essa estimação não paramétrica da função de volatilidade, conforme mencionado, é usada no procedimento de apreçamento em um modelo HJM Gaussiano, e os resultados são comparados com as formas usuais de apreçamento.

Outro trabalho que motiva essa dissertação é o de Vieira (1999), em que há uma boa exposição da modelagem da estrutura a termo da taxa de juros e a avaliação de contratos derivativos. O autor desenvolve, por meio do modelo Vasicek, a estimação da estrutura a termo e o apreçamento dos ativos em função dos parâmetros do processo da taxa de curto prazo e do preço de mercado do risco (ou prêmio pelo risco). No presente trabalho, são também precificadas opções sobre o IDI, porém utilizando o método HJM Gaussiano.

A principal contribuição deste último método é que, diferentemente do modelo Vasicek, ele não considera que o risco de mercado tenha origem em apenas um único fator de risco. Assim, é analisada neste trabalho a modelagem multifatorial da volatilidade da série dos contratos DI de um dia, que são a base de cálculo do IDI. Além disso, o modelo HJM utiliza como instrumento principal a taxa a termo, enquanto o Vasicek faz uso da taxa de juros de curto prazo para estimar a estrutura a termo. A classe de modelos HJM apresenta como um de seus casos particulares a estrutura proposta por Vasicek. A vantagem do HJM em relação ao Vasicek, portanto, consiste na possibilidade de se ajustar qualquer tipo de curva de rendimento pelo HJM, o que confere flexibilidade ao apreçamento de derivativos.

O trabalho de Barbedo, Lion e Vicente (2010), por sua vez, realiza o apreçamento das opções sobre o IDI utilizando também o modelo HJM, porém a estimação da volatilidade da série de juros é feita a partir da análise de componentes principais (PCA, em inglês), e não pelo método de kernel como é feito no presente trabalho. O artigo utiliza dados de negociação do IDI na Bolsa de Valores e no mercado de balcão e desenvolve o modelo considerando um, dois ou três fatores influenciando o índice. Na parte da aplicação do nosso modelo, utilizamos a estimação da volatilidade por PCA como referência de comparação, além de realizarmos duas regressões de kernel com componentes da variância calculados por PCA.

Quanto aos resultados, o trabalho conclui que a análise dos componentes principais não é suficiente para explicar os movimentos de preços de derivativos, já que foi realizado um teste para comparar as distribuições com um, dois ou três fatores e a hipótese nula de igualdade das distribuições foi rejeitada. Há erros quadráticos médios maiores na estimação dos índices negociados em balcão, já que a volatilidade nesse mercado parece ser

maior que a da Bolsa. Com relação às estimações do HJM, parece que o modelo tende a sobreavaliar opções com tempo até a maturidade maior que um ano, talvez devido à baixa liquidez do ativo. Além disso, o artigo realiza também regressões lineares do erro quadrático médio do apreçamento contra variáveis como a maturidade e moneyness, e constata que há uma relação direta entre o tempo até a maturidade e o erro de apreçamento e uma relação inversa entre esse erro e o moneyness. Lembrando que por moneyness nos referimos a quão "no dinheiro" está a opção, ou seja, a proximidade entre o preço de exercício e o valor do ativo objeto.

Com relação ao método de estimação da volatilidade, vale a pena discorrer sobre as vantagens do método não paramétrico em relação ao paramétrico. Sabe-se que os métodos paramétricos comumente utilizados na literatura somente captam o que está ocorrendo com a média da distribuição, ao invés da distribuição como um todo. Além disso, os métodos paramétricos exigem suposições a priori sobre a distribuição real da variável ou do formato das funções utilizadas, ao contrário dos não paramétricos. Assim, métodos não paramétricos são robustos à forma funcional de eventuais funções do modelo e à suposta distribuição da variável.

O trabalho de Ait-Sahalia (1996), por exemplo, realiza a estimação não paramétrica da função instantânea da volatilidade. O artigo desenvolve uma técnica para estimar o modelo usado na construção teórica (que é toda em tempo contínuo), utilizando apenas dados disponíveis (que estão em tempo discreto). O autor elabora o processo da taxa de juros baseado numa equação diferencial de Itô, na qual a função de volatilidade deve ser especificada por meio da técnica do density-matching. A estimação da dinâmica da taxa a termo apresentada pelo autor é feita de forma contrária às estimações tradicionais: ao invés de obter a distribuição da taxa a partir de um processo pré-determinado, como o de Ornstein- Uhlenbeck ou de Cox-Ingersoll-Ross, por exemplo, obtêm-se os parâmetros μ e σ^2 a partir da densidade estimada.

O trabalho utiliza dados de taxas do depósito Eurodollar do Bank of America entre 1973 e 1995 e realiza a estimação não paramétrica da volatilidade da série. É feita, também, a estimação não paramétrica de kernel da densidade, que por sua vez é comparada com as densidades obtidas pelo modelo CIR (Gamma) e Vasicek (Gaussiana). Assim, estima-se a estrutura a termo da taxa de juros e, então, é possível precificar as opções sobre títulos. É realizado o apreçamento com distintas taxas de juros anuais, levando em conta diferentes períodos até a maturidade e preços de exercício, fato que é feito também no presente trabalho. Ao final do artigo, o autor conclui que a estimação por

density-matching pode ser usada em diversos processos em finanças e economia.

Dessa forma, com base nesses trabalhos, desenvolveremos a aplicação do modelo HJM Gaussiano para os dados do IDI da BM&F, com o intuito de utilizar a volatilidade estimada pelo método de kernel para precificar as opções sobre esse ativo. Além dessa aplicação empírica, realizamos também o apreçamento de opções sobre títulos zero cupom, de modo a comparar o ajuste do modelo proposto com os modelos tradicionais de apreçamento, em especial o método de PCA.

3 Abordagem Teórica

A ideia desse capítulo do trabalho é apresentar a teoria que embasa a aplicação do modelo proposto, sendo que as informações contidas aqui são extraídas basicamente de livros e artigos acadêmicos da área. Vale lembrar que o conteúdo específico que utilizamos na aplicação do modelo no programa econométrico está exposto no próximo capítulo, intitulado Apreçamento não paramétrico de opções.

3.1 Modelos de Finanças em tempo contínuo

Nessa seção apresentamos a teoria de finanças em tempo contínuo, que serve de base para entender o conceito de processo de Itô que, por sua vez, depende do conceito de movimento Browniano. Assim, como consideramos na aplicação do modelo que a taxa de juros a termo segue um processo de Itô, é importante apresentar o desenvolvimento teórico desse conceito.

Segundo Munk (2003), na teoria moderna de finanças, muitas vezes processos estocásticos são usados para modelar a evolução de variáveis incertas no tempo. Assim, conhecer o funcionamento desses processos é de suma importância para desenvolver modelos da estrutura a termo da taxa de juros, que será um dos focos deste trabalho.

A metodologia será desenvolvida assumindo-se o processo em tempo contínuo, já que os cálculos normalmente são simplificados nesse caso. Um dos processos fundamentais com essa propriedade é o chamado movimento Browniano padrão ou processo de Wiener, a partir do qual podem-se derivar muitos outros processos de trajetória contínua que podem ser utilizados em modelagem financeira. De acordo com Valls e Vieira (2002), modelos de finanças em tempo contínuo normalmente apresentam 4 partes básicas: um conjunto de datas, um espaço de probabilidade filtrado, um processo de Wiener padrão m-dimensional e um processo de preços (ou taxas de juros, como é o caso do presente trabalho) n+1 dimensional.

O conjunto de datas refere-se basicamente à data inicial e final do modelo, além das datas em que os investidores podem negociar seus ativos. No caso contínuo, há infinitas datas em que é possível realizar negócios. Obviamente, na realidade, o tempo das negociações é discreto, mas quanto menor o intervalo entre trades, mais próximo da realidade será a aproximação em período contínuo. Vale lembrar que não é possível dizer, a priori, qual seria um espaço de tempo suficientemente pequeno para que a aproximação em período contínuo seja satisfatória, sendo que esse fato varia em cada caso, conforme constatado por Valls e Vieira (2002).

O espaço de probabilidade filtrado, por sua vez, normalmente é representado por uma quádrupla $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t \in T}, \mathbb{P})$, em que $(\mathbb{F}_t)_{t \in T}$ é uma filtração contida em \mathbb{F} . No caso, a filtração $(\mathbb{F}_t)_{t \in T}$ seria uma coleção de σ -álgebras do espaço amostral Ω , que satisfaça $\mathbb{F}_s \subseteq \mathbb{F}_t$ para todo s < t em T. Em outras palavras, a filtração é um conjunto de informações, que incorpora dados passados ao longo do tempo. Nesse contexto, é importante definir o conceito de "espaço de valor", que é o menor conjunto que contém todos os valores que o processo estocástico pode assumir.

Já o processo de Wiener (ou movimento Browniano) nada mais é que um passeio aleatório em tempo contínuo. Considerando, por simplificação, que um processo estocástico $z=(z_t)_{t\geq 0}$ tenha dimensão igual a um, ele será considerado um movimento Browniano padrão se:

- 1. $z_0 = 0$
- 2. Para todo $t, t' \geq 0$ com $t < t' : z_{t'} z_t \sim N(0, t' t)$
- 3. Para todo $0 \le t_0 < t_1 < ... < t_n$, as variáveis aleatórias $z_{t_1} z_{t_0}, ..., z_{t_n} z_{t_{n-1}}$ são mutuamente independentes
- 4. z apresenta trajetória contínua, mas não diferenciável quase certamente

Antes de explicar mais detalhadamente o movimento Browniano, é importante discorrer sobre equações diferenciais estocásticas (EDE), já que o processo da taxa de juros a termo segundo o modelo HJM segue uma EDE. De acordo com Oksendal (2003), o caso geral de uma EDE é dado por:

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t)W_t \tag{3.1}$$

em que W_t é um ruído branco. A interpretação de Itô dessa equação é que X_t satisfaça a seguinte equação integral estocástica:

$$X_{t} = X_{0} + \int_{0}^{t} b(s, X_{s})ds + \int_{0}^{t} \sigma(s, X_{s})dB_{s}$$
(3.2)

ou na forma diferencial:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t \tag{3.3}$$

Para resolver esse tipo de equação, um importante método a ser utilizado é a fórmula de Itô, que representa a regra de derivação para cálculos estocásticos. A interpretação econômica da equação estocástica geral apresentada acima é que o processo X_t pode ser dividido em uma parte determinista $b(t, X_t)dt$ e uma parte estocástica $\sigma(t, X_t)dB_t$, que são, no caso específico deste trabalho, fatores deterministas e estocásticos que influenciam a taxa de juros.

Dados esses conceitos de EDE, retomemos o detalhamento do processo de Wiener. A mudança esperada em um movimento Browniano padrão possui média igual a zero e a variância igual ao tamanho desse intervalo. Generalizando, assumimos que $z=(z_t)_{t\geq 0}$ seja um movimento Browniano padrão de dimensão igual a um e definimos um novo processo estocástico $x=(x_t)_{t\geq 0}$ como:

$$x_t = x_0 + \mu t + \sigma z_t, t \ge 0$$

Nessa equação, x_0 , μ , e σ são constantes. Das propriedades do movimento Browniano, sabe-se que $x_t \sim N(\mu t, \sigma^2 t)$. Já a mudança no valor do processo entre dois momentos arbitrários t e t' é:

$$x_{t'} - x_t = \mu(t' - t) + \sigma(z_{t'} - z_t)$$

Uma mudança no processo em um intervalo de tempo infinitesimal normalmente é representada como:

$$dx_t = \mu dt + \sigma dz_t$$

Em que dz_t pode ser interpretado como uma variável aleatória normalmente distri-

buída com média 0 e variância dt. O processo x, nesse caso, pode ser chamado de movimento Browniano generalizado. O parâmetro μ reflete o chamado "deslocamento", ou seja, a mudança esperada no processo por unidade de tempo, enquanto o parâmetro σ reflete a incerteza sobre valores futuros da variável.

Já o processo de taxas de juros do modelo financeiro apresentado nesse trabalho é um processo de Itô. Um processo estocástico unidimensional x_t é chamado de processo de Itô se a sua dinâmica apresentar o seguinte formato:

$$dx_t = \mu_t dt + \sigma_t dz_t$$

Nesse caso, o deslocamento μ_t e a volatilidade σ_t são também processos estocásticos. Em casos mais gerais de processos de Itô, o deslocamento e a volatilidade também podem depender de valores passados do processo x. Esse tipo de processo normalmente não é Markov nem martingale, e a evolução dele pode ser representada como:

$$x_{t'} - x_t = \int_t^{t'} \mu_u du + \int_t^{t'} \sigma_u dz_u$$

3.2 Mudança de medida de probabilidade

No presente trabalho, estamos realizando o apreçamento de ativos quando inexistem condições de arbitragem, que estabelece que os preços dos ativos sejam martingales na medida neutra ao risco. Assim, é necessário realizar uma mudança na medida de probabilidade para representarmos uma situação em que não há possibilidades de arbitragem, ou seja, não existem lucros positivos com risco zero.

Especificamente, na aplicação do modelo proposto, realizamos a mudança de probabilidade apenas no cálculo do preço das opções sobre títulos zero cupom, sendo que na parte empírica do trabalho, do apreçamento de opções sobre o IDI, só se observa a medida de probabilidade real.

Quando modelamos uma variável econômica por um processo estocástico, estamos impondo uma medida de probabilidade \mathbb{P} . Num processo de raiz quadrada, por exemplo, \mathbb{P} deve ser tal que $z=(z_t)$ seja um movimento Browniano padrão. Mas, de acordo com Munk (2003), para fins matemáticos, pode ser conveniente estipular uma medida de probabilidade diferente daquela do mundo real, para refletir uma situação em que os agentes sejam neutros ao risco. Assim, atendemos ao requisito de inexistência de

arbitragem necessário para o modelo de apreçamento que utilizamos.

Denotando um espaço amostral por $\Omega = \{\omega_1, \dots, \omega_n\}$ e o vetor de probabilidades \mathbb{P} por (p_1, \dots, p_n) , pode-se definir p_i como:

$$p_i = \mathbb{P}(\{\omega_i\}), i = 1, \dots, n$$

Então $p_i \in [0,1]$ e $\sum_{i=1}^n p_i = 1$. Mas como estamos interessados em uma medida de probabilidade equivalente \mathbb{Q} , devemos ter que $p_i > 0 \Leftrightarrow q_i > 0$ para todo $i = 1, \ldots, n$. Essa afirmação é equivalente a dizer que as medidas de probabilidade são equivalentes, ou seja, que elas atribuem probabilidade zero de ocorrência para os mesmos eventos. Essa mudança de probabilidade de \mathbb{P} para \mathbb{Q} pode ser representada pelo vetor $\xi = (\xi_1, \ldots, \xi_n)$, em que:

$$\xi_i = \frac{q_i}{p_i}, i = 1, \dots, n$$

O vetor ξ pode ser chamado de derivada de Radon-Nikodym de \mathbb{Q} com relação a \mathbb{P} e é denotada por $\frac{d\mathbb{Q}}{d\mathbb{P}}$. Note que $\xi_i > 0$ para todo i e que a expectativa em \mathbb{P} é:

$$E^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = E^{\mathbb{P}}[\xi] = \sum_{i=1}^{n} p_i \xi_i = \sum_{i=1}^{n} p_i \frac{q_i}{p_i} = \sum_{i=1}^{n} q_i = 1$$

Considerando uma variável aleatória X que assume o valor x_i quando o estado i é realizado, tem-se que o valor esperado de x sob a medida de probabilidade \mathbb{Q} é:

$$E^{\mathbb{Q}}[X] = \sum_{i=1}^{n} q_i x_i = \sum_{i=1}^{n} p_i \frac{q_i}{p_i} x_i = \sum_{i=1}^{n} p_i \xi_i x_i = E^{\mathbb{P}}[\xi x]$$

Definindo $\xi_t = E_t^{\mathbb{P}} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \right]$ tem-se que o processo ξ é \mathbb{P} -Martingale, já que para $t < t' \leq T$:

$$E_t^{\mathbb{P}}[\xi_{t'}] = E_t^{\mathbb{P}} \left[E_{t'}^{\mathbb{P}} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \right] \right] = E_t^{\mathbb{P}} \left[\frac{d\mathbb{Q}}{d\mathbb{P}} \right] = \xi_t$$

Seja $x = (x_t)_{t \in [0,T]}$ um processo estocástico, então sabe-se que:

$$E_t^{\mathbb{Q}}[x_{t'}] = E_t^{\mathbb{P}} \left[\frac{\xi_{t'}}{\xi_t} x_{t'} \right]$$
(3.4)

Seja $\lambda=(\lambda_t)_{t\in[0,T]}$ um processo estocástico. Pode-se gerar uma medida equivalente

de probabilidade \mathbb{Q}^{λ} definindo $\xi_t^{\lambda} = \exp\left\{-\int_0^t \lambda_s dz_s - \frac{1}{2} \int_0^t \lambda_s^2 ds\right\}$. A partir da definição de ξ_t^{λ} , pode-se definir uma medida equivalente de probabilidade \mathbb{Q}^{λ} por meio da derivada de Radon-Nikodym:

$$\frac{d\mathbb{Q}^{\lambda}}{d\mathbb{P}} = \xi_t^{\lambda} = \exp\left\{-\int_0^T \lambda_s dz_s - \frac{1}{2} \int_0^T \lambda_s^2 ds\right\}$$

Da equação 3.4, tem-se que:

$$E_t^{\mathbb{Q}^{\lambda}}[x_{t'}] = E_t^{\mathbb{P}} \left[\frac{\xi_{t'}^{\lambda}}{\xi_t^{\lambda}} x_{t'} \right] = E_t^{\mathbb{P}} \left[x_{t'} exp \left\{ -\int_t^{t'} \lambda_s dz_s - \frac{1}{2} \int_t^{t'} \lambda_s^2 ds \right\} \right]$$

para qualquer processo estocástico $x = (x_t)_{t \in [0,T]}$. Nesse contexto, segundo Munk (2003), é necessário apresentar o Teorema de Girsanov, no sentido de explicar o efeito da mudança na medida de probabilidade sobre o processo estocástico analisado.

