

# O Prêmio pela Maturidade na Estrutura a Termo das Taxas de Juros Brasileiras\*

Ricardo D. Brito<sup>†</sup>      Angelo Jose Mont' Alverne Duarte<sup>‡</sup>

Osmani Teixeira de Carvalho Guillén<sup>§</sup>

Fevereiro de 2003

Preliminar

## Abstract

Este artigo estuda a validade da Hipótese das Expectativas (HE) racionais para o Brasil, utilizando dados diários de Julho de 1996 a Dezembro de 2002 para prazos entre 1 dia e 1 ano. (i) Os coeficientes do diferencial rendimento (*yield spread*) estimados na equação de mudança de curto prazo da taxa de longo prazo são imprecisos e não permitem conclusão. (ii) A estimação da equação de mudança de curto prazo da taxa curta indica que o diferencial tem poder preditivo. (iii) Os coeficientes do diferencial na equação de mudança de longo prazo da taxa curta são mais precisamente estimados e não são significativamente diferentes da unidade. (iv) A previsão de expectativas racionais da mudança de longo prazo da taxa curta é altamente correlacionada com o diferencial, e (iv) o diferencial observado é mais volátil que o diferencial teórico. Os resultados dos testes rejeitam a hipótese particular coeficiente unitário, porém reconhecem o poder preditivo do diferencial observado.

Palavras Chave: estrutura a termo das taxas de juros, hipótese das expectativas, prêmio pelo prazo.

Classificação do JEL: E43, G13.

---

\*Gostaríamos de agradecer a Eurilton Araújo (Ibmec) e Fernando Garcia (EAESP/FGV-SP) pelas sugestões e comentários e aos colegas Marcelo Zeuli e Ricardo Maia Clemente (BACEN) pela assistência na coleta dos dados e montagem das séries.

<sup>†</sup>IBMEC.

<sup>‡</sup>Banco Central do Brasil e EPGE/FGV-RJ, 21 2559-6000 ramal 9798, [aduarte@fgvmail.br](mailto:aduarte@fgvmail.br)

<sup>§</sup>Banco Central do Brasil e EPGE/FGV-RJ.

# 1 Introdução

A relação entre as taxas de juros de curto e longo prazos, ou o processo de formação da estrutura a termo das taxas de juros, interessa tanto aos participantes do mercado preocupados em inferir o valor das oportunidades de investimento, quanto às autoridades monetárias preocupadas em monitorar as expectativas dos agentes.

A mais conhecida teoria sobre estrutura a termo das taxas de juros, a Hipótese das Expectativas (HE), afirma que uma taxa de juros à vista de longo prazo é a média de longo prazo das expectativas das futuras taxas à vista de curto prazo mais um “prêmio pelo prazo” (*term premium*) invariante no tempo:

$$R_t^{(n)} = \frac{1}{k} \sum_{i=0}^{k-1} E_t R_{t+mi}^{(m)} + \psi_k, \quad (1)$$

onde:  $R_t^{(n)}$  é a taxa de juros de longo prazo com  $n$  períodos,  $R_t^{(m)}$  é a taxa de juros de curto prazo de  $m$  períodos, ambas tomadas em nível,  $k = n/m$  é um inteiro maior que a unidade,  $E_t$  é a esperança condicionada a informação conhecida na data  $t$ , e a constante  $\psi_k$  é o prêmio pelo prazo. Se HE vale, ou seja, se o prêmio pelo prazo é constante, é fácil entender a relação entre as expectativas das futuras taxas e as taxas correntes. A inclinação da estrutura a termo de juros à vista é determinada pela expectativa do mercado quanto aos futuros valores dos juros à vista e mudanças desta expectativa refletem-se em movimentos da curva. Se, por outro lado, o prêmio pelo prazo varia ao longo do tempo de forma imprevisível, fica difícil discernir entre mudanças das taxas à vista causadas por revisão de expectativa das futuras taxas ou mudanças das taxas à vista causadas por variações dos prêmios pelo prazo.