Conforme exposto em Munk (2003), o Teorema de Girsanov mostra que um processo de Itô, ao ser analisado sob uma medida de probabilidade distinta da original, poderá apresentar um coeficiente de deslocamento diferente, mas a volatilidade do processo se mantém a mesma. Assim, a classe da distribuição do processo em questão não se altera com a mudança da medida de probabilidade, mas os momentos podem ser diferentes.

De acordo com Oksendal (2003), esse teorema mostra que se alterarmos o coeficiente de deslocamento de um dado processo de Itô com coeficiente de difusão não nulo, a dinâmica do processo não muda drasticamente. Na realidade, a dinâmica do novo processo será absolutamente contínua com relação à dinâmica do processo original e, portanto, podemos utilizá-lo sem prejuízo à análise em modelos de finanças. Assim, para ilustrar matematicamente a constatação desse teorema, apresentamos a seguir uma breve demostração que consta em Munk (2003):

Seja $z_t^{\lambda} = z_t + \int_0^t \lambda_s ds$ um movimento Browniano padrão sob a medida de probabilidade \mathbb{Q}^{λ} , em que $0 \le t \le T$. Se $x = (x_t)$ for um processo de Itô com dinâmica $dx_t = \mu_t dt + \sigma_t dz_t$, então:

$$dx_t = \mu_t dt + \sigma_t (dz_t^{\lambda} - \lambda_t dt) = (\mu_t - \sigma_t \lambda_t) dt + \sigma_t dz_t^{\lambda}$$

Assim, o deslocamento do processo sob a medida \mathbb{Q}^{λ} é $\mu - \sigma \lambda$, que é diferente do deslocamento sob a medida de probabilidade original \mathbb{P} , a não ser que σ ou μ sejam idênticos a zero. A variância, por sua vez, se mantém igual.

3.3 Modelos de curto prazo para a taxa de juros

Nessa seção do trabalho, apresentamos os modelos Vasicek unifatorial, Vasicek com dois fatores e Cox-Ingersoll-Ross. Esses modelos de curto prazo são apresentados com o objetivo de mostrar que os modelos unifatoriais podem não ser os mais adequados para encontrar a estrutura a termo da taxa de juros e, por esse e outros motivos, utilizamos o modelo HJM para essa finalidade.

Além disso, a exposição dos modelos a seguir visa embasar a metodologia utilizada na parte aplicada do trabalho, dado que, para encontrar o preço teórico das opções sobre títulos zero cupom, utilizamos o modelo Vasicek com dois fatores. Ademais, o modelo Cox-Ingersoll-Ross é também utilizado na aplicação do trabalho, quando realizamos a suposição de variância não constante no tempo, conforme apresentado na seção de aplicação do modelo.

Em linhas gerais, são apresentadas as suposições dos modelos, os processos estocásticos utilizados em cada um deles e as propriedades desses processos.

3.3.1 Modelo Vasicek

De acordo com Vieira (1999), esse modelo apresenta alguns avanços em relação ao modelo de Merton que, por sua vez, é o primeiro estudo sobre a estrutura a termo em que os preços dos títulos são obtidos em função do processo estocástico seguido pela taxa de juros de curto prazo. Um ponto importante que deve ser destacado é que, no modelo Vasicek, a taxa de juros não pode crescer indefinidamente, como no modelo de Merton, já que o processo estipulado para a taxa possui um termo θ que representa seu valor esperado de longo prazo.

Assim como no modelo de Merton, no Vasicek também pode ser demonstrado, por meio da aplicação do Lema de Itô sobre a dinâmica da taxa de juros de curto prazo, que a referida taxa segue uma distribuição normal. Então, conforme mencionamos na Revisão Bibliográfica, a taxa de juros no modelo Vasicek pode assumir valores negativos (fato que ocorre também no estudo de Merton). Porém, conforme argumentado em Vieira (1999), o modelo possui grande tratabilidade analítica.

Conforme apresentado na referida tese de Vieira (1999), se os investidores forem avessos ao risco (sendo que o prêmio de mercado do risco, λ , seja maior que zero), o retorno instantâneo esperado dos títulos aumenta linearmente com a volatilidade da taxa de juros

de curto prazo e não linearmente com o prazo para o vencimento do contrato. O modelo implica também que, na data de vencimento do contrato, a volatilidade dos preços dos títulos seja nula, o que faz bastante sentido prático.

A apresentação matemática do modelo em questão é feita em Back (2010), que mostra que o processo estocástico considerado para a taxa de juros de curto prazo é o Ornstein-Uhlenbeck, que possui a característica de reversão à média. Assim, a taxa de curto prazo converge para um valor de longo prazo nesse modelo, sendo que choques aleatórios podem desviar a taxa desse equilíbrio.

Nesse contexto é importante ressaltar que o modelo Vasicek unifatorial realiza uma boa modelagem da estrutura a termo da taxa de juros, sendo que Vieira (1999) utilizou o processo estocástico utilizado pelo modelo Vasicek em seu trabalho. Porém, como o modelo HJM permite uma flexibilização maior da estrutura a termo da taxa, faremos uso dele para desenvolver nosso trabalho. No caso, assumimos que a taxa de juros segue um processo de Itô, e encontramos sua função de volatilidade por meio de diversas regressões de kernel. Detalhes desse desenvolvimento são apresentados na seção de aplicação do modelo desse trabalho.

Conforme mencionado, faremos um desdobramento desse modelo para o caso do Vasicek com dois fatores, já que utilizamos esse modelo para encontrar o preço teórico das opções sobre títulos zero cupom no referido capítulo sobre a aplicação do nosso modelo.

3.3.1.1 Modelo Vasicek com dois fatores

Primeiramente, é importante mencionar que, para aplicar esse modelo (assim como alguns outros também) no programa econométrico R, é necessário realizar a discretização de Euler, que nada mais é que um procedimento para tornar o processo em tempo discreto. De acordo com Zivot e Wang (2003), são raros os casos em que uma equação diferencial estocástica apresenta solução explícita. Assim, há vários métodos para aproximar soluções usando discretização, tais como o método de Euler e os métodos de Platen.

O método de Euler é o mais simples, e a discretização de um processo estocástico X_{t_n} , que definimos como Y_t , sendo $t_n = t_0 + n\delta t$, é dado por:

$$Y_{n+1} = Y_n + a(t_n, Y_n)\delta t + b(t_n, Y_n)\delta W_n$$
(3.5)

Em que $Y_0 = X_0$ é a condição inicial. Os incrementos no movimento Browniano,

 $\delta W_n = W_{t_{n+1}} - W_{t_n}$, podem ser simulados usando variáveis normais geradas de forma pseudo aleatória, ou seja, a geração dos dados apresenta algum algoritmo implícito. Assim, cada etapa do procedimento de Euler requer M variáveis aleatórias normais.

Tendo em mente essa técnica de discretização, podemos analisar o modelo que utilizamos como referência para os preços das opções sobre títulos zero cupom. De acordo com Nawalkha, Beliaeva e Soto (2007), o modelo de Vasicek com dois fatores é um método alternativo à derivação de Hull-White, segundo o qual a taxa de juros de curto prazo é especificada como segue:

$$r(t) = \delta(t) + Y_1(t) + Y_2(t), \quad Y_1(0) = 0 \quad e \quad Y_2(0) = 0$$
 (3.6)

Os dois fatores mencionados no nome do modelo, que são variáveis de estado, são dados pelas seguintes equações:

$$dY_1(t) = -\tilde{\kappa}_1 Y_1(t) dt + \nu_1 d\tilde{W}_1, \quad Y_1(0) = 0$$
(3.7)

$$dY_2(t) = -\tilde{\kappa}_2 Y_2(t) dt + \nu_2 d\tilde{W}_2, \quad Y_2(0) = 0$$
(3.8)

Com os fatores latentes modelados, aplicamos o Lema de Itô sobre o preço dos títulos e tomamos a expectativa sob a medida neutra ao risco. Assim, resulta a equação diferencial estocástica para os preços dos títulos:

$$\frac{\partial P}{\partial t} - \frac{\partial P}{\partial Y_1} \tilde{\kappa}_1 Y_1(t) - \frac{\partial P}{\partial Y_2} \tilde{\kappa}_2 Y_2(t) + \frac{1}{2} \frac{\partial^2 P}{\partial Y_1^2} \nu_1^2 + \frac{1}{2} \frac{\partial^2 P}{\partial Y_2^2} \nu_2^2 + \frac{\partial^2 P}{\partial Y_1 \partial Y_2} \nu_1 \nu_2 \rho_{12} = r(t) P(t, T)$$
(3.9)

No caso, ρ_{12} é a correlação entre os dois processos de Wiener das equações das variáveis de estado. Sendo assim, consideremos a seguinte solução para o preço dos títulos:

$$P(t,T) = e^{A(\tau) - B_1(\tau)Y_1(t) - B_2(\tau)Y_2(t) - H(t,T)}$$
(3.10)

Em que:

$$H(t,T) = \int_{t}^{T} \delta(v)dv \quad e \quad \tau = T - t \tag{3.11}$$

Tomando a derivada parcial do preço do título e substituindo-a na equação encontrada pelo Lema de Itô, temos as seguintes equações diferenciais ordinárias:

$$\frac{\partial A(\tau)}{\partial \tau} = \frac{1}{2} B_1^2(\tau) \nu_1^2 + \frac{1}{2} B_2^2(\tau) \nu_2^2 + B_1(\tau) B_2(\tau) \nu_1 \nu_2 \rho_{12}$$
(3.12)

$$\frac{\partial B_1(\tau)}{\partial \tau} = 1 - B_1(\tau)\tilde{\kappa}_1 \tag{3.13}$$

$$\frac{\partial B_2(\tau)}{\partial \tau} = 1 - B_2(\tau)\tilde{\kappa_2} \tag{3.14}$$

Essas equações estão sujeitas a $A(0) = B_1(0) = B_2(0) = 0$, e há uma solução fechada para os fatores $A(\tau)$, $B_1(\tau)$ e $B_2(\tau)$, apresentada em Nawalkha, Beliaeva e Soto (2007). Há também uma solução para δ obtida tomando-se a derivada do logaritmo do preço do título na data 0 com relação à maturidade e substituindo as equações diferenciais ordinárias dadas acima para obter:

$$\delta(t) = f(0,t) + \frac{1}{2}B_1^2(t)\nu_1^2 + \frac{1}{2}B_2^2(t)\nu_2^2 + B_1(t)B_2(t)\nu_1\nu_2\rho_{12}$$
(3.15)

A solução para o termo H(t,T) pode ser dada usando a equação de solução do preço dos títulos da seguinte forma:

$$H(t,T) = \int_{t}^{T} \delta(v)dv = H(0,T) - H(0,t) = A(T) - A(t) - \ln P(0,T) + \ln P(0,t)$$
 (3.16)

Em que A(t) e A(T) são dados na solução de forma fechada desse fator e P(0,t) e P(0,T) são os preços iniciais de títulos zero cupom.

Devemos também apresentar o processo do preço do título de acordo com esse modelo. Ele é obtido por meio do Lema de Itô da seguinte forma:

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt - \sum_{i=1}^{i=2} B_i(T-t)\nu_i d\tilde{W}_i(t)$$
 (3.17)

A dinâmica da taxa a termo também é apresentada em Nawalkha, Beliaeva e Soto (2007), assim como a volatilidade da dinâmica dessa taxa.

A principal vantagem dos modelos Gaussianos de dois fatores é que a curva de volatilidade da taxa a termo pode assumir valores mais realistas. Para o processo de apreçamento de opções, esse fator é bastante importante, pois o valor do derivativo depende diretamente da função de volatilidade do ativo objeto.

Especificamente, o apreçamento de uma opção de compra sob o modelo Vasicek de dois fatores, que estamos utilizando para encontrar o valor teórico do derivativo na aplicação deste trabalho, é dada por:

$$c(t) = P(t,T)N(d_1) - KP(t,S)N(d_2)$$
(3.18)

Em que N(x) é a distribuição normal acumulada avaliada em x. As variáveis d_1 e d_2 são dadas por:

$$d_1 = \frac{\ln([P(t,T)]/[P(t,S)K]) + V/2}{\sqrt{V}}$$
(3.19)

$$d_2 = \frac{\ln([P(t,T)]/[P(t,S)K]) - V/2}{\sqrt{V}}$$
(3.20)

No nosso caso, a volatilidade "V" nas expressões acima é a integral da volatilidade do rendimento do título zero cupom calculado por kernel, que será a variância do preço dos títulos.

$$V = \sum_{i=1}^{2} \nu_i^2 \int_t^S (B_i(T-u) - B_i(S-u))^2 du + 2\nu_1 \nu_2 \rho_{12}$$

$$\int_t^S (B_1(T-u) - B_1)(S-u))(B_2(T-u) - B_2(S-u)) du$$
(3.21)

O resultado da volatilidade integrada acima é apresentada também em Nawalkha, Beliaeva e Soto (2007), e tem o seguinte formato:

$$V = \sum_{i=1}^{2} \left(\frac{\nu_i^2}{2} B_i^2(U) B_i(2s) \right) + 2\nu_1 \nu_2 \rho_{12} B_1(U) B_2(U) \left(\frac{1 - e^{-(\tilde{\kappa}_1 + \tilde{\kappa}_2)s}}{\tilde{\kappa}_1 + \tilde{\kappa}_2} \right)$$
(3.22)

Nesse modelo, o preço inicial do título P(0,T) é especificado de forma exógena. Os cinco parâmetros de neutralidade ao risco $\tilde{\kappa}_1$, $\tilde{\kappa}_2$, ν_1 , ν_2 e ρ_{12} são obtidos calibrando o modelo a uma cross section dos preços dos derivativos de taxas de juros. Em Laurini e Ohashi (2013), os parâmetros citados são assumidos como $\kappa_1 = .8$, $\kappa_2 = .7$, $\nu_1 = .1$, $\nu_2 = .1$ e $\rho_{12} = -.3$, que são os valores que assumiremos à aplicação deste modelo em nosso trabalho.

3.3.2 Modelo Cox-Ingersoll-Ross

Esse modelo de estrutura a termo que considera a taxa de curto prazo é o que utilizamos para definir as oscilações da variância em um dos casos da aplicação do nosso modelo. Conforme mencionado na introdução, estamos considerando que a variância do nosso modelo teórico, que é o Vasicek com dois fatores, pode ser constante ou não no tempo. Assim, nossos resultados são comparados entre essas duas situações para mostrar o ajuste das diferentes regressões que utilizamos para encontrar a função de volatilidade da série.

A apresentação matemática desse modelo, exposta em Back (2010), evidencia que o processo gerador de dados considerado para a taxa de juros de curto prazo é um processo de raiz quadrada, que não permite taxas negativas. Assim, como a variância do processo em questão tende a zero quando a taxa diminui, ela se afasta da origem e seu valor é sempre positivo.

Vale lembrar que a já mencionada discretização de Euler também é usada nessa definição das oscilações da variância, já que o modelo CIR é apresentado na forma contínua, e o programa econométrico necessita de *inputs* na forma discreta.

3.4 Heath, Jarrow e Morton

Nessa seção será apresentado o modelo que utilizamos para encontrar a estrutura a termo da taxa de juros, que foi escolhido pelos motivos já explicitados até aqui. Vale lembrar que o modelo HJM é considerado completo para o estudo da estrutra a termo, de acordo com Gluckstern (2001), sendo que os modelos anteriores propostos principalmente por Black Scholes e Merton, Vasicek, CIR e Ho-Lee apresentam algumas restrições ao uso, seja pela possibilidade de gerar taxas de juros negativas, seja pelas suposições acerca da volatilidade dos preços dos títulos. O HJM é, na realidade, uma classe de modelos que abrange diversas especificações para a taxa de juros, conferindo flexibilidade para a

análise.

O trabalho dos autores Heath, Jarrow e Morton (1992) mostra que essa classe de parametrização da estrutura a termo é baseada diretamente nas curvas a termo (forward rate), assumindo que cada curva é gerada por um processo estocástico distinto, o que equivale a assumir que a estrutura a termo completa é um processo estocástico de dimensão infinita.

Assim, de acordo com Munk (2003), a forma mais natural de atingir consistência ao estimar preços é partir de uma estrutura a termo observada e depois modelar a evolução de toda a estrutura a termo de uma maneira que impeça a arbitragem. Essa é a abordagem apresentada por Heath, Jarrow e Morton (HJM). Assumimos um espaço de probabilidade filtrado que, normalmente, é representado por uma quádrupla $(\Omega, \mathbb{F}, (\mathbb{F}_t)_{t \in T}, \mathbb{P})$, em que $(\mathbb{F}_t)_{t \in T}$ é uma filtração contida em \mathbb{F} . No caso, a filtração $(\mathbb{F}_t)_{t \in T}$ seria uma coleção de σ -álgebras do espaço amostral Ω , que satisfaça $\mathbb{F}_s \subseteq \mathbb{F}_t$ para todo s < t em T, assumindo que as condições de regularidade usuais sejam atendidas (FILIPOVIC, 2009).

Conforme Musiela e Marek (2005), no modelo HJM, a taxa de juros a termo medida na data t com maturidade em T deve satisfazer a equação diferencial estocástica:

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t)$$

Em que α e σ são adaptados e W é um movimento Browniano sob uma medida neutra ao risco. O termo "adaptados", nesse contexto, significa que os coeficientes α e σ fazem parte da informação disponível em t.

Nesse contexto, é importante delinear algumas das suposições básicas do modelo. Supondo que a estrutura a termo da taxa de juros no instante zero seja conhecida e representada pela função $T \to f_0^T$, assume-se que, para um "T" fixo, a taxa a termo de maturidade T evolui de acordo com:

$$df_t^T = \alpha(t, T, (f_t^s)_{s \ge t})dt + \sum_{i=1}^n \beta_i(t, T, (f_t^s)_{s \ge t})dz_{it}, 0 \le t \le T$$

em que z_1, \ldots, z_n são n movimentos Brownianos padrão sob a medida de probabilidade real. De acordo com Heath, Jarrow e Morton (1992), atendendo a algumas restrições, nesse processo estocástico, n movimentos Brownianos independentes determinam a flutuação estocástica de toda a curva da taxa a termo começando de uma curva inicial fixa $\{f(0,T): T \in [0,\tau]\}$. A sensibilidade da mudança na maturidade de uma determinada

taxa a termo em relação a cada movimento Browniano é observada diferenciando os coeficientes de volatilidade.

Diferentes especificações para esses coeficientes de volatilidade geram características qualitativas significativamente diferentes do processo da taxa a termo. Nesse contexto, é importante lembrar que a especificação do coeficiente de volatilidade neste trabalho será feita pelo do método de kernel, que é apresentado nas próximas seções.

é possível apresentar também a dinâmica da taxa à vista, com base na dinâmica da taxa a termo:

$$dr_t = \alpha(t, t, (f_t^s)_{s \ge t})dt + \sum_{i=1}^n \beta_i(t, t, (f_t^s)_{s \ge t})dz_{it}, t \in [0, T]$$

Segundo Heath, Jarrow e Morton (1992), a principal diferença da taxa a termo e da taxa à vista é que nessa última os argumentos do tempo e da maturidade variam simultaneamente.

A ideia básica da modelagem da estrutura a termo da taxa de juros é utilizar a função de desconto $T \to B_t^T$, a curva de rendimento $T \to y_t^T$ ou ainda a função da taxa a termo $T \to f_t^T$, de acordo com as seguintes relações:

$$B_t^T = e^{-\int_t^T f_t^s ds} = e^{-y_t^T (T-t)}$$

$$f_t^T = -\frac{\partial ln B_t^T}{\partial T} = y_t^T + (T - t)\frac{\partial y_t^T}{\partial T}$$

$$y_t^T = \frac{1}{T-t} \int_t^T f_t^s ds = -\frac{1}{T-t} ln B_t^T$$

O método em questão faz uso da taxa a termo como instrumento de modelagem, já que o preço dos títulos zero cupom e as curvas de rendimento envolvem somas e integrais de taxas a termo. Sabe-se também que uma das formas de precificar derivativos é encontrar o retorno esperado descontado sob uma medida martingale à vista, em que o desconto é feito em termos da taxa de juros de curto prazo. A relação da taxa de curto prazo com a taxa a termo, a curva de rendimento e a função de desconto é dada por:

$$r_t = f_t^t = \lim_{T \to t} y_t^T = -\lim_{T \to t} \frac{\partial ln B_t^T}{\partial T}$$

Outro motivo para se escolher a taxa a termo como instrumento de modelagem é a ausência de necessidade de se impor determinadas restrições à estrutura de volatilidade, que seriam essenciais caso utilizássemos os preços dos títulos zero cupom (a volatilidade desses preços deve ser próxima de zero quando eles são próximos de um, por exemplo, que é o valor de face do título). Assim, podemos utilizar a suposição de volatilidade das taxas a termo como constante no tempo. A variância dos preços dos títulos, por sua vez, nada mais é que a integral da função de volatilidade das taxas. Como realizamos a suposição de variância das taxas constante no tempo, sua integral em relação à maturidade resulta em $\sigma_f(T-t)$, que é igual a σ_p .

3.5 Dinâmica do preço dos títulos e a restrição de deslocamento

Retomando o tema da mudança de probabilidade, pode-se mostrar matematicamente como é possível realizar a mudança para a medida *martingale* à vista, com base na teoria apresentada em Munk (2003).

Sob a dinâmica da taxa a termo imposta pelo modelo HJM, a dinâmica dos preços dos títulos zero cupom será:

$$dB_t^T = B_t^T [\mu^T (t, (f_t^s)_{s \ge t}) dt + \sum_{i=1}^n \sigma_i^T (t, (f_t^s)_{s \ge t}) dz_{it}]$$

em que

$$\mu^{T}(t, (f_{t}^{s})_{s \ge t}) = r_{t} - \int_{t}^{T} \alpha(t, u, (f_{t}^{s})_{s \ge t}) du + \frac{1}{2} \sum_{i=1}^{n} \left(\int_{t}^{T} \beta_{i}(t, u, (f_{t}^{s})_{s \ge t}) du \right)^{2}$$

e

$$\sigma_i^T(t, (f_t^s)_{s \ge t}) = -\int_t^T \beta_i(t, u, (f_t^s)_{s \ge t}) du$$

Para provar essa relação, é necessário introduzir um processo estocástico auxiliar:

$$Y_t = \int_t^T f_t^u du$$

Como os títulos são continuamente capitalizados, os preços serão: $B_t^T=e^{-Yt}$. Encontrando a dinâmica da curva de rendimento pela regra de Leibinitz para integrais esto-

cásticas, temos:

$$dY_t = \left(-r_t + \int_t^T \alpha(t, u, (f_t^s)_{s \ge t}) du\right) dt + \left(\int_t^T \beta(t, u, (f_t^s)_{s \ge t}) du\right) dz_t$$

Pelo Lema de Itô, sabendo que $B_t^T = g(Y_t)$ e $g(Y) = e^{-Y}$, a dinâmica do preço do título é encontrada.

Já se considerarmos a medida de probabilidade martingale à vista \mathbb{Q} , a taxa a termo terá os mesmos termos de volatilidade $\beta_i(t,T,(f_t^s)_{s\geq t})$, mas deslocamentos diferentes. O processo $z^{\mathbb{Q}}$ de dimensão n definido por

$$dz_{it}^{\mathbb{Q}} = dz_{it} + \lambda_{it}dt$$

é um movimento Browniano padrão sob a medida \mathbb{Q} , em que os processos λ_i são os preços de mercado do risco. A dinâmica da taxa a termo será:

$$df_t^T = \hat{\alpha}(t, T, (f_t^s)_{s \ge t})dt + \sum_{i=1}^n \beta_i(t, T, (f_t^s)_{s \ge t})dz_{it}^{\mathbb{Q}}$$

em que
$$\hat{\alpha}(t, T, (f_t^s)_{s \ge t}) = \alpha(t, T, (f_t^s)_{s \ge t}) - \sum_{i=1}^n \beta_i(t, T, (f_t^s)_{s \ge t}) \lambda_{it}$$
.

Com essa informação, podemos encontrar a dinâmica dos preços dos títulos zero cupom (o deslocamento da taxa). Sabendo que o deslocamento é igual a r_t (taxa à vista), devemos ter que

$$\int_{t}^{T} \hat{\alpha}(t, u, (f_{t}^{s})_{s \ge t}) du = \frac{1}{2} \sum_{i=1}^{n} \left(\int_{t}^{T} \beta_{i}(t, u, (f_{t}^{s})_{s \ge t}) du \right)^{2}$$

Diferenciando com relação a T, resulta:

$$\hat{\alpha}(t, T, (f_t^s)_{s \ge t}) = \sum_{i=1}^n \beta_i(t, T, (f_t^s)_{s \ge t}) \int_t^T \beta_i(t, u, (f_t^s)_{s \ge t}) du$$

Essa relação é chamada restrição de deslocamento HJM. Essa restrição apresenta algumas consequências importantes:

1. A taxa a termo não precisa ser especificada de forma exógena, já que o comportamento da taxa é totalmente caracterizado pela curva inicial, o número de fatores e os termos de volatilidade β_i .

2. Como preços de derivativos dependem da evolução da estrutura a termo da taxa de juros sob a medida martingale à vista, esses preços dependem apenas da curva da taxa a termo inicial e das funções de volatilidade $\beta_i(t, T, (f_t^s)_{s \ge t})$.

3.6 Modelo HJM Gaussiano

Segundo Musiela e Marek (2005), se a função de volatilidade σ for determinista, a classe do modelo HJM que deve ser utilizada é a Gaussiana, já que o resultado da avaliação de derivativos que dependem da taxa de juros será, nesse caso, independente do prêmio de risco. Modelos Gaussianos são aqueles em que os valores futuros das taxas a termo são normalmente distribuídos. De acordo com Munk (2003), uma desvantagem dessa classe de modelos é que eles permitem taxas de juros reais negativas, mas são modelos bastante tratáveis, em especial permitindo fórmulas analíticas para o apreçamento de títulos e derivativos.

Um modelo HJM é Gaussiano se a volatilidade das taxas a termo β_i são funções deterministas do tempo e maturidade:

$$\beta_i(t, T, (f_t^s)_{s > t}) = \beta_i(t, T), i = 1, 2, \dots, n$$

Para observar esse fato, note que o deslocamento da taxa a termo sob a medida martingale à vista \mathbb{Q} é também uma função determinista do tempo e da maturidade:

$$\hat{\alpha}(t,T) = \sum_{i=1}^{n} \beta_i(t,T) \int_{t}^{T} \beta_i(t,u) du$$

Assim, para qualquer T, as taxas a termo de maturidade T evoluem de acordo com:

$$f_t^T = f_0^T + \int_0^t \hat{\alpha}(u, T) du + \sum_{i=1}^n \int_0^t \beta_i(u, T) dz_{iu}^{\mathbb{Q}}$$

Como $\beta_i(u,T)$ depende apenas do tempo, as integrais estocásticas são normalmente distribuídas, então as taxas a termo futuras são normalmente distribuídas sob \mathbb{Q} . Já a dinâmica da taxa de juros de curto prazo, $r_t = f_t^t$ será:

$$r_{t} = f_{0}^{t} + \int_{0}^{t} \hat{\alpha}(u, t)du + \sum_{i=1}^{n} \int_{0}^{t} \beta_{i}(u, t)dz_{iu}^{\mathbb{Q}}, 0 \le t$$
(3.23)

que é também normalmente distribuída sob \mathbb{Q} .

Com relação às especificações consideradas, conforme Musiela e Marek (2005), o modelo HJM Gaussiano considera que, em qualquer modelo de estrutura a termo, a volatilidade de qualquer ativo objeto seja determinista, ou seja, ela não será afetada por choques aleatórios. Uma das principais vantagens do modelo HJM Gaussiano, de acordo com Brace e Musiela (1994), é o cálculo confiável do preço de derivativos at-the-money por meio de fórmulas que podem ser derivadas matematicamente e programadas de forma simples.

Como discutido anteriormente, o uso de uma estrutura de volatilidade determinista, implicando em um modelo HJM Gaussiano, é uma simplificação importante na análise de derivativos. Embora não permitam a maior generalidade possível nesta classe de modelos, são computacionalmente simples e evitam o uso de métodos de árvores não recombinantes ou simulações de Monte Carlo no apreçamento de derivativos, sendo adequadas aos propósitos práticos de mercado. Essa formulação também é importante já que permite a adoção natural do método de estimação não paramétrico para a função de volatilidade discutida a seguir.

4 Abordagem teórica da estimação pelo Método de kernel

A ideia desse capítulo é apresentar a teoria que embasa a aplicação do método de kernel nesse trabalho. Conforme já discutido, utilizamos essa técnica para estimar as funções de volatilidade da série da taxa de juros. No processo de apreçamento de opções que realizamos, a função de volatilidade da série é o *input* principal, já que o estudo de Heath, Jarrow e Morton (1992) mostra que os deslocamentos do processo da taxa a termo dependem somente da estrutura inicial e da volatilidade da taxa. Sendo assim, fazemos uma exposição acerca da metologia de kernel, incluindo as técnicas de seleção de "janela" e a teoria de regressões de kernel. Vale lembrar que a referida "janela" diz respeito ao parâmetro de suavização, ou seja, ao peso que se atribui a cada observação na estruturação da regressão.

Estruturamos o presente capítulo de forma a apresentar primeiramente a teoria de kernel na estimação de funções densidade de probabilidade e seus respectivos procedimentos de seleção de janela que, apesar de não ser diretamente utilizado na aplicação do nosso modelo proposto, é útil para entender o funcionamento da técnica em questão. Logo após, na seção 4.2, é apresentada a teoria sobre as regressões não paramétricas de kernel, que é a técnica que utilizamos para estimar as funções de volatilidade das curvas de juros, tanto na seção sobre as simulações das curvas e apreçamento de opções sobre títulos zero cupom quanto na seção empírica, em que realizamos o apreçamento de opções sobre o IDI.

Assim, com relação à forma funcional especificada, essa técnica de kernel que utilizamos é não paramétrica. Os métodos não paramétricos têm sido bastante utilizados na literatura recente, uma vez que mesmo a teoria econômica estando bem avançada em determinadas áreas, ela muitas vezes não oferece uma forma funcional que possa ser usada para a modelagem estatística.

Há diversos métodos não paramétricos, mas o que utilizamos nesse trabalho são as

técnicas de "suavização", que são métodos que podem ser considerados como "descritivos". é importante primeiramente discorrer a respeito das propriedades da distribuição empírica, antes de estudar a estimação da densidade das distribuições.

De acordo com Florens, Marimoutou e Peguin-Feissolle (2007), o método básico da estatística não paramétrica são os modelos i.i.d com distribuição de probabilidade desconhecida. A ideia desse modelo é que as observações são elementos do \mathbb{R}^m gerados independentemente por uma distribuição de probabilidade desconhecida Q. Nesse contexto, o estimador "natural" de Q é a distribuição empírica definida como:

$$\hat{Q}_n = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$$

Em que δ_{x_i} é uma medida Dirac em x_i . Lembrando que se $E \subset \mathbb{R}^m$, $\delta_{x_i}(E)$ é igual a 1 se $x_i \in E$ e igual a 0 caso contrário. Esse modelo é desenvolvido sem impor restrições sobre a distribuição de probabilidade Q. No entanto, a estimação por kernel é feita impondo-se a suposição de que Q admite a seguinte densidade f:

$$Q(E) = \int_{E} f(u)du$$

O objetivo nesse modelo é estimar f, ou seja, a densidade. Para realizar esse procedimento, é possível recorrer a diversas técnicas, tais como desenhar o histograma das observações ou utilizar o método de kernel. Há também estimações por funções ortogonais, funções spline, dentre outras. A técnica de spline cúbico é usada na aplicação do nosso modelo, no sentido de encontrar uma função de volatilidade diária a partir dos pontos estimados pela função de kernel.

De acordo com Racine e Li (2007), a grande desvantagem de métodos paramétricos de estimação, como a máxima verossimilhança, por exemplo, é a necessidade de se supor uma distribuição de probabilidade a priori para estimar os parâmetros desconhecidos. Assim, para estimar a função densidade de probabilidade (fdp) sem a necessidade de realizar pressupostos paramétricos sobre a sua forma funcional, pode-se partir da definição de fdp:

$$f(x) = \frac{d}{dx}F(x)$$

Sabendo que a função de distribuição acumulada (fda) pode ser estimada como o

número de observações de dada variável menores ou iguais a um valor pré-determinado dividido pelo número de observações, um estimador para a fdp pode ser representado como:

$$\hat{f}(x) = \frac{F_n(x+h) - F_n(x-h)}{2h} \tag{4.1}$$

em que h é um incremento positivo pequeno.

Definindo a função de kernel uniforme como dada por:

$$k(z) = \begin{cases} \frac{1}{2} se |z| \le 1\\ 0 c.c. \end{cases}$$

Então, a fdp em 4.1 também pode ser expressa como:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} k\left(\frac{X_i - x}{h}\right) \tag{4.2}$$

O termo "h" na equação acima é a chamada janela do kernel, ou parâmetro de suavização. Essa equação mostra o estimador de kernel uniforme, já que ele corresponde a uma fdp uniforme, mas existem diversas outras funções de kernel $k(\cdot)$ que podem ser utilizadas como, por exemplo, a função normal padrão.

Conforme provado em Racine e Li (2007), a função $\hat{f}(x)$ definida em 4.2 é um estimador consistente da fdp desde que a função kernel atenda às seguintes condições:

- $1. \int k(v)dv = 1$
- 2. k(v) = k(-v)
- $3. \int v^2 k(v) dv = \kappa_2 > 0$

Para mensurar a proximidade do estimador da função à função real, é preciso observar o valor do erro quadrático médio (EQM), e não apenas da variância, pois o estimador pode ser viesado. Temos que:

$$EQM\left(\hat{f}(x)\right) \equiv E\left\{\left[\hat{f}(x) - f(x)\right]^{2}\right\}$$

Essa medida tende para zero quando $k\to\infty$ e $k/n\to0$. Mas vale lembrar que o EQM é computado em apenas um ponto da distribuição analisada, sendo que uma forma

mais acurada de se analisar o erro do estimador é por meio do EQM integrado (EQMI), ou seja:

$$EQMI(\hat{f}) \equiv \int E\left[\hat{f}(x) - f(x)\right]^2 dx$$

Quando X é uniformemente distribuído, é possível mostrar que o parâmetro de suavização ótimo " h_{ot} " apresenta uma taxa de convergência igual a $n^{-1/3}$.

4.1 A Seleção do Parâmetro de Suavização Univariado

O parâmetro de suavização ótimo, que nada mais é que o valor que será atribuído ao parâmetro que dá peso às diferentes observações na estimação, depende da integral da segunda derivada da densidade desconhecida. Racine e Li (2007) mostram que esse parâmetro é obtido a partir da relação

$$h_{ot} = \kappa^{-2/5} \kappa^{1/5} \left\{ \int \left[f^{(2)}(x) \right]^2 dx \right\}^{-1/5} > 0 \tag{4.3}$$

Uma maneira de encontrar h_{ot} é estabelecer um "valor piloto" de h e calcular de forma não paramétrica $\int \left[f^{(2)}(x)\right]^2 dx$, para posteriormente usar esse valor para obter h_{ot} . O mais comum é supor que f(x) seja de uma família paramétrica de distribuições, e depois computar h usando a relação 4.3, encontrando o referido "valor piloto". Assim, calcula-se $\int \left[\hat{f}^{(2)}(x)\right]^2 dx$, que será usado para encontrar h_{ot} por meio da equação 4.3. Também é comum na literatura usar o próprio "valor piloto" como a janela do modelo.

Há distintos métodos de seleção desse parâmetro de acordo com a teoria, sendo que utilizamos nesse trabalho a técnica conhecida como validação cruzada por mínimos quadrados e a técnica de AIC_c , que é o critério de Akaike modificado para a utilização em métodos não paramétricos. A seguir, apresentamos a validação cruzada por mínimos quadrados e a validação cruzada por máxima verossimilhança, que são normalmente utilizadas na aplicação de kernel à estimação da função densidade de probabilidade. Em nosso caso específico, em que realizamos regressões de kernel para estimar a função de volatilidade da curva de juros, nosso método de seleção de janela é apresentado nas seções 4.3.2 e 4.3.3 a seguir.

4.1.1 Seleção do parâmetro de suavização: métodos de validação cruzada

A escolha de uma janela para a estimação por kernel pode ser associada à escolha do número de termos na aproximação de uma série. Se o número de termos é grande (ou o k é pequeno), o modelo resultante é mais flexível. Mas o aumento da flexibilidade (redução do potencial viés) leva a um aumento na variabilidade. Assim, é necessário analisar o EQM da estimação resultante para tomar uma decisão.