Em acordo com o paradigma de Mercados Eficientes, a literatura empírica para os EUA e Europa testa a HE sob expectativas racionais. Por exemplo, Mankiw e Miron (1986) utilizam títulos do tesouro americano de três e seis meses para testar a validade de HE para o período de 1890 a 1979. Rejeitam a HE para todos os subperíodos, exceto 1890-1914, anterior a organização do FED. Campbell e Shiller (1991) examinam dados do pós-guerra da estrutura a termo americana e encontram um comportamento inconsistente com a HE. Para qualquer maturidade entre um mês e dez anos, um diferencial de rendimento (*yield spread*) elevado entre a taxa longa e a curta prevê um crescimento de longo prazo da taxa de curto prazo e uma redução de curto prazo da taxa de longo prazo. Mais recentemente, trabalhando com séries de *REPO-agreement* (operações de venda com compromisso de recompra) como proxies para as taxas sem risco de curto prazo americanas, Longstaff (2000) não rejeita a HE para vencimentos até 3 meses. Mankiw (1986), utilizando dados dos Estados Unidos, Canadá, Reino Unido e Alemanha, rejeita a HE ao testar várias implicações derivadas desta mesma. O autor reforça suas conclusões mostrando que variações de volatilidade dos títulos, variações no risco não-diversificável, ou variações na oferta de ativos não podem explicar satisfatoriamente a grande variação no prêmio pela maturidade. Hardouvelis (1994) estuda o comportamento das taxas de juros dos países do G7 e, utilizando variáveis instrumentais, consegue reverter a correlação negativa entre o diferencial de rendimento da estrutura a termo e a variação da taxa de longo prazo. Em suma, HE é quase sempre rejeitada para os EUA e não-rejeitada para outros países do G7 <sup>1</sup>.

A freqüente rejeição de HE fez crescer o interesse pelas formulações de prêmio pelo prazo variável. Mankiw (1986) sugere que a taxa longa é observada com ruído, Modigliani

---

<sup>1</sup>Ver Anderson et al. (1996) Cap. 9 para um resumo da evidência empírica internacional sobre a HE.

e Sutch (1966) apontam para as variações na oferta de títulos longos motivadas pela política de gestão da dívida pública, e Engle et al. (1987) elaboram um risco variável no tempo. Uma alternativa ao prêmio pelo prazo variável é que a falha de HE seja devida a persistentes erros de expectativa. Froot (1989) utiliza pesquisas sobre expectativas de taxas de juros para mostrar a relevância dos erros de expectativas sistemáticos em horizontes longos. Campbell e Shiller (1991) sugerem uma sobre-reação das taxas de longo prazo às expectativas das futuras taxas curtas.

No Brasil, a literatura sobre o tema é incipiente, assim como é recente a formação de uma estrutura a termo passível de teste. Tabak e Andrade (2001) estudam a HE sob expectativas racionais para a estrutura a termo brasileira com dados diários e prazos entre dois meses e doze meses. Utilizando o diferencial de rendimento defasado como instrumento para o diferencial contemporâneo, capturam uma dependência temporal do prêmio pela maturidade e concluem pela rejeição de HE sob expectativas racionais. Lima e Issler (2002) testam HE sob expectativas racionais para dados de frequência mensal e prazos de um mês, 180 dias e 360 dias no contexto dos modelos de valor presente desenvolvidos em Campbell e Shiller (1987) e Campbell e Shiller (1991). Após testarem as implicações do modelo de valor presente concluem que as evidências são apenas parcialmente favoráveis a HE <sup>2</sup>.