4.1.2 Validação Cruzada por Mínimos Quadrados

Esse método é baseado no princípio de seleção da janela que minimiza o EQMI da estimação resultante. Lembrando que esse método de seleção do parâmetro que estamos abordando nessa seção é aplicável à estimação da função densidade de probabilidade por kernel, que não é diretamente utilizado na aplicação do modelo proposto. A ideia de minimização do EQMI, porém, também é utilizada na seleção de janela das regressões de kernel, que são utilizadas na estimação das funções de volatilidade das curvas de juros. Assim, é importante realizar a exposição dessa técnica para entender o procedimento utilizado na seleção da janela do nosso modelo proposto.

O erro quadrático médio integrado, ou EQMI, é expresso por:

$$\int \left[\hat{f}(x) - f(x)\right]^2 dx = \int \hat{f}(x)^2 dx - 2 \int \hat{f}(x)f(x)dx + \int f(x)^2 dx$$

Como o último termo não depende de h, o objetivo é minimizar os dois primeiros termos com relação a h. Conforme apresentado em Racine e Li (2007), essa minimização se dá a partir da relação abaixo:

$$CV_f(h) = \frac{1}{n^2 h} \sum_{i=1}^n \sum_{j=1}^n \bar{k} \left(\frac{X_i - X_j}{h} \right) - \frac{2}{n(n-1)h} \sum_{i=1}^n \sum_{j \neq i, j=1}^n k \left(\frac{X_i - X_j}{h} \right)$$

e normalmente se realizam algoritmos de busca numéricos.

Dado que $CV_f(h) = CV_{f_0} + s.o.$, em que s.o. significa termos de ordem menor e que não são relacionados com "h". Pode ser mostrado que $\hat{h} = h^0 + o_p(h^o)$, ou seja,

$$\frac{\hat{h} - h^0}{h^0} \equiv \frac{\hat{h}}{h^0} - 1 \to 0$$

em que a seta acima indica convergência em probabilidade, \hat{h} é o valor de h que minimiza $CV_f(h)$ e h^0 é o valor que minimiza CV_{f_0} .

Racine e Li (2007) apresenta também a seguinte convergência para o método de validação cruzada por mínimos quadrados:

$$\frac{\int \left[\hat{f}(x,\hat{h}) - f(x)\right]^2 dx}{\inf_h \int \left[\hat{f}(x,h) - f(x)\right]^2 dx} \to 1$$

em que $\hat{f}(x, \hat{h})$ denota o estimador de kernel de f(x) com \hat{h} selecionado por validação cruzada e $\hat{f}(x, h)$ é o estimador de kernel com h genérico.

4.1.3 Validação Cruzada por Verossimilhança

Esse método é menos utilizado que a validação cruzada por mínimos quadrados, dado que pode apresentar problemas nas caudas da distribuição. Assim, não utilizamos essa técnica diretamente na aplicação do modelo proposto, mas apresentamos o método como referência de mais uma forma de selecionar o parâmetro de suavização de kernel. A escolha de h, nesse caso, busca maximizar a função de log verossimilhança dada por:

$$L = lnL = \sum_{i=1}^{n} ln\hat{f}_{-i}(X_i)$$

em que $\hat{f}_{-i}(X_i)$ é o estimador leave-one-out de kernel de $f(X_i)$ definido como $\hat{f}_{-i}(X_i) = \frac{1}{(n-1)h} \sum_{j=1, j \neq i}^{n} k\left(\frac{X_i - X_j}{h}\right)$.

A ideia é, portanto, utilizar a função densidade de probabilidade estimada pelo método de kernel para aplicar o conceito de máxima verossimilhança. Como a função de verossimilhança depende de h, conforme mencionado, escolhe-se o parâmetro de forma a maximizar a função.

4.2 Regressão não paramétrica

Conforme mencionado, a estimação da função de volatilidade das curvas de juros realizadas na aplicação do modelo proposto é conduzida pelo método de kernel, por meio de regressões não paramétricas. São essas funções de volatilidade que utilizamos como *input* para o apreçamento das opções em questão, sejam as opções sobre os títulos zero cupom, sejam as opções sobre o IDI. Nesse contexto, é importante discorrer sobre os métodos de regressão paramétricos e os não paramétricos para justificar a escolha do modelo não paramétrico neste trabalho. Sabe-se que a forma funcional verdadeira da regressão quase nunca é conhecida, sendo essa uma desvantagem do método paramétrico, que exige que se especifique a forma paramétrica do modelo exata antes da estimação. Se a função de média condicional, a possível heterocedasticidade ou a função de correlação serial forem incorretamente especificadas, podem surgir problemas de parâmetros estimados inconsistentes e inferência inválida.

Assim, o método de regressão não paramétrico pode ser usado para evitar má especificação da forma funcional da regressão, mas cabe ressaltar que esse método exige um número maior de observações para atingir o mesmo grau de precisão que uma regressão paramétrica corretamente especificada.

4.2.1 Estimação de kernel constante local

De acordo com Hansen (2009), a regressão de kernel constante local, também chamada de regressão de Nadaraya-Watson, consiste em estimar um parâmetro y dados os valores de uma variável $X_i = x$, que chamamos de função g(x). Uma forma de realizar essa estimação é considerar uma vizinhança do ponto x, observando que X_i apresenta uma probabilidade positiva de assumir o valor x. Deve-se observar o número de observações nessa vizinhança (que será crescente com a amostra), e extrair a média dos valores de y_i para esses dados. Um aspecto importante é determinar o tamanho da vizinhança de modo a otimizar o trade off entre a variação na função g(x) (viés da estimação) contra o número de observações na vizinhança (variância da estimação).

Nesse contexto, consideremos o seguinte modelo de regressão não paramétrico:

$$y_i = g(x_i) + u_i, \quad i = 1, \dots, n$$

em que $g(\cdot)$ é a função desconhecida a que nos referimos acima. Em Racine e Li (2007)

é provado que E(Y/X) é o preditor ótimo de Y dado X, uma vez que o erro quadrático médio (EQM) desse preditor é o menor possível:

$$E\{[Y - r(X)]^2\} \ge E\{[Y - E(Y|X)]^2\}$$

em que r(X) é qualquer outro preditor de Y dado X.

Sabe-se da definição de esperança de variáveis contínuas que:

$$E(Y|X=x) = \int y f_{y|x}(y|x) dy = \frac{\int y f_{y,x}(x,y) dy}{f(x)} = g(x)$$

Assim, é necessário estimar $f_{y,x}(x,y)$ e f(x) de acordo com a estimação de densidade de kernel já apresentada neste trabalho.

Conforme apresentado em Racine e Li (2007), a estimação de g(x) será:

$$\hat{g}(x) = \frac{\int y.\hat{f}_{y,x}(x,y)dy}{\hat{f}(x)} = \frac{\sum_{i=1}^{n} Y_i.K\left(\frac{X_i-x}{h}\right)}{\sum_{i=1}^{n} K\left(\frac{X_i-x}{h}\right)}$$

que nada mais é do que uma média ponderada de Y_i :

$$\hat{g}(x) = \sum_{i=1}^{n} Y_i \cdot \omega_i$$

É apresentada também em Racine e Li (2007) a prova para a seguinte relação:

$$\hat{g}(x) - g(x) = O_p \left(\sum_{s=1}^q h_s^2 + (nh_1 \dots h_q)^{-1/2} \right)$$

da qual depreendemos que se cada janela h_s tiver a mesma ordem de magnitude, a escolha ótima que minimiza $EQM[\hat{g}(x)]$ é $h_s \sim n^{-1/(q+4)}$.

Pode-se estabelecer também a normalidade assintótica de $\hat{g}(x)$ por meio do Teorema Central do Limite (TCL) de Liapunov. A convergência é dada por:

$$\sqrt{n \cdot h_1 \dots h_q} \left(\hat{g}(x) - g(x) - \sum_{s=1}^q h_s^2 B_s(x) \right) \stackrel{D}{\to} N \left(0, \kappa^q \sigma^2(x) / f(x) \right)$$

em que
$$B_s(x) = \frac{\kappa_2}{2} \left\{ 2f_s(x)g_s(x) + f(x)g_{ss}(x) \right\} / f(x).$$

4.3 Seleção de janela constante local

Como se sabe, o método de kernel baseia-se na ponderação dos valores observados da variável Y_i . Assim, é necessário utilizar um parâmetro de ponderação, ou suavização, para realizar a estimação. Portanto, nessa seção discutimos três principais métodos de seleção do parâmetro de suavização:

- 1. Método Rule-of-thumb e Plug-in
- 2. Método de validação cruzada de mínimos quadrados
- 3. Método AIC_c , ou AIC modificado

Lembrando que os métodos que utilizamos neste trabalho são os dois últimos enumerados acima. O método Rule-of-Thumb é computacionalmente simples, mas ele trata as covariadas simetricamente, sem diferenciar distintos componentes que causam mudanças na função de regressão g(x). Já o método Plug-in pode ser computacionalmente bastante intenso, sendo que decidimos por utilizar os outros métodos de seleção de janela.

4.3.1 Métodos Rule-of-thumb e Plug-in

O método Rule-of-thumb consiste em escolher o parâmetro h_s como $c_s X_{s,sd}$ $n^{-1/(4+q)}$ com (s = 1, ..., q), em que c_s é uma constante e $X_{s,sd}$ é o desvio padrão amostral de $\{X_{is}\}_{i=1}^n$. Essa escolha se deve à ordem do parâmetro ótimo de suavização, que é $O(n^{-1/(4+q)})$.

Esse método apresenta a vantagem de ser simples (computacionalmente), mas ele trata todos os componentes (covariadas) de x simetricamente. Isso significa que esse método não apresenta flexibilidade para tratar diferentemente um componente x_1 e um x_2 , por exemplo, que podem representar componentes que causam mudanças na função de regressão g(x).

Outro método utilizado é o *Plug-in*, que busca minimizar o EQM ponderado integrado, que apresenta o formato $\int E[\hat{g}(x) - g(x)]^2 V(x) dx$, em que V(x) é uma função de ponderação não negativa que garante que, assintoticamente, EQMPI seja finito.

Nesse método de escolha do parâmetro, pode-se estimar o termo a_s^0 , que consta na expressão de "h":

$$h_s^0 = n^{-1/(q+4)} \cdot a_s^0$$
, com $s = 1, \dots, q$

Lembrando que essa expressão apresenta o "h" que minimiza o EQMPI.

Essa estimação do termo a_s^0 pode ser computacionalmente intensa. Porém, mesmo que exista uma forma fechada para a expressão de a_s^0 , ainda serão necessárias estimações não paramétricas para componentes deste termo, o que complica o uso desse método.

Normalmente, o método *Plug-in* não é usado em aplicações empíricas não paramétricas multivariadas, como discutido em Racine e Li (2007) e, assim, usamos outros métodos em nosso trabalho.

4.3.2 Validação cruzada de mínimos quadrados

A ideia desse método de escolha do parâmetro de suavização é escolher h_1, \ldots, h_q de modo a minimizar a função objetivo:

$$CV_{cl}(h_1, \dots, h_q) = n^{-1} \sum_{i=1}^n (Y_i - \hat{g}_{-i}(X_i))^2 M(X_i)$$
 (4.4)

em que $\hat{g}_{-i}(X_i) = \sum_{l \neq i}^n Y_l K((X_i - X_l)/h) / \sum_{l \neq i}^n K((X_i - X_l)/h)$ é o estimador de kernel leave-one-out de $g(X_i)$ e $0 \leq M(\cdot) \leq 1$ é uma função de peso.

Vale lembrar que o kernel *leave-one-out* consiste em realizar a estimação por kernel na amostra excluindo uma observação de cada vez, de modo que se realizem tantas estimações de kernel quantas forem as observações da amostra.

A ideia desse método é prever a amostra excluída pelo modelo estimado e comparar esta previsão com o valor real da observação, computando-se a diferença. Deve-se buscar minimizar o viés encontrado e ajustar a variância correta do real processo gerador de dados, de modo que não haja "sobre suavização" ou "sub suavização" da curva estimada.

Para pesquisadores da área aplicada, os métodos de validação cruzada são interessantes já que não é necessário conhecer as expressões explícitas para os parâmetros de suavização ótimos, sendo suficiente minimizar a função objetivo definida em 4.4, que pode ser feito utilizando um procedimento de otimização numérica padrão.

4.3.3 Critério AIC_c para seleção de janela

Esse método de seleção de janela de kernel é um dos que utilizamos na parte aplicada do nosso modelo proposto, sendo que foram apresentadas suas propriedades em amostras finitas no trabalho original que propôs o método, de Hurvich, Simonoff e Tsai (1998). A ideia da técnica é modificar o critério de informação de Akaike no sentido de utilizá-lo para métodos não paramétricos. A abordagem de Hurvich, Simonoff e Tsai (1998) é válida para estimadores que podem ser escritos como combinações lineares de outras variáveis e, portanto, é diretamente aplicável a uma vasta gama de estimadores não paramétricos.

O critério de informação em questão é apresentado em Racine e Li (2007) na seguinte forma:

$$AIC_c = \ln(\hat{\sigma}^2) + \frac{1 + tr(H)/n}{1 - \{tr(H) + 2\}/n}$$
(4.5)

com

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \left\{ Y_i - \hat{g}(X_i)^2 \right\}$$
 (4.6)

Sabe-se que o trabalho de Hurvich, Simonoff e Tsai (1998) desenvolve simulações para mostrar que essa técnica de seleção de janela funciona bem quando comparada ao método plug-in já apresentado e com uma série de métodos de validação cruzada generalizados.

Na parte aplicada do nosso trabalho, além dessa técnica de seleção de janela, utilizamos também a validação cruzada por mínimos quadrados. Os autores Hurvich, Simonoff e Tsai (1998) realizaram testes e mostraram que esses dois métodos apresentam as mesmas propriedades assintóticas, fato que foi corroborado por Li e Racine (2004). Esse último trabalho inclusive mostrou que, para amostras pequenas, o critério AIC_c normalmente apresenta resultados melhores que a validação cruzada por mínimos quadrados.

4.4 Seleção de janela linear local: validação cruzada por mínimos quadrados

A principal diferença da estimação de kernel constante local e a estimação linear local é que a primeira pode apresentar problemas de grande viés quando estima-se a função de regressão próximo ao limite (fronteira). Já o estimador linear local apresenta uma

correção para esse problema, dado que o viés nesse caso não é função da densidade f(x).

Dada essa diferença, faremos nessa seção a apresentação da seleção da janela desse outro tipo de estimação de kernel pela validação cruzada por mínimos quadrados.

Seja $\hat{g}_{i,L}(X_i)$ o estimador linear local leave-one-out. O par (\hat{a}_i, \hat{b}_i) é a solução de (a, b) no seguinte problema de minimização:

$$\min_{\{a,b\}} \sum_{j \neq i, j=1}^{n} \left[Y_j - a - (X_j - X_i)' b \right]^2 K \left(\frac{X_i - X_j}{h} \right)$$

Então $\hat{a}_i \equiv \hat{g}_{-i,L}(X_i)$ é o estimador de kernel linear local leave-one-out de $g(X_i)$.

A ideia da validação cruzada para a seleção da janela do kernel nesse caso é escolher os valores de h_s que minimizam:

$$CV_{ll}(h_1, ..., h_q) = \min_{h} \frac{1}{n} \sum_{i} [Y_i - \hat{g}_{-i,L}(X_i)]^2 M(X_i)$$

em que $M(\cdot)$ é uma função de ponderação. Os autores Li e Racine apresentam que o parâmetro de suavização linear local escolhido pela validação cruzada converge para o parâmetro de suavização ótimo, e a taxa de convergência é a mesma que no caso da validação cruzada constante local. A convergência é dada por:

$$n^{1/(q+4)}\hat{h}_s \stackrel{P}{\to} a_{s,l}^0$$

4.5 Regressão polinomial local - o caso univariado

Pode-se estimar também g(x) por meio de um estimador polinomial local. No caso deste trabalho, estamos estimando uma função g(x) de volatilidade, em que a variável x representa os regressores que serão apresentados no próximo capítulo. Quando X é univariado, um estimador de kernel polinomial local de p-ésima ordem é baseado na minimização da seguinte função objetivo:

$$\min_{\{b_0,b_1,\dots,b_q\}} \sum_{j=1}^n (Y_j - b_0 - b_1(X_j - x) - \dots - b_p(X_j - x)^p)^2 \cdot K\left(\frac{X_j - x}{h}\right)$$

Sejam \hat{b}_l os valores de b_l $(l=0,1,\ldots,p)$ que minimizam a equação acima. Então \hat{b}_0

estima g(x) e $l!\hat{b}_l$ estima $g^{(l)}(x)$, em que $g^{(l)}(x) = \frac{d^l g(x)}{dx^l}$ é a derivada de l-ésima ordem de g(x) com $l = 1, \ldots, p$.

Em nosso caso, o cálculo das derivadas não é necessário, já que não há restrições específicas acerca da função de volatilidade a ser estimada. Uma restrição que pode ser aplicada à função em questão é a de positividade, já que trata-se de uma variância. Vale lembrar que nas análises de simulação e na análise empirica que desenvolvemos, o método de estimador polinomial local e o método local constante apresentam os mesmos resultados, e por isso apresentamos somente os resultados do método local constante.

É apresentado também em Racine e Li (2007) que pode-se estabelecer a normalidade assintótica de $\hat{g}(x)$ e $\hat{g}^{(l)}(x)$ por meio do TCL de Liapunov.

4.6 O caso multivariado

Esse caso se aplica quando há mais de um fator modelado que afeta a série em questão. No caso desse trabalho, como utilizamos cinco regressões distintas para encontrar a função de volatilidade da série, estamos abordando justamente o caso multivariado. Assim, observamos nessa seção algumas características acerca da função g(x), já apresentada nas seções anteriores.

O estimador polinomial local de p-ésima ordem é apresentado em Racine e Li (2007), em que se mostra a taxa de convergência quase-certa uniforme e o resultado assintótico de normalidade para o estimador polinomial local de g(x) e suas derivadas até a ordem "p".

A notação utilizada por Racine e Li (2007) é:

$$r = (r_1, \dots, r_q), \quad r! = r_1! \cdot \dots \cdot r_q!, \quad \bar{r} = \sum_{j=1}^q r_j$$

$$x^r = x_1^{r_1} \cdot \ldots \cdot x_q^{r_q}, \quad \sum_{0 \le \bar{r} \le p} = \sum_{j=0}^p \sum_{r_1=0}^j \ldots \sum_{r_q=0}^j$$

em que $\bar{r} \equiv r_1 + \ldots + r_q \equiv j$

Além disso,
$$(D^r g)(x) = \frac{\partial^r g(x)}{\partial x_1^{r_1} ... \partial x_q^{r_q}}$$
.

Usando essa notação, pode-se aproximar g(z) localmente usando um polinômio multivariado de ordem p dado por:

$$g(z) \cong \sum_{0 \le \bar{r} \le p} \frac{1}{r!} (D^r) g(v)|_{v=x} (z-x)^r$$
 (4.7)

Pode-se definir uma função multivariada de mínimos quadrados ponderados por:

$$\sum_{i=1}^{n} \left\{ Y_i - \sum_{0 \le \bar{r} \le p} b_r(x) (X_i - x)^r \right\}^2 K\left(\frac{X_i - x}{h}\right)$$
 (4.8)

Minimizando 4.8 com relação a cada b_r fornece uma estimativa $\hat{b}_r(x)$, e por 4.7 sabe-se que $r!\hat{b}_r(x)$ estima $(D^rg)(x)$ tal que $(D^r\hat{g})(x) = r!\hat{b}_r(x)$. Assim, essa minimização fornece um estimador de kernel polinomial local de p-ésima ordem, ou seja, esse é o procedimento de estimação das próprias regressões de kernel que utilizamos no capítulo seguinte, sobre a aplicação do modelo proposto.

A estimação da função de volatilidade para uma certa maturidade é realizada através da seguinte especificação, seguindo Jeffrey et al. (2004):

$$\tilde{\gamma}^2(x,\tau) = \frac{1}{\Delta} \frac{\sum_{i=1}^n K_{h_x} (X_{t_i} - x) (y_{t_{i+1}}(\tau) - y_{t_i}(\tau))^2}{\sum_{i=1}^n K_{h_x} (X_{t_i} - x)}$$
(4.9)

Assim, a ideia geral é estimar a função de variância fazendo uma regressão não paramétrica dos desvios quadráticos das taxas observadas y_t em relação a sua média, que é a proxy da variância não observada da taxa de juros, em relação a um conjunto de variáveis explicativas X_{t_i} , e obter a função de volatilidade como a raiz desta estimativa.

$egin{array}{lll} 5 & Apreçamento ~n\~{a}o ~param\'{e}trico ~de \ &op\~{c}\~{o}es \end{array}$

Nesta seção discutimos os aspectos específicos da implementação da metodologia de kernel para a estimação da função de volatilidade em modelos HJM Gaussianos e a consequente utilização deste modelo para o apreçamento de opções. Para isso realizamos duas análises - a primeira é uma análise de Monte Carlo comparando o desempenho das diversas espeficações para a obtenção de preços de opções de compra sobre títulos zero cupom, assumindo que o processo gerador de dados é um modelo Vasicek com dois fatores. Na análise empírica calculamos o preço de opções de compra sobre o IDI no mercado brasileiro, usando dados de mercado.

5.1 Simulações

Como será discutido a seguir, a metodologia de estimação da função de volatilidade pode ser realizada usando diferentes especificações para a regressão por kernel na estimação da função de volatilidade. A análise de Monte Carlo permite comparar o desempenho dessas diferentes especificações com o preço teórico do modelo, sendo que esse procedimento de simulação assume como processo gerador de dados um modelo Vasicek de dois fatores. De acordo com a especificação usada em Nawalkha, Beliaeva e Soto (2007), temos que a taxa de juros de curto prazo, nesse processo, é dada por:

$$r(t) = \delta(t) + Y_1(t) + Y_2(t), \quad Y_1(0) = 0 \quad e \quad Y_2(0) = 0$$
 (5.1)

Sendo que os dois fatores Y_1 e Y_2 apresentados são variáveis de estado e seguem as seguintes dinâmicas:

$$dY_1(t) = -\tilde{\kappa}_1 Y_1(t) dt + \nu_1 d\tilde{W}_1, \quad Y_1(0) = 0$$
(5.2)

$$dY_2(t) = -\tilde{\kappa}_2 Y_2(t) dt + \nu_2 d\tilde{W}_2, \quad Y_2(0) = 0$$
(5.3)

Dados os fatores latentes, pode-se aplicar o Lema de Itô sobre o processo de preços dos títulos e tomar a expectativa sob a medida neutra ao risco. Como esse modelo é afim, encontramos a dinâmica dos preços dos títulos como sendo:

$$P(t,T) = e^{A(\tau) - B_1(\tau)Y_1(t) - B_2(\tau)Y_2(t) - H(t,T)}$$
(5.4)

Em que:

$$H(t,T) = \int_{t}^{T} \delta(v)dv = H(0,T) - H(0,t) = A(T) - A(t) - \ln P(0,T) + \ln P(0,t) \quad (5.5)$$

Nesse contexto, é importante ressaltar algumas vantagens de se utilizar um modelo com dois fatores para encontrar o preço teórico das opções. Segundo Stehlikova e Sevcovic (2008), nos modelos de um fator, a estrutura a termo da taxa de juros é completamente determinada pela origem, ou seja, a taxa de curto prazo. No sentido de permitir diferentes estruturas a termo, partindo de uma mesma origem, surgiram modelos com mais de um fator. O artigo menciona, com base em outros estudos, que os modelos com dois fatores apresentam maior acuracidade no apreçamento de opções, quando comparados a modelos de um único fator.

Assim, a principal vantagem deste modelo Gaussiano de dois fatores é que a curva de volatilidade da taxa a termo pode assumir valores mais realistas e, assim, o processo gerador de dados é uma representação apropriada para a análise da metodologia proposta. Fórmulas fechadas para a obtenção de preços de derivativos, em especial opções sobre titulos zero cupom, são apresentadas em Nawalkha, Beliaeva e Soto (2007), permitindo a comparação direta com os preços obtidos por nossa metodologia.

Assumimos duas especificações para o processo de Vasicek. Na primeira simulação os parâmetros são dados por $\kappa_1=.8,\,\kappa_2=.7,\,\nu_1=.1,\,\nu_2=.1$ e $\rho_{12}=-.3$, seguindo a notação apresentada em Nawalkha, Beliaeva e Soto (2007). Na segunda simulação assumimos que os parâmetros ν_1 e ν_2 são variantes no tempo, seguindo um processo de Feller com parâmetros de reversão a média .2 e variância da variância igual a .1. Esta especificação com estes parâmetros variantes no tempo tem o objetivo de verificar a robustez do método de HJM Gaussiano quando o verdadeiro processo é não homogêneo

no tempo.

Assumimos maturidades com 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 90, 108 e 120 meses com um delta de 1/252 (dados diários) entre cada observação da curvas. Para obter curvas mais realistas, adicionamos um erro de observação aditivo aos preços de títulos simulados, mimetizando efeitos de bid-ask spread. Realizamos uma suposição de normalidade desses erros, com média nula e desvio padrão igual a 0,0035, valor calibrado para reproduzir o bid-ask spread observados em títulos do Tesouro Americano. A robustez do método de kernel em relação a este erro de medida é uma das vantagens fundamentais do método, como discutido em Jeffrey et al. (2004).

Na análise de simulação geramos curvas com 250 e 500 observaçõe no tempo, e realizamos 500 replicações do experimento de Monte Carlo. Nas tabelas de resultados apresentamos os resultados do erro quadrático médio entre o preço da opção obtido com o modelo HJM Gaussiano usando as funções de volatilidade estimadas e o verdadeiro preço da opção usando os verdadeiros parâmetros do processo Vasicek. Em cada Tabela analisamos a estimação de cinco diferentes espeficações para as variáveis independentes, considerando o método de kernel local constante, e analisamos o efeito da escolha de janela ótima por validação cruzada generalizada usando Mínimos Quadrados e por AIC. A estimação em questão, tanto para esse modelo quanto para o segundo, é realizada por meio de uma função de kernel gaussiana.

A estimação da volatilidade a que nos referimos deve ser realizada considerando a suposição de variância constante no tempo, porém variante entre as maturidades. Assim, é importante ressaltar que realizamos a estimação da função de volatilidade das taxas a termo em função das diferentes covariadas escolhidas em cinco distintas especificações da regressão não paramétrica. A função de volatilidade de interesse, porém, é a dos preços dos títulos zero cupom associados a essas taxas, sendo que, para tanto, é necessario integrar a função de volatilidade da taxa a termo em relação à maturidade. Assim, sob a referida suposição de constância da volatilidade no tempo, essa integral resulta no produto entre o tempo até a maturidade e a própria variância da taxa.

A intuição acerca dessa integração é apresentada no trabalho de Jeffrey et al. (2004), em que o estimador proposto para uma função de volatilidade de observações de um processo discreto em uma dada maturidade é o somatório das ponderações das observações pelo método de kernel ao longo do tempo. Assim, como em nosso trabalho estamos considerando as variáveis como contínuas, é necessário aplicar a integral, e não o somatório, ao longo das distintas maturidades, já que consideramos diversas maturidades em um dado

dia.

Conforme mencionado, nós analisamos cinco diferentes especificações para a regressão não paramétrica usada na estimação da função de volatilidade. O primeiro modelo, que denominamos Modelo 1, realiza uma regressão do rendimento menos a sua média ao quadrado contra as maturidades das taxas em um dia, usando uma especificação similar à proposta em Mercurio e Moraleda (2000), que assume uma função de volatilidade como função das maturidades da curva. Neste procedimento estimamos esta regressão dia a dia na curva de juros, e assumimos que a função de volatilidade é a média de todas estas curvas no tempo.

O segundo modelo usa a mesma especificação para a estimação cross section, mas usando uma ponderação distinta para as diversas funções de volatilidade estimadas, usando pesos obtidos através de uma ponderação definida pela função de kernel quadrática espectral, estudada por Andrews (1991) para a estimação de matrizes de variância de longo prazo, usando uma função que assume processos dependentes no tempo, conforme discutido em Zeileis (2004). Para tanto, o modelo utiliza pesos $\omega(j)$ derivados da estimação não paramétrica da variância de longo prazo \hat{V}_{lr}^A estimada para os rendimentos quadráticos:

$$\hat{V}_{lr}^{A} := \sum_{j=-(n-1)}^{(n-1)} \omega(j) \widehat{\gamma(j)}$$

$$(5.6)$$

 $\operatorname{com}\widehat{\gamma(j)}$ denotando a função de autocovariância e $\omega(j)$ sendo uma sequência de pesos na forma $\omega(j) = K(j/b)$. No caso, $K(\cdot)$ é uma função de kernel contínua com K(0) = 1 e b é um parâmetro de janela com $b \to \infty$ quando $n \to \infty$, estimado usando a regra definida em Andrews (1991).

O objetivo deste método de ponderação é robustificar a estimação do modelo HJM Gaussiano na presença de uma estrutura de volatilidade variante no tempo. Desta forma, a função de volatilidade estimada é uma média das curvas de volatilidade mais próximas, média calculada levando em conta a dependência temporal entre estas curvas estimadas usando-se as propriedades assintóticas da variância de longo prazo deste processo. Esperase que o ajuste dos preços com essa ponderação seja superior ao do primeiro modelo quando não se faz a suposição de volatilidade constante, questão que é discutida em breve nos resultados do modelo.

O terceiro modelo, por sua vez, já é uma regressão no tempo, em que a variância para

cada maturidade é estimada regredindo de forma não paramétrica as taxas centradas ao quadrado contra seu valor no período passado. Nesse caso, estamos considerando que a variância segue um processo autorregressivo de ordem um, e realizamos uma estimação para cada maturidade observada, obtendo os valores ajustados como previsão da função de variância do processo.

Um conjunto de regressores que teoricamente forneceria bastantes informações na explicação da volatilidade da taxa de juros seria o valor passado de todas as taxas de diferentes maturidades. No entanto, o modelo 3 mencionado utiliza apenas o valor defasado de uma variável (apenas uma maturidade). O motivo de não realizarmos um modelo completo de vetor autorregressivo para desenvolver essa regressão é basicamente que a velocidade de convergência dos estimadores para o parâmetro populacional seria ainda mais lento do que normalmente já é nos modelos não paramétricos, e isso poderia gerar viés na análise, problema conhecido como maldição da dimensionalidade. Assim, consideramos que a taxa em dada maturidade depende de seu valor passado, mas não depende das defasagens das taxas em outras maturidades.

Para tentar adicionar mais informação na estimação da função de volatilidade, mas ao mesmo tempo evitando o problema de dimensionalidade, usamos uma decomposição de componentes principais na curva e extraímos os dois primeiros componentes principais. Esta informação é utilizada nas estimações dos dois próximos modelos. Neste caso realizamos uma decomposição de componentes principais para a curva de juros, e regredimos o rendimento centrado ao quadrado de cada maturidade contra os componentes principais ao quadrado. No Modelo 4 das tabelas, a variável independente é o primeiro componente da decomposição mencionada, enquanto o Modelo 5 utiliza os dois primeiros componentes.

Para ilustrar as regressões descritas, podemos apresentar a estimação da volatilidade que consta em Jeffrey et al. (2004):

$$\tilde{\gamma}^2(x,\tau) = \frac{1}{\Delta} \frac{\sum_{i=1}^n K_{h_x} (X_{t_i} - x) (y_{t_{i+1}}(\tau) - y_{t_i}(\tau))^2}{\sum_{i=1}^n K_{h_x} (X_{t_i} - x)}$$
(5.7)

Nessa equação, $\tilde{\gamma}^2$ é a volatilidade estimada, $K_h(\cdot)$ é a função de kernel gaussiana e h_x é uma sequencia de janelas de kernel. Em nosso trabalho, a variável X_{t_i} são as maturidades no caso dos modelos 1 e 2, o valor defasado do próprio rendimento do título para o modelo 3 e, para os modelos 4 e 5, os regressores serão o primeiro e os dois primeiros componentes principais da série, respectivamente.

Conforme descrito, as regressões de kernel utilizadas apresentam os seguintes formatos, respectivamente, para os cinco modelos, seguindo a notação de Jeffrey et al. (2004):

$$\tilde{\gamma}^{2}(\tau) = \frac{1}{\Delta} \frac{\sum_{i=1}^{n} K_{h_{\tau}}(\tau) (y_{t_{i+1}}(\tau) - y_{t_{i}}(\tau))^{2}}{\sum_{i=1}^{n} K_{h_{\tau}}(\tau)}$$
(5.8)

$$\tilde{\gamma}^2(y,\tau) = \frac{1}{\Delta} \frac{\sum_{i=1}^n K_{h_y} (y_{t-1_i} - y)^2 (y_{t_{i+1}}(\tau) - y_{t_i}(\tau))^2}{\sum_{i=1}^n K_{h_y} (y_{t-1_i} - y)^2}$$
(5.9)

$$\tilde{\gamma}^2(pca_1, \tau) = \frac{1}{\Delta} \frac{\sum_{i=1}^n K_{h_{pca1}}(PCA_1t_i^2)(y_{t_{i+1}}(\tau) - y_{t_i}(\tau))^2}{\sum_{i=1}^n K_{h_{pca1}}(PCA_1t_i^2)}$$
(5.10)

$$\tilde{\gamma}^2(pca_1, pca_2, \tau) = \frac{1}{\Delta} \frac{a}{b} \tag{5.11}$$

Em que:

$$a = \sum_{i=1}^{n} K_{h_{pea1}}(PCA_1t_i^2) K_{h_{pea2}}(PCA_2t_i^2)(y_{t_{i+1}}(\tau) - y_{t_i}(\tau))^2$$
 (5.12)

$$b = \sum_{i=1}^{n} K_{h_{pca1}}(PCA_1t_i^2) K_{h_{pca2}}(PCA_2t_i^2)$$
(5.13)

Lembrando que a primeira regressão refere-se aos dois primeiros modelos, já que a diferença entre eles está basicamente na ponderação das diferentes volatlidades calculadas para cada dia (dessa forma, a regressão é a mesma).

Para cada um desses modelos, utilizamos duas técnicas de seleção de janela ótima apresentadas por Racine e Li (2007), que são denominadas nas tabelas de resultado como MQ e AIC. As janelas em questão são respectivamente calculadas por validação cruzada de mínimos quadrados e pelo critério modificado de AIC, apresentados anteriormente.

5.2 Apreçamento de opções sobre títulos zero cupom

No procedimento de simulação, analisamos os preços de opções de compra sobre títulos zero cupom com valores de face iguais a um. Como discutimos anteriormente, temos

fórmulas analíticas para este preço em modelos afins, em especial o modelo Vasicek que é assumido como o processo gerador de dados. O apreçamento de opções sobre títulos zero cupom, de acordo com Andersen e Piterbarg (2010), é feita no modelo HJM Gaussiano de acordo o valor que a opção de compra europeia paga na data da maturidade:

$$V(T) = (P(T, T^*) - K)^+, \ T^* > T \tag{5.14}$$

Quando o preço do ativo na data da maturidade $P(T,T^*)$ for maior que o preço de exercício da opção K, ela é exercida. No caso do modelo HJM Gaussiano, a opção em questão é precificada da seguinte forma:

$$V(T) = P(t, T^*)\Phi(d_+) - P(t, T)K\Phi(d_-)$$
(5.15)

Em que:

$$d_{\pm} = \frac{\ln(P(t, T^*)/(KP(t, T))) \pm v/2}{\sqrt{v}}$$
 (5.16)

$$v = \int_{t}^{T} |\sigma_{P}(u, T^{*}) - \sigma_{P}(u, T)|^{2} du$$
 (5.17)

com T^* denotando o período em que é realizado o apreçamento, T a maturidade do titulo e $\sigma_P(u,T^*)$ e $\sigma_P(u,T)$ denotam as volatilidades do título zero cupom nos períodos T^* e T. Note que a função de volatilidade é usada no cálculo do termo v, que é a integral da função de volatilidade do preço do título entre esses dois períodos, notando que a volatilidade do preço do título em um modelo HJM Gaussiano é obtida como a integral da função determinista da volatilidade da taxa, o que neste caso representa multiplicar a volatilidade da taxa pelo diferencial de tempo entre T^* e T, como discutido em Filipovic (2009).