No presente artigo construímos a estrutura a termo brasileira de um dia a um anos examinamos a HE para o Brasil como em Campbell e Shiller (1991) e Hardouvelis (1994): (i) por Mínimos Quadrados Ordinários (MQO) estimamos o coeficiente angular do diferencial de rendimento (*term spread*) na equação de mudança de curto prazo da taxa de longo prazo e na equação de mudança de curto prazo da taxa de curto prazo; (ii) estimamos por MQO coeficiente angular do diferencial de rendimento na equação de mudança de longo prazo da taxa de curto prazo; e (iii) utilizando um modelo VAR, calculamos a previsão de expectativas racionais da mudança de longo prazo da taxa de curto prazo. Uma vez documentado o aparente desvio da HE, tentamos racionalizar os resultados em termos (iv) da presença de ruído branco nas taxas observadas reestimando (i) por Variáveis Instrumentais (VI); ou (v) de reação exagerada (*overreaction*) das taxas de longo prazo.

O presente artigo encontra-se dividido em seis seções, incluindo esta introdução. A segunda seção indica os procedimentos adotados na construção da curva de juros brasileira e faz uma análise descritiva da mesma. O suporte teórico utilizado neste trabalho é brevemente discutido na terceira seção. Na quarta seção testamos a HE para o Brasil. Na quinta seção aventamos motivos para a rejeição da HE. Finalmente, na sexta seção, as conclusões são elencadas.

## 2 Análise dos Dados e Teste Preliminar

Na figura 1 apresentamos as taxas à vista anualizadas para todos os vencimentos utilizados em frequência diária, calculadas para o período de 01/07/1996 a 31/12/2001 segundo descrito no Apêndice I. Podem ser visualizados três pontos importantes: as crises da Ásia em outubro de 1997, da Rússia em agosto de 1998 e por último, em janeiro de 1999, a mudança de regime cambial no Brasil.

---

<sup>2</sup>Tabak e Andrade (2001) e Lima e Issler (2002) serão analisados em mais detalhe na seção ?.

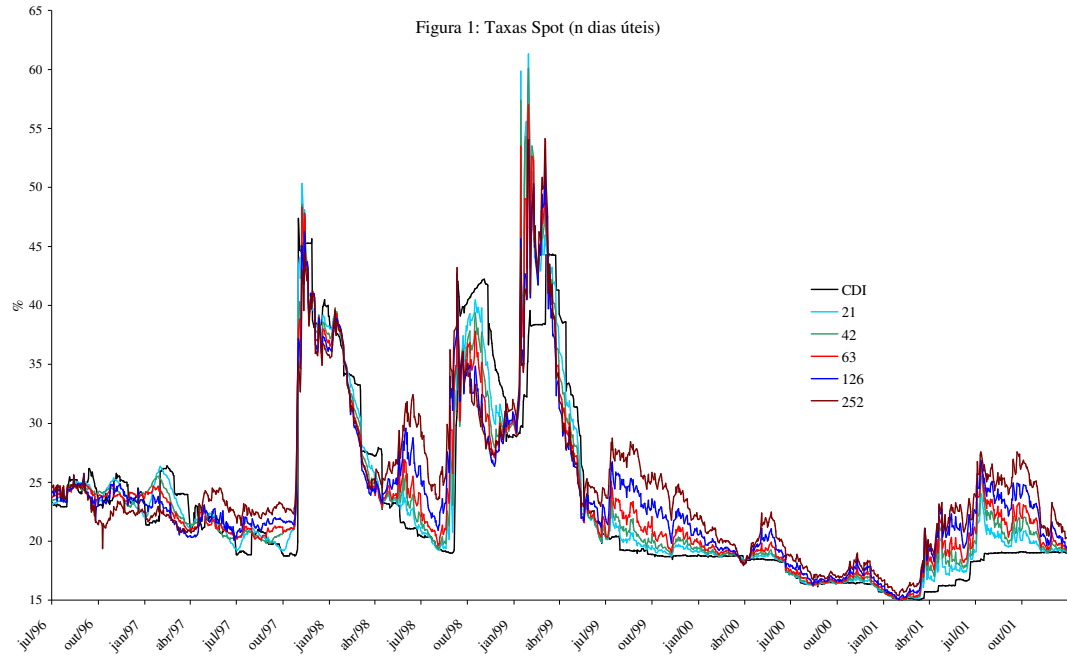


Tabela 1  
Estatísticas Descritivas das Taxas à Vista Anualizadas Brasileiras de 01/07/1996 até 31/12/2001