Dada a explicação dos modelos utilizados na estimação da volatilidade e no apreçamento das opções, explicitamos alguns resultados. A seguir apresentamos alguns gráficos gerados pelas simulações, em que cada curva apresentada representa uma maturidade da taxa:

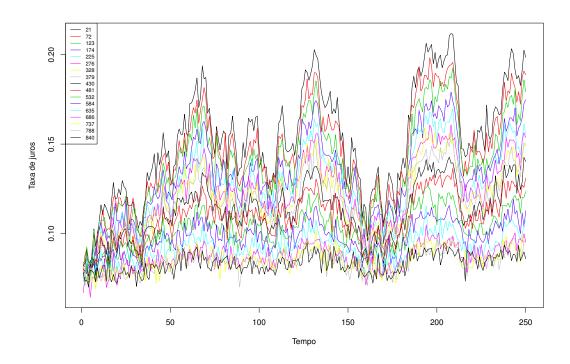


Figura 1: Simulações com variância constante - amostra $250\,$

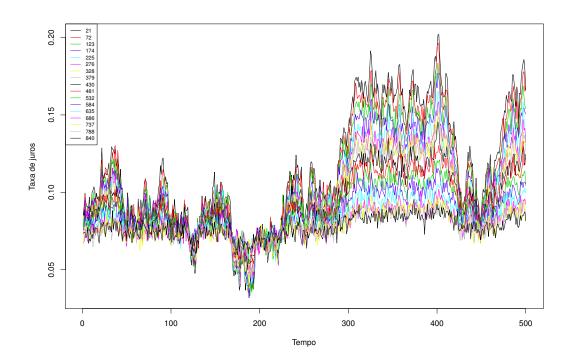


Figura 2: Simulações com variância constante - amostra $500\,$

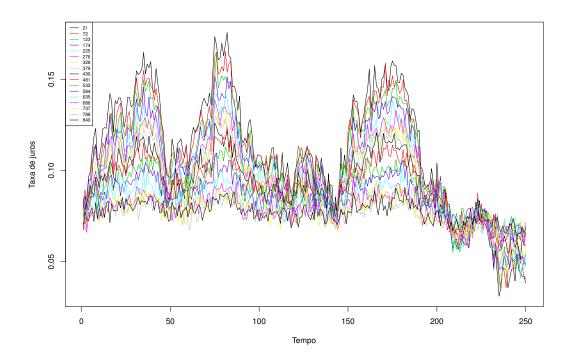


Figura 3: Simulações com variância alterando no tempo - amostra $250\,$

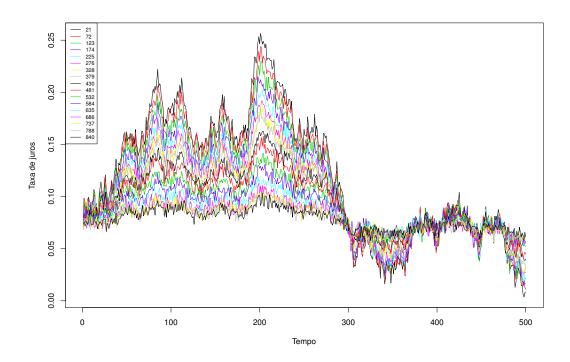


Figura 4: Simulações com variância alterando no tempo - amostra $500\,$

Nas figuras 1-4 mostramos exemplos das estruturas a termo simuladas pelo processo Vasicek com dois fatores, sendo que cada uma das distintas maturidades da taxa está designada com uma cor diferente e exposta na legenda dos gráficos. Apresentamos as simulações com 250 e 500 observações para os processos com parâmetros ν_1 e ν_2 constantes e variantes no tempo, sendo que é possível notar que os processos simulados geram curvas bastante complexas e similares às observadas para processos reais de juros.

No exercício de simulação assumimos o apreçamento de opções de compra sobre títulos zero cupom com maturidade de 10 anos, assumindo-se tempo até a maturidade da opção de 1 ano, .5 e .25 anos. Em cada caso assumimos diversos preços de exercício para o opção, que são aproximados para ter casos de opções in, at e out-of-the money em cada tempo até a maturidade, como discutido em Mercurio e Moraleda (2000), que realiza um experimento similar.

Nas tabelas 1-4, são apresentadas as raízes dos erros quadráticos médios da comparação entre o preço simulado pelo modelo Vasicek com dois fatores das opções sobre títulos zero cupom e o preço calculado por cada um dos cinco modelos, além do modelo de referência usando o HJM Gaussiano calibrado pela análise de componentes principais (PCA), que é uma forma não paramétrica bastante comum de calibração de modelos HJM Gaussianos, como discutido em Jarrow (2002) e Filipovic (2009).

A tabela 1 mostra a raiz do erro quadrático médio do apreçamento de opções sobre um título zero cupom, cuja dinâmica de preços e taxas é encontrada a partir do modelo Vasicek com dois fatores. No caso, consideramos a volatilidade da série de juros como constante no tempo e a amostra considerada é de 250 dias. Analisando os modelos em cross section (modelos 1 e 2), notamos que o valor da REQM aumenta com o preço de exercício da opção e diminui com o prazo até a maturidade. Na tabela 1, pode-se observar os valores do modelo 1 em vermelho corroborando essa constatação de aumento do erro com o preço de exercício maior. Já em azul pode-se observar a redução do erro de apreçamento com a elevação do prazo até a maturidade.

Desses primeiros resultados pode ser depreendido que a REQM diminui com o prazo até a maturidade provavelmente porque uma opção com prazo até a maturidade mais longo deve ter maior probabilidade de exercício (dado um mesmo *strike*), sendo que a acuracidade do cálculo nesse caso deve ser maior. À medida que o prazo até a maturidade diminui, e a opção está *out-of-the-money*, sua probabilidade de exercício se reduz.

Corroborando a análise anterior, outra constatação importante é que a REQM aumenta à medida que o preço de exercício da opção aumenta, fato que pode se dever ao

Tempo até a Maturidade	Preço de exercício	0.45	0.5	0.55	0.6
0.25 Anos	PCA	0.056	0.069	0.076	0.075
	Modelo 1 - MQ	0.002593	0.004392	0.005557	0.006184
	Modelo 1 - AIC	0.002606	0.004413	0.005580	0.006190
	Modelo 2 - MQ	0.000645	0.002854	0.006051	0.008365
	Modelo 2 - AIC	0.000645	0.002850	0.006039	0.008351
	Modelo 3 - MQ	0.001326	0.003080	0.004556	0.005452
	Modelo 3 - AIC	0.001319	0.003064	0.004527	0.005421
	Modelo 4 - MQ	0.001014	0.002552	0.004586	0.005855
	Modelo 4 - AIC	0.001013	0.002547	0.004582	0.005854
	Modelo 5 - MQ	0.001013	0.002550	0.004584	0.005855
	Modelo 5 - AIC	0.001012	0.002543	0.004576	0.005854
Tempo até a Maturidade	Preço de exercício	0.48	0.53	0.58	0.63
	PCA	0.037	0.049	0.059	0.062
	Modelo 1 - MQ	0.000719	0.001626	0.003930	0.006944
	Modelo 1 - AIC	0.000725	0.001634	0.003926	0.006926
	Modelo 2 - MQ	0.000323	0.002140	0.006524	0.009918
	Modelo 2 - AIC	0.000323	0.002140	0.006517	0.009908
0.5 Anos	Modelo 3 - MQ	0.000237	0.000926	0.002595	0.004465
	Modelo 3 - AIC	0.000234	0.000921	0.002593	0.004463
	Modelo 4 - MQ	0.000205	0.000871	0.002640	0.004613
	Modelo 4 - AIC	0.000205	0.000871	0.002651	0.004629
	Modelo 5 - MQ	0.000206	0.000872	0.002645	0.004618
	Modelo 5 - AIC	0.000206	0.000872	0.002657	0.004648
Tempo até a Maturidade	Preço de exercício	0.5	0.55	0.6	0.65
1 Ano	PCA	0.029	0.041	0.051	0.056
	Modelo 1 - MQ	0.000362	0.001398	0.004849	0.008258
	Modelo 1 - AIC	0.000365	0.001398	0.004836	0.008238
	Modelo 2 - MQ	0.000297	0.002219	0.007275	0.010601
	Modelo 2 - AIC	0.000297	0.002218	0.007271	0.010593
	Modelo 3 - MQ	0.000091	0.000906	0.003582	0.005797
	Modelo 3 - AIC	0.000091	0.000907	0.003587	0.005800
	Modelo 4 - MQ	0.000089	0.000939	0.003729	0.006037
	Modelo 4 - AIC	0.000090	0.000943	0.003743	0.006054
	Madala 5 MO	0.000090	0.000939	0.003733	0.006042
	Modelo 5 - MQ	0.000090	0.000333	0.000100	0.000012

Tabela 1: REQM - apreçamento de opção sobre um título zero cupom - Volatilidade constante no tempo - Amostra $250\,$

menor número de exercícios das opções em questão, o que reduz o número de pontos para cálcular os preços, aumentando a variância. Assim, a menor probabilidade de exercício de opções *out-of-the-money* pode reduzir a acuracidade do cálculo.

Com relação aos modelos em série temporal, alguns resultados são muito similares. A REQM dos modelos 3, 4 e 5 também tendem a aumentar com o preço de exercício da opção e diminuir com o prazo até a maturidade do título em questão. Entre as técnicas, podemos perceber que o modelo 5 é, normalmente, similar ao modelo 4, evidenciando que o acréscimo do segundo fator da análise de componentes principais não parece contribuir significativamente para a acuracidade do cálculo.

Na tabela 2 apresentamos os resultados para a estimação usando o tamanho de amostra igual a 500. Os resultados são similiares aos encontrados na tabela 1. Podemos constatar uma maior diferença entre os valores dos erros utilizando cada um dos métodos de seleção de janela de kernel. Em geral, o AIC_c apresenta erro um pouco menor que o outro critério, para quase todos os modelos. Essa característica pode ser observada pela diferença entre os valores em vermelho (janela selecionada pelo critério de mínimos quadrados) e valores em azul (janela selecionada pelo critério AIC_c). O modelo 4 foi escolhido para mostrar essa característica, mas outros modelos também apresentam essa constatação na maioria das especificações de preços de exercício e tempo até a maturidade.

Para buscar uma explicação para esse fato, podemos nos referir ao trabalho de Li e Racine (2004), segundo o qual as propriedades assintóticas de ambos os métodos são as mesmas, mas o critério AIC_c normalmente apresenta resultados melhores que a validação cruzada por mínimos quadrados em amostras finitas. Em ambas as tabelas pode-se observar essa leve superioridade do critério AIC_c , mas esse comportamento evidencia-se um pouco mais na análise com 500 observações.

Nas tabelas 3 e 4 apresentamos os resultados para o processo gerador de dados que considera a função de volatilidade variante no tempo, i.e. ν_1 e ν_2 variam no tempo de acordo com um processo de Feller. Em determinadas combinações de preços de exercício e tempo até a maturidade, o modelo 2 de fato apresenta REQM menor que o primeiro modelo, mas isso não acontece na maior parte dos casos. Pode-se observar em destaque na tabela 3 alguns exemplos em que o erro do modelo 2 (em azul) são maiores que os do modelo 1 (em vermelho). Assim, pode-se concluir com base nos testes que a ponderação das estimações por kernel, nesse caso, não agrega informação considerável ao apreçamento das opções. Vale lembrar que a comparação que fazemos é relativa entre os dois primeiros modelos, e não absoluta entre a REQM nos dois casos (com volatilidade constante e não constante), já que os parâmetros do modelo Vasicek se alteram de um caso para outro, alterando também os resultados absolutos.

Outra constatação importante é que os erros são, em geral, maiores nessas tabelas 3 e

Tempo até a Maturidade	Preço de exercício	0.45	0.5	0.55	0.6
0.25 Anos	PCA	0.059	0.071	0.076	0.073
	Modelo 1 - MQ	0.004285	0.007211	0.008827	0.008236
	Modelo 1 - AIC	0.004306	0.007244	0.008867	0.008269
	Modelo 2 - MQ	0.002712	0.004383	0.007199	0.007943
	Modelo 2 - AIC	0.002718	0.004386	0.007189	0.007932
	Modelo 3 - MQ	0.005479	0.008267	0.010043	0.009259
	Modelo 3 - AIC	0.005467	0.008238	0.010004	0.009222
	Modelo 4 - MQ	0.002744	0.005864	0.007516	0.00714
	Modelo 4 - AIC	0.002730	0.005836	0.007483	0.007118
	Modelo 5 - MQ	0.005353	0.008023	0.009718	0.008958
	Modelo 5 - AIC	0.005344	0.007998	0.009681	0.008924
Tempo até a Maturidade	Preço de exercício	0.48	0.53	0.58	0.63
	PCA	0.039	0.051	0.060	0.062
	Modelo 1 - MQ	0.001308	0.002663	0.004513	0.005948
	Modelo 1 - AIC	0.001318	0.002681	0.004525	0.00594
	Modelo 2 - MQ	0.001325	0.002917	0.007535	0.00945
	Modelo 2 - AIC	0.001330	0.002920	0.007528	0.00944
0.5 Anos	Modelo 3 - MQ	0.002233	0.003811	0.004779	0.00507
	Modelo 3 - AIC	0.002229	0.003801	0.004763	0.00505
	Modelo 4 - MQ	0.000375	0.001378	0.002880	0.00455
	Modelo 4 - AIC	0.000370	0.001366	0.002865	0.00454
	Modelo 5 - MQ	0.002161	0.003677	0.004603	0.00499
	Modelo 5 - AIC	0.002157	0.003670	0.004593	0.005003
Tempo até a Maturidade	Preço de exercício	0.5	0.55	0.6	0.65
1 Ano	PCA	0.031	0.043	0.052	0.056
	Modelo 1 - MQ	0.000693	0.001777	0.004621	0.00672
	Modelo 1 - AIC	0.000700	0.001785	0.004612	0.00670
	Modelo 2 - MQ	0.000918	0.002808	0.008202	0.01003
	Modelo 2 - AIC	0.000922	0.002810	0.008196	0.01003
	Modelo 3 - MQ	0.001315	0.002520	0.003630	0.00466
	Modelo 3 - AIC	0.001313	0.002518	0.003631	0.00465
	Modelo 4 - MQ	0.000178	0.000779	0.002320	0.00491
	Modelo 4 - AIC	0.000181	0.000785	0.002322	0.00491
	Modelo 5 - MQ	0.001260	0.002463	0.003658	0.00478
	Modelo 9 - MQ	0.001200	0.002400	0.000000	0.0010

Tabela 2: REQM - apreçamento de opção sobre um título zero cupom - Volatilidade constante no tempo - Amostra $500\,$

4 quando comparados aos erros das tabelas 1 e 2. Como fazemos a suposição de variância constante na primeira análise, é esperado que o erro do apreçamento dos derivativos seja de fato menor no primeiro caso. A variância que se altera no tempo pode ser uma fonte

Tempo até a Maturidade	Preço de exercício	0.45	0.5	0.55	0.6
	PCA	0.054	0.067	0.073	0.072
	Modelo 1 - MQ	0.002809	0.004827	0.006175	0.007255
	Modelo 1 - AIC	0.002821	0.004846	0.006189	0.007254
	Modelo 2 - MQ	0.001136	0.003738	0.007586	0.010060
0.25 Anos	Modelo 2 - AIC	0.001135	0.003735	0.007575	0.010047
	Modelo 3 - MQ	0.001047	0.002628	0.004673	0.005891
	Modelo 3 - AIC	0.001045	0.002622	0.004667	0.005883
	Modelo 4 - MQ	0.001014	0.002552	0.004586	0.005855
	Modelo 4 - AIC	0.001013	0.002547	0.004582	0.005854
	Modelo 5 - MQ	0.001013	0.002550	0.004584	0.005855
	Modelo 5 - AIC	0.001012	0.002543	0.004576	0.005854
Tempo até a Maturidade	Preço de exercício	0.48	0.53	0.58	0.63
	PCA	0.035	0.048	0.057	0.060
	Modelo 1 - MQ	0.000951	0.002648	0.005605	0.008807
	Modelo 1 - AIC	0.000955	0.002651	0.005598	0.008787
	Modelo 2 - MQ	0.000748	0.003179	0.008151	0.011668
	Modelo 2 - AIC	0.000748	0.003178	0.008145	0.011662
0.5 Anos	Modelo 3 - MQ	0.000367	0.001556	0.004249	0.006767
	Modelo 3 - AIC	0.000368	0.001562	0.004265	0.006779
	Modelo 4 - MQ	0.000367	0.001586	0.004370	0.006950
	Modelo 4 - AIC	0.000367	0.001588	0.004380	0.006962
	Modelo 5 - MQ	0.000367	0.001585	0.004371	0.006958
	Modelo 5 - AIC	0.000367	0.001589	0.004387	0.006978
Tempo até a Maturidade	Preço de exercício	0.5	0.55	0.6	0.65
	PCA	0.028	0.039	0.049	0.053
	Modelo 1 - MQ	0.000820	0.002753	0.006748	0.010231
	Modelo 1 - AIC	0.000822	0.002752	0.006734	0.010210
	36 11 0 360				0.010460
	Modelo 2 - MQ	0.000804	0.003403	0.009058	0.012469
	Modelo 2 - MQ Modelo 2 - AIC	0.000804 0.000804	0.003403 0.003403	0.009058 0.009053	0.012469 0.012464
1 Ano	<u> </u>				
1 Ano	Modelo 2 - AIC	0.000804	0.003403	0.009053	0.012464
1 Ano	Modelo 2 - AIC Modelo 3 - MQ	0.000804 0.000380	0.003403 0.001926	0.009053 0.005503	$\begin{array}{c} 0.012464 \\ 0.008228 \end{array}$
1 Ano	Modelo 2 - AIC Modelo 3 - MQ Modelo 3 - AIC	0.000804 0.000380 0.000380	0.003403 0.001926 0.001931	0.009053 0.005503 0.005519	0.012464 0.008228 0.008240
1 Ano	Modelo 2 - AIC Modelo 3 - MQ Modelo 3 - AIC Modelo 4 - MQ	0.000804 0.000380 0.000380 0.000383	0.003403 0.001926 0.001931 0.001964	0.009053 0.005503 0.005519 0.005660	0.012464 0.008228 0.008240 0.008449

Tabela 3: REQM - apreçamento de opção sobre um título zero cupom - Volatilidade variante no tempo - Amostra $250\,$

de imprecisão no cálculo, já que estamos assumindo um processo HJM Guassiano que assume uma função de volatilidade determinista, fato que é também observado na análise empírica deste trabalho, apresentada na próxima seção.

Com relação a outras constatações, é perceptível que alguns padrões observados na análise sob variância constante também são observados aqui: a REQM tende a aumentar com o preço de exercício e diminuir com o prazo até a maturidade. Além disso, a diferença entre os resultados das análises com ambas as janelas de kernel consideradas também seguem, via de regra, o padrão observado na primeira análise: os valores são bastante próximos, mas a janela de AIC_c performa marginalmente melhor para todos os modelos, com exceção do modelo 1. Esse fato pode ser justificado pelas constatações apresentadas em Hurvich, Simonoff e Tsai (1998) que, conforme mencionado no capítulo 4 deste trabalho, rodaram simulações para mostrar que essa técnica de seleção de janela funciona bem quando comparada a uma série de métodos de validação cruzada generalizados.

Assim, podemos consolidar as principais conclusões a respeito dos resultados apresentados. Fica evidente em todas as tabelas que a raiz do erro quadrático médio do apreçamento dos modelos propostos é consistentemente menor que a REQM do apreçamento realizado por meio da estimação da volatilidade por PCA. Assim, o procedimento de simulação mostra por meio da repetição que o modelo proposto apresenta um ajuste mais acurado no apreçamento do derivativo em questão, quando comparado ao método convencional de análise de componentes principais.

Comparando os dois modelos em *cross section*, é possível perceber que a REQM do modelo 2 é normalmente menor que o do modelo 1 para preços de exercício relativamente pequenos, mas esse fato se inverte para valores maiores de preços de exercício. Além disso, parece haver outro padrão: à medida que o prazo até a maturidade do título aumenta, a diferença entre a REQM do primeiro modelo e do segundo parece diminuir.

Uma possível explicação para o fato de o erro computado ser menor, em média, para o modelo 2 em relação ao modelo 1 quando os preços de exercício são mais baixos pode ser a maior acuracidade do cálculo para *strikes* menores (opções mais *in-the-money*). Como o modelo 2 realiza a ponderação de kernel para estimar a função de volatilidade média entre os dias analisados, ele pode apresentar um resultado superior em determinados casos, fato que se evidencia em opções de preço de exercício menor. No entanto, o fato de essa diferença entre os modelos diminuirem em certos casos (quando o prazo até a maturidade aumenta, por exemplo), pode ser um indicativo de que a ponderação por kernel aplicada ao modelo 2, para estimar a função de volatilidade média, pode não agregar muitas informações relevantes.

Com relação aos modelos em série temporal, alguns resultados são muito similares. A REQM dos modelos 3, 4 e 5 também tendem a aumentar com o preço de exercício da opção

Tempo até a Maturidade	Preço de exercício	0.45	0.5	0.55	0.6
0.25 Anos	PCA	0.057	0.068	0.073	0.071
	Modelo 1 - MQ	0.003780	0.006299	0.008013	0.00831
	Modelo 1 - AIC	0.003796	0.006322	0.008035	0.00832
	Modelo 2 - MQ	0.001916	0.004807	0.008559	0.01053
	Modelo 2 - AIC	0.001916	0.004805	0.008551	0.01052
	Modelo 3 - MQ	0.004902	0.008228	0.010314	0.00940
	Modelo 3 - AIC	0.004885	0.008210	0.010294	0.00937
	Modelo 4 - MQ	0.004779	0.008053	0.010098	0.00917
	Modelo 4 - AIC	0.004765	0.008038	0.010081	0.00915
	Modelo 5 - MQ	0.004775	0.008048	0.010093	0.0091'
	Modelo 5 - AIC	0.004764	0.008036	0.010081	0.00915
Tempo até a Maturidade	Preço de exercício	0.48	0.53	0.58	0.63
	PCA	0.038	0.049	0.058	0.059
	Modelo 1 - MQ	0.001306	0.003348	0.006463	0.00854
	Modelo 1 - AIC	0.001312	0.003355	0.006459	0.0085
	Modelo 2 - MQ	0.001110	0.004074	0.008993	0.0120'
	Modelo 2 - AIC	0.001110	0.004075	0.008990	0.0120
0.50 Anos	Modelo 3 - MQ	0.002094	0.004807	0.007205	0.0071'
	Modelo 3 - AIC	0.002088	0.004807	0.007217	0.00718
	Modelo 4 - MQ	0.002036	0.004757	0.007193	0.0072
	Modelo 4 - AIC	0.002032	0.004759	0.007204	0.0072
	Modelo 5 - MQ	0.002034	0.004754	0.007189	0.0072
	Modelo 5 - AIC	0.002031	0.004759	0.007210	0.0072
Tempo até a Maturidade	Preço de exercício	0.5	0.55	0.6	0.65
1 Ano	PCA	0.030	0.041	0.050	0.053
	Modelo 1 - MQ	0.001032	0.003312	0.007356	0.0096'
	Modelo 1 - AIC	0.001034	0.003312	0.007343	0.0096
	Modelo 2 - MQ	0.001069	0.004206	0.009873	0.0126
	Modelo 2 - AIC	0.001069	0.004207	0.009870	0.0126'
	Modelo 3 - MQ	0.001719	0.004597	0.007307	0.00762
	Modelo 3 - AIC	0.001719	0.004604	0.007329	0.00764
	Modelo 4 - MQ	0.001698	0.004614	0.007410	0.00778
	Modelo 4 - AIC	0.001698	0.004622	0.007430	0.0078
	Modelo 5 - MQ	0.001697	0.004612	0.007407	0.00778
	Modelo 5 - AIC	0.001697	0.004622	0.007437	0.00782

Tabela 4: REQM - apreçamento de opção sobre um título zero cupom - Volatilidade variante no tempo - Amostra $500\,$

e diminuir com o prazo até a maturidade do título em questão. Entre as técnicas, podemos perceber que a comparação do modelo 4 com o modelo 5 mostra que há pouca informação agregada pelo acréscimo do segundo fator da análise de componentes principais, já que o

resultado desses dois modelos é bastante similar em todas as tabelas. Em destaque, na tabela 4, pode-se observar essa constatação comparando os valores em vermelho (modelo 4) com os valores em azul (modelo 5).

O modelo 3, por sua vez, apresenta um resultado relativamente maior que o 4 e o 5 na maioria das tabelas, com exceção de alguns pontos em que o preço de exercício é maior, em especial nas tabelas cujo número de observações é igual a 500. Assim, de uma maneira geral, e modelo 3 é preterido em relação ao 4 e ao 5. Essa questão pode ser devida ao fato de não utilizarmos as defasagens de todas as maturidades nesse modelo, mas apenas de uma maturidade. Apesar de evitarmos que a convergência dos estimadores fosse bastante lenta, fato que poderia ocorrer caso usassemos um VAR na estimação, a REQM computada por esse modelo é maior.

Assim, corrobora-se a conclusão que o modelo proposto apresenta resultados superiores em termos de acuracidade no apreçamento de opções sobre títulos zero cupom com taxas simuladas por um processo de Monte Carlo, já que a raiz do erro quadrático médio computada é, em média, menor que aquela computada pelo modelo PCA, bastante utilizado na literatura.

5.3 Aplicação empírica do modelo - apreçamento das opções sobre o IDI

Para mostrar uma aplicação do método para problemas reais, usamos o método para realizar o apreçamento de opções de compra sobre o índice DI, um derivativo de taxas de juros bastante importante no Brasil. Primeiramente, é importante discorrer sobre o funcionamento do mercado de Depósito Interfinanceiro para contextualizar o trabalho. As opções sobre o índice de Depósito Interfinanceiro (IDI) são derivativos cotados em pontos do índice, em que cada ponto é definido pela BM&F com o valor de R\$ 1,00. Elas são chamadas de opções asiáticas, dado que seu retorno depende do preço médio do ativo objeto ao longo do período analisado, e não de seu valor no dia de vencimento da opção (como acontece com muitas outras opções no mercado).

Seu ativo objeto, o IDI, nada mais é que a acumulação das médias das taxas de depósito interfinanceiro de um dia. Essas taxas, por sua vez, são calculadas com base nos contratos de DI futuro, que são instrumentos cujo valor de referência é de 100.000 pontos na data de vencimento, e o preço de compra ou venda é o valor presente desse montante de referência (que é de R\$ 100.000, dado que cada ponto possui o valor de R\$ 1,00).

Assim, a volatilidade que calculamos nessa seção do trabalho por meio das regressões não paramétricas de kernel nada mais é que o estimador da função de volatilidade dos PU (preços unitários) dos contratos DI1.

Os contratos DI Futuro são altamente negociados no mercado brasileiro, sendo o principal derivativo de juros baseado no CDI. Esse contratos possuem uma sistemática de ajustes diários e há contratos vencendo todos os meses, sendo que eles refletem a taxa do CDI do momento inicial até o vencimento do contrato.

Já as opções sobre o IDI possuem uma liquidez bem menor que os contratos DI Futuros. Esse fato pode se dever ao elevado valor relativo das opções, que pode afugentar investidores menores, e também à falta de um modelo de apreçamento amplamente aceito e com menos deficiências teóricas. Como o volume negociado é baixo, pesquisadores normalmente enfrentam dificuldades na aplicação prática de modelos sobre esse derivativo.

Como referência, podemos observar a tabela 5, em que são apresentadas as médias de volume negociado tanto das opções de compra sobre o IDI quanto dos contratos futuros do DI de um dia (DI1) ao longo do mês de outubro de 2013. Consta na tabela também a média de contratos negociados de cada ativo, e pode-se observar que o volume de negócios das opções de compra sobre o IDI é bastante reduzido quando comparado ao volume do DI futuro.

	Número médio	Média de contratos	Volume médio
Ativo	de negócios	negociados	negociado em R\$
	$(\mathrm{Out}/2013)$	$(\mathrm{Out}/2013)$	$(\mathrm{Out}/2013)$
Opção de compra	13	78326	R\$ 10.014.353,47
sobre o IDI	10	10020	10.014.555,47
Contrato DI	10974	1060016	R \$ 91.469.179.304,91
Futuro	10974	1000010	10 \$ 91.409.179.504,91

Tabela 5: Média do volume de negócios em outubro de 2013 - IDI e DI Futuro

Nesse contexto, na busca por um modelo mais acurado de apreçamento de opções sobre o IDI, diversas discussões vêm sendo apresentadas na literatura brasileira com os mais diversos modelos. A formulação teórica original do apreçamento de opções sobre o IDI no contexto de um modelo Vasicek foi proposta em Vieira e Valls (1999) e Valls e Vieira (2001); o apreçamento em um contexto de modelos Cox-Ingersoll-Ross foi proposto em Barbachan e Ornelas (2003), o uso do modelo de Hull-White foi discutido em Gluckstern (2001) e uma abordagem baseada na calibração de modelos HJM Gaussianos usando análise de componentes principais foi discutida em Barbedo, Lion e Vicente (2010).

Para realizar o apreçamento das opções utilizando os resultados da estimação da volatilidade, realizamos a calibração da função de volatilidade sobre o DI e fazemos a integral dessa função com relação à maturidade para encontrar a volatilidade do IDI. Lembrando que isso é possível pois o IDI é uma acumulação dos contratos DI de um dia e, portanto, é similar a um título zero cupom. Como estamos no modelo HJM Gaussiano, podemos modificar a função de apreçamento de uma opção de compra sobre título zero cupom para o caso de uma opção de compra sobre o IDI, dado que o fato de a volatilidade ser determinista implica que ela varia de forma não aleatória com a maturidade, possibilitando seu uso como input da fórmula de apreçamento em ambos os casos. Neste caso a fórmula é uma generalização do modelo de Black apresentado no apreçamento de opções sobre títulos zero cupom, usando um parâmetro de volatilidade integrada v dado por:

$$v = \int_{t}^{T} |\sigma_{P}(u, T^{0}) - \sigma_{P}(u, T)|^{2} du$$
 (5.18)

com T^0 denotando o período de apreçamento da opção e T sendo a maturidade da opção. Nesse caso, o parâmetro v é calculado como a integral da função de volatilidade entre um período 0 e um período de maturidade 0. Como o IDI é a acumulação das taxas DI, ele se comporta de forma análoga a um título zero cupom e, assim, a volatilidade do IDI pode ser calculada como a volatilidade de um título zero cupom com maturidade 0 calculado com a estrutura a termo das curvas DI com maturidades 0 e 0. Note que nesse caso estimamos a função de volatilidade realizando a estimação não paramétrica sobre as curvas de contratos de DI futuro no mercado brasileiro, e por não arbitragem a estimação da função de volatilidade nesse caso serve para calcular a volatilidade do IDI entre o período inicial e a maturidade da opção.

Para o cálculo da função volatilidade, os cinco modelos descritos na seção anterior, sobre simulações, também são utilizados nessa aplicação sobre o IDI. Vale relembrar os regressores de cada uma das funções utilizadas: nos denominados modelos 1 e 2, utilizamos a própria maturidade das taxas analisadas como regressores. A diferença básica entre esses modelos está na ponderação das distintas funções de volatilidade estimadas para cada dia da amostra, no sentido de encontrar uma função de volatilidade que é utilizada no apreçamento das opções. Já no modelo 3 o regressor é o valor da própria taxa quadrática centrada em sua média no período anterior, considerando que a volatilidade segue um processo autorregressivo. Para os modelos 4 e 5, respectivamente, consideramos como regressores o primeiro e os dois primeiros termos da análise de componentes principais.

Neste caso, usamos diferentes janelas (amostras) de estimação de curvas de DI para o

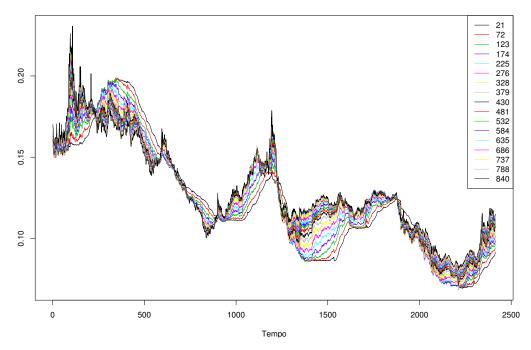


Figura 5: Estrutura a Termo de DI Futuros - 1999-2013

cálculo de funções de volatilidade, para verificar o impacto dessa escolha sobre os preços das opções. A análise é feita com janelas de 22 dias (que é aproximadamente um mês em dias úteis), 125 (aproximadamente 6 meses), 250 (aproximadamente um ano) e 500 (cerca de 2 anos). A análise é realizada tomando-se os últimos 300 pregões antes de 26/06/2013, e o apreçamento é feito para todas as opções de compra sobre o IDI que tenham sido negociadas naquele dia. Assim precificamos os últimos 300 pregões, estimando a função de volatilidade com janelas móveis com os últimos 22, 125, 250 e 500 curvas de juros de DI anteriores a cada pregão.

Como a curva de DI não é observada com maturidades fixas, interpolamos as curvas de DI observada na BM&F com vencimentos variando de 21 a 840 dias, de modo a obter 17 maturidades igualmente espaçadas nesse intervalo. Assim, a estimação da volatilidade dessa série com 17 maturidades é realizada por meio das regressões de kernel e as ponderações são feitas para encontrar uma função de volatilidade que será usada no apreçamento, assim como é feito na seção das simulações. A figura 5 mostra a evolução da estrutura de DI Futuros negociada na BM&F entre 1999 e 2013, nas maturidades interpoladas descritas acima.

Já a tabela 6 mostra um exemplo de precificação do método de kernel, usando o método 3 discutido anteriormente, para precificar as opções de compra sobre IDI com diferentes maturidades e tempo até vencimento no pregão do dia 04/04/2012, comparadas ao preço de fechamento de cada opção. Nesta tabela temos o valor do IDI na data inicial,

o tempo até a maturidade em anos, o valor de exercício do contrato (*strike*) e os preços calculados pelo método de kernel não paramétrico/HJM Gaussiano e, na última coluna, o valor de fechamento dessa opção neste pregão.

Os três valores apresentados em azul na tabela 6 representam as maturidades em que o erro de apreçamento das opções sobre o IDI foi menor. Já em vermelho são apresentados os três piores resultados. Dessas constatações, pode-se depreender que o preço de mercado nem sempre é bem aproximado por meio do modelo proposto. Assim, cabe ressaltar que o preço de mercado está sujeito a alterações nas expectativas, além de ter sido calculado por meio de um modelo aceito (normalmente Black and Scholes), e não reflete, necessariamente, o preço que a opção de fato deveria assumir. Por esse motivo, é natural que o modelo proposto não encontre sempre preços próximos aos de mercado. O bom ajuste do modelo é apresentado na seção de simulações desse trabalho, que compara o valor do apreçamento pelo modelo proposto com o valor teórico da opção.