Estatísticas	Taxas em nível						Diferença diária das taxas					
	1	21	42	63	126	252	1	21	42	63	126	252
Média	.2055	.2086	.2092	.2108	.2143	.2208	.0000	.0000	.0000	.0000	.0000	.0000
Méd.Segunda	.2052	.2079	.2085	.2102	.2137	.2202	-.0009	-.0002	.0000	.0001	.0000	.0000
Méd.Terça	.2064	.2081	.2086	.2102	.2137	.2201	-.0001	.0002	.0002	.0002	.0001	.0001
Méd.Quarta	.2050	.2082	.2087	.2103	.2140	.2206	-.0001	.0000	-.0001	-.0001	.0000	.0000
Méd.Quinta	.2049	.2096	.2103	.2119	.2153	.2219	.0002	.0001	.0001	.0000	.0000	.0000
Méd.Sexta	.2059	.2093	.2098	.2112	.2145	.2212	.0001	.0000	.0000	.0000	.0000	.0000
Desvio Padrão	.0590	.0591	.0567	.0540	.0505	.0498	.0074	.0082	.0079	.0075	.0068	.0065
D.P.Segunda	.0586	.0580	.0555	.0528	.0492	.0485	.0155	.0146	.0140	.0135	.0127	.0121
D.P.Terça	.0596	.0569	.0544	.0519	.0490	.0484	.0164	.0131	.0135	.0134	.0133	.0131
D.P.Quarta	.0588	.0587	.0565	.0537	.0504	.0496	.0161	.0148	.0151	.0148	.0150	.0147
D.P.Quinta	.0582	.0612	.0589	.0560	.0522	.0514	.0162	.0201	.0196	.0190	.0171	.0165
D.P.Sexta	.0603	.0612	.0587	.0559	.0522	.0515	.0191	.0169	.0158	.0153	.0147	.0140
$\rho_1$	.9920	.9900	.9900	.9900	.9910	.9910	-.0250	-.0230	.0110	.0430	.1100	.1190
$\rho_2$	.9850	.9810	.9800	.9800	.9800	.9800	.0200	-.0580	-.0680	-.0880	-.1130	-.1200
$\rho_3$	.9580	.9620	.9600	.9590	.9570	.9580	.0030	.1530	.1450	.1210	.0630	.0520
$\rho_4$	.9070	.9130	.9140	.9130	.9100	.9140	.0140	.0320	.0470	.0410	.0330	.0260
$\rho_5$	.7880	.7930	.7910	.7850	.7740	.7810	-.0200	.0330	.0400	.0300	.0280	.0260
Teste ADF	-2.58	-2.68	-2.78	-2.97*	-3.07*	-3.07*	-8.79**	-8.2**	-7.92**	-7.56**	-7.5**	-7.38**
Teste KPSS	1.89**	1.68**	1.59**	1.53**	1.44**	1.30**	0.03	0.04	0.04	0.04	0.03	0.04

Notas:

- (i) Tamanho das amostras em nível: 1380 observações para Média, 275 para Segunda, 276 para Terça, 283 para Quarta, 271 para Quinta e 275 para Sexta.
- (ii) Tamanho das amostras em diferença: 1379 observações para Média, 263 para Segunda, 266 para Terça, 280 para Quarta, 256 para Quinta e 264 para Sexta.
- (iii) ADF testa  $H_0$ : série não-estacionária. Para nível diário incluiu intercepto e 20 diferenças defasadas. Para diferença diária