	IDI	Maturidade	Strike	Preço Kernel	Preço Fechamento
1	137976.660	0.238	140800.000	0.726	16.090
2	137976.660	0.238	140900.000	0.519	0.230
3	137976.660	0.492	143800.000	50.288	75.210
4	137976.660	0.492	143900.000	46.728	36.290
5	137976.660	0.492	144000.000	43.392	14.680
6	137976.660	0.738	146400.000	239.083	387.490
7	137976.660	0.738	146500.000	231.225	305.290
8	137976.660	0.738	146800.000	208.964	113.070
9	137976.660	0.738	146900.000	201.966	72.360
10	137976.660	0.738	147000.000	195.170	43.260
11	137976.660	0.738	147100.000	188.572	24.040
12	137976.660	0.738	147200.000	182.168	12.360
13	137976.660	0.738	147300.000	175.953	5.850
14	137976.660	0.738	147400.000	169.922	2.550
15	137976.660	1.226	152500.000	867.312	882.140
16	137976.660	1.226	153000.000	810.415	590.780
17	137976.660	1.226	153300.000	777.828	447.850

Tabela 6: Apreçamento de opção de compra do IDI - Pregão 04/04/2012 - Método 3

As figuras 6-9 mostram alguns exemplo de funções de volatilidade estimadas pelas diversas metodologias de kernel para a série de DI. O dia da estimação em questão foi 15/10/2012, comparando diversos métodos e janelas de estimação. Pode-se observar que o modelo 1 com janela de 500 dias (figura 6) apresenta uma função de volatilidade com formato atípico se comparada com a literatura. No artigo de Mercurio e Moraleda (2000), por exemplo, há uma discussão acerca do formato típico de funções de volatilidade de estruturas a termo da taxa de juros, deixando claro que o formato arqueado (humped-

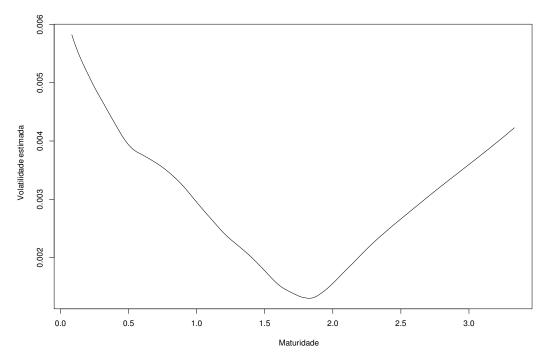


Figura 6: Função de volatilidade pelo modelo 1 - Dia 2012-15-10 - Amostra com janelas de 500 dias

shaped) é normalmente encontrado em trabalhos empíricos. De acordo com os autores, muitos trabalhos teóricos não captam essa característica realista da função. Podemos notar que, usando a janela de 500 com o modelo 4 (figura 7), conseguimos capturar esse padrão típico de função de volatilidade. No entanto, o formato da curva depende bastante da janela de estimação utilizada. Por exemplo, usando janelas mais curtas para o método 4 (figuras 8-9) o formato da função de volatilidade se altera bastante, capturando a mudança no padrão de volatilidade observado nas série de DI.

As tabelas 7 -10 mostram os resultados acumulados do apreçamento usando as quatro distintas amostras e cinco métodos, agregando os resultados de apreçamento de todas as opções de compra negociadas na BM&F nos últimos 300 pregões antes de 26/06/2013, com distintas maturidades e *strikes*. Com relação aos resultados obtidos, o primeiro importante ponto a ser mencionado é a magnitude dos valores apresentados. Sabe-se que os prêmios das opções sobre o IDI são bastante elevados, dado que os lotes de contratos DI1 são de alto valor. O lote padrão de negócios de DI1 é de 5 contratos, sendo que cada contrato equivale ao valor presente de R\$ 100.000,00, conforme discutido no início dessa seção. Assim, para algumas observações das tabelas, o erro pode ser percentualmente pequeno. Porém, conforme discutido, os valores de mercado das opções não necessariamente refletem o valor real que a opção deveria assumir, sendo que erros grandes não implicam em um mal ajuste do modelo.

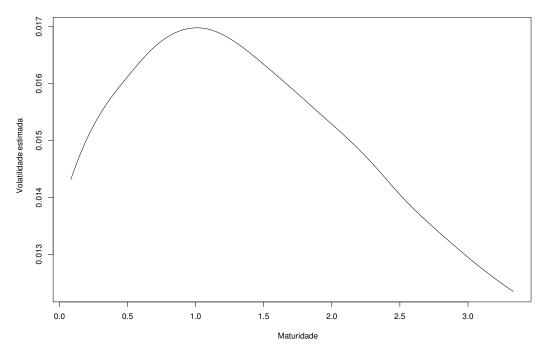


Figura 7: Função de volatilidade pelo modelo 4 - Dia 2012-15-10 - Amostra com janelas de 500 dias

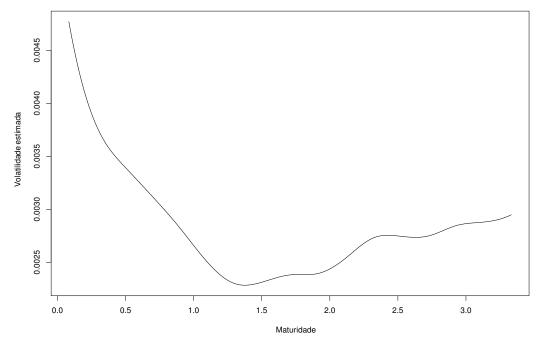


Figura 8: Função de volatilidade pelo modelo 4 - Dia 2012-15-10 - Amostra com janelas de 22 dias

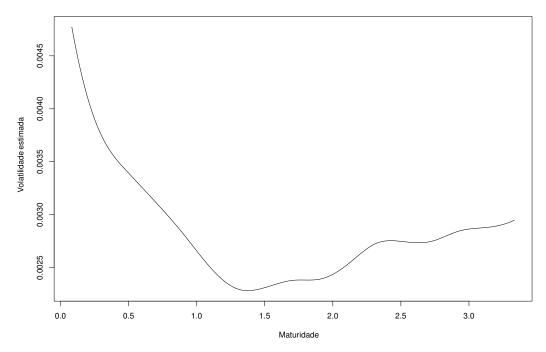


Figura 9: Função de volatilidade pelo modelo 4 - dia
2012-15-10 - Amostra com janelas de 125 dias

	EM	REQM	EAM
Modelo 1	-46.761166	254.829162	87.349967
Modelo 2	-45.023716	262.044923	90.160737
Modelo 3	60.006274	254.507967	60.499900
Modelo 4	59.998458	254.611075	60.512218
Modelo 5	60.088302	254.944957	60.578615

Tabela 7: Apreçamento de opções sobre IDI - Amostra com janelas de 22 dias

	EM	REQM	EAM
Modelo 1	-25.296524	238.437783	75.950551
Modelo 2	-2.900870	232.073306	66.116314
Modelo 3	8.061941	268.613400	76.666328
Modelo 4	5.460418	269.301026	77.834928
Modelo 5	8.027997	268.414380	76.665955

Tabela 8: Apreçamento de opções sobre IDI - Amostra com janelas de 125 dias

	EM	REQM	EAM
Modelo 1	-3.411557	219.023462	62.305738
Modelo 2	27.753802	220.838958	49.669030
Modelo 3	-146.713758	493.835845	198.140595
Modelo 4	-150.323672	497.891363	200.396225
Modelo 5	-146.590064	493.783066	198.203904

Tabela 9: Apreçamento de opções sobre IDI - Amostra com janelas de 250 dias

	EM	REQM	EAM
Modelo 1	-2.285512	212.014386	59.754696
Modelo 2	-108.768341	388.191697	166.246454
Modelo 3	-58.067898	336.785319	70.408238
Modelo 4	-58.811975	338.797761	70.996572
Modelo 5	-58.033748	336.679289	70.378564

Tabela 10: Apreçamento de opções sobre IDI - Amostra com janelas de 500 dias

Outros aspectos importantes a serem relatados são as medidas de ajuste apresentadas nas tabelas. O erro médio (EM) reflete o viés do modelo em relação aos preços de mercado das opções consideradas, enquanto a raiz do erro quadrático médio (REQM), considerada na análise das simulações, reflete não só o viés, mas a variância também. O erro absoluto médio (EAM), por sua vez, é a média dos módulos dos erros, de modo que não haja valores negativos. Com relação à seleção de janela de kernel, foi utilizado nessa parte empírica um kernel Gaussiano e o critério de validação cruzada por mínimos quadrados. Os resultados usando o método AIC são similares e, portanto, não são apresentados.

Para cada uma das medidas de erro mencionadas, é destacado em azul nas tabelas os melhores resultados. Pode-se observar que os modelos que performam melhor para aproximar os preços de mercado variam de acordo com a amostra e variam entre as medidas de erro também. Assim, para a amostra de 22 dias, o modelo 3 apresenta resultado mais aproximado ao mercado, de acordo com duas métricas de erro. Já com 125 dias, o modelo 2 performa melhor, sendo que para os outros dois tamanhos de amostra, 250 e 500 dias, o melhor resultado é o do modelo 1. Sendo assim, o modelo que parece aproximar melhor os valores de mercado de acordo com a análise realizada é o primeiro. Mas é importante analisar mais a fundo essas constatações, levando em consideração outros trabalhos da literatura também.

Conforme observado nos destaques das tabelas, comparando as diversas especificações da estimação de kernel, obtemos os melhores resultados para essa amostra usando o modelo 1 em termos de erro médio (amostra com 250 observações), e o modelo 2 obtém os melhores resultados em termos de raiz do erro quadrático e erro absoluto médio (amostra com 250 observações). É possível interpretar este resultado como a robustez destes dois métodos em uma curva que apresenta um formato bastante complexo, com mudanças claras no padrão de volatilidade no período observado.

Adicionalmente, pode-se observar dos resultados que, de uma maneira geral, os dois primeiros modelos tendem a diminuir seus erros à medida que a amostra de estimação aumenta. Uma clara exceção pode ser observada na tabela 10, em que os erros do modelo

2 são bastante elevados. Já com relação aos outros três modelos, a tendência geral é de aumento nos erros até a análise com 250 dias, sendo que quando a amostra é de 500 dias, esses valores tendem a se reduzir.

Estes efeitos podem ser explicados por dois efeitos. Um primeiro é a questão do tradeoff entre viés e variância de estimadores quando usamos um estimador que assume um
processo constante (no nosso caso assumindo um modelo HJM Gaussiano com função de
volatilidade determinista) quando o verdadeiro processo está mudando no tempo. Nesta
situação podemos minizar o viés usando uma amostra de estimação mais curta, mas isso
reduz a eficiência da estimação, já que usamos uma amostra menor. De forma geral
podemos verificar que os melhores resultados em termos de ajuste aos preços observados
de mercado são dados com a amostra de estimação de 125 dias, o que equivale a usar
aproximadamente seis meses para estimar as funções de volatilidade.

Embora não seja diretamente possível calcular os resultados obtidos com trabalhos anteriores na literatura de precificação, como Vieira e Valls (1999), Barbachan e Ornelas (2003), Gluckstern (2001) e Barbedo, Lion e Vicente (2010), devido ao nosso uso de um banco de dados mais recente, e as diversas métricas usadas em cada trabalho, algumas comparações qualitativas gerais podem ser feitas. Os modelos calibrados usando janelas de 22 e 125 dias apresentam um bom ajuste aos dados, e de forma geral o viés e erro quadrático médio são bastante baixos. Isso contrasta com os resultados encontrados em Barbachan e Ornelas (2003) e Barbedo, Lion e Vicente (2010), que encontram erros sistemáticos de previsão, com subapreçamento sistemático ao utilizar os modelos Vasicek e Cox-Ingersoll-Ross e superapreçamento para os modelos HJM calibrado por PCA. Embora novamente seja uma amostra diferente, e com isso os valores nocionais dos contratos sejam distintos, nossos resultados com as janelas de estimação de 22 e 125 dias também mostram um resultado médio em termos de erros absoluto médios sistematicamente inferiores aos obtidos em Gluckstern (2001) que comparou modelos de Black, Hull-White e Black-Karasinski no apreçamento de opções sobre o IDI.

5.4 Conclusões

O presente trabalho teve por objetivo aplicar uma técnica não paramétrica no apreçamento de opções sobre os contratos relacionados à estrutura a termo de taxas de juros, mostrando os resultados desse apreçamento sobre as opções do IDI. A metodologia utilizada foi basicamente a proposição de cinco modelos de regressão de kernel para encontrar a função de volatilidade de estruturas a termo de taxas de juros, e a aplicação destas funções de volatilidade no apreçamento de derivativos usando modelos HJM Gaussianos. Assim, adotando a estrutura de não arbitragem da classe de modelos de Heath-Jarrow-Morton, o apreçamento de opções sobre títulos de renda fixa depende basicamente da modelagem da função de volatilidade do ativo objeto.

Nosso trabalho contribui de duas formas para a literatura de precificação não paramétrica de opções. A primeira é analisando diferentes especificações possíveis para estimação da função de volatilidade, propondo formas cross-section e usando das séries temporais da curva de juros, e assim generalizando a formulação proposta em Jeffrey et al. (2004), que usa somente os valores passados das taxas como regressores na especificação de kernel. Adicionalmente discutimos um modificação de Jeffrey et al. (2004), usando a ponderação ótima derivada da estimação não paramétrica de matrizes de variância de longo prazo, mais adequadas a processos dependentes e possivelmente heterogêneos.

A segunda contribuição do trabalho é verificar o desempenho dos estimadores não paramétricos da função de volatilidade no apreçamento de derivativos, uma análise que está ausente em Jeffrey et al. (2004). Neste objetivo utilizamos uma versão analiticamente simples da classe de modelos HJM, assumindo uma função de volatilidade determinista e assim levando a um modelo HJM Gaussiano. Esta forma evita as dificuldades computacionais existentes em modelos não homogêneos, que necessitam de calibração de árvores não recombinantes ou então de precificação via simulações de Monte Carlo. Embora esta classe seja uma forma restrita da classe geral de modelos HJM, mostramos que seus resultados no apreçamento de opções são bastante robustos, mesmo quando o processo gerador dos dados está associado a funções de volatilidade variantes no tempo, como analisado na seção de estudos de simulação e a análise empírica no apreçamento de opções sobre o IDI.

A busca pelo refinamento do cálculo das opções sobre o IDI é mais um ponto que deve ser destacado como contribuição deste trabalho. Conforme discutido, um bom ajuste no apreçamento desse derivativo pode auxiliar instituições que necessitem desenvolver *hedge* em posições atreladas ao índice, fato que serve de motivação para o estudo de métodos que contribuam com essa literatura.

Referências

- AIT-SAHALIA, Y. Nonparametric pricing of interest rate derivative securities. *Econometrica*, v. 64, p. 527–560, 1996.
- ANDERSEN, L. F. B.; PITERBARG, V. V. Interest Rate Modeling, Volume I: Foundations and Vanilla Models. [S.l.]: Atlantic Financial Press, 2010.
- ANDREWS, D. W. K. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, v. 59, p. 817–858, 1991.
- BACK, K. E. Asset Pricing and Portfolio Choice Theory. [S.l.]: Oxford University Press, 2010.
- BARBACHAN, J. S. F.; ORNELAS, J. R. H. Apreçamento de opções de idi usando o modelo cir. *Estudos Econômicos*, v. 33, p. 287–323, 2003.
- BARBEDO, C. H.; LION, O. B.; VICENTE, J. V. Pricing asian interest rate options with a three-factor hjm model. *Revista Brasileira de Finanças*, v. 8, p. 9–13, 2010.
- BRACE, A.; MUSIELA, M. Swap derivatives in a gaussian HJM framework. *Preprint*, *University of New South Wales*, 1994.
- CHAN, K. C. G. et al. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance*, v. 67, p. 1209–1227, 1992.
- FILIPOVIC, D. Term Structure Models: A graduate course. [S.l.]: Springer, 2009.
- FLORENS, J.; MARIMOUTOU, V.; PEGUIN-FEISSOLLE, A. Econometric Modelling and Inference. [S.l.]: Cambridge University Press, 2007.
- GLUCKSTERN, M. C. Aplicação do modelo de Hull-White à precificação de opções sobre IDI. Tese (Doutorado) EAESP Fundação Getulio Vargas, 2001.
- HANSEN, B. E. Lecture notes on nonparametrics. University of Winconsin. 2009.
- HEATH, D.; JARROW, R.; MORTON, A. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econ*, v. 60, p. 77–105, 1992.
- HURVICH, C. M.; SIMONOFF, F. S.; TSAI, C.-L. Smoothing parameter selection in nonparametric regression using an improved akaike information criterion. *Journal Royal Statistical Society*, v. 60, p. 271–293, 1998.
- JARROW, R. A. Modeling Fixed-Income Securities and Interest Rate Options. [S.l.]: Stanford Economics and Finance, 2002.

JEFFREY, A. et al. Nonparametric estimation of a multifactor heath-jarrow-morton model: An integrated approach. *Journal of Financial Econometrics*, v. 2, n. 2, p. 251–289, 2004.

LAURINI, M. P.; OHASHI, A. A noisy principal component analysis for forward rate curves. June 2013.

LI, Q.; RACINE, J. F. Cross-validated local linear nonparametric regression. *Statistica Sinica*, v. 14, p. 485–512, 2004.

MERCURIO, F.; MORALEDA, F. M. An analytically tractable interest rate model with humped volatility. *European Journal of Operational Research*, v. 120, p. 205–214, 2000.

MUNK, C. Fixed Income Analysis: Securities, Pricing and Risk Management. [S.l.: s.n.], 2003.

MUSIELA, M.; MAREK, R. Martingale Methods in Financial Modelling. [S.l.]: Springer, 2005.

NAWALKHA, S. K.; BELIAEVA, N. A.; SOTO, G. M. Dynamic Term Structure Modeling. [S.l.]: John Wiley & Sons, Inc., 2007.

OKSENDAL, B. Stochastic Differential Equations: An Introduction with Applications. [S.l.]: Springer, 2003.

RACINE, J. S.; LI, Q. *Nonparametric Econometrics*. [S.l.]: Princeton University Press, 2007.

STEHLIKOVA, B.; SEVCOVIC, D. Mathematical Analysis of Term Structure Models. Tese (Doutorado) — Faculty of Mathematics, Physics and Informatics Comenius University Bratislava, 2008.

VALLS, P. L. V.; VIEIRA, C. A. Review of major results of martingale theoryapplied to the valuation of contingent claims. *Brazilian Review of Econometrics*, v. 21, n. 2, p. 120–135, 2001.

VALLS, P. L. V.; VIEIRA, C. A. Uma resenha sobre os principais resultados da teoria de martingais aplicada à avaliação de títulos contigentes. 2002.

VIEIRA, C. A. Modelagem da Estrutura a Termo da Taxa de Juros e Avaliação de Contratos Derivativos. Tese (Doutorado) — Universidade de São Paulo - Faculdade de Economia, Administração e Contabilidade, 1999.

VIEIRA, C. A.; VALLS, P. L. V. Derivação de uma fórmula para cálculo do preço livre de arbitragem de opções sobre idi. $Resenha\ BmF$, n. 139, p. 37–42, 1999.

ZEILEIS, A. Econometric computing with hc and hac covariance matrix estimators. Journal of Statistical Software, v. 11, 2004.

ZIVOT, E.; WANG, J. Modelling Financial Time Series with S-Plus. [S.l.]: Springer, 2003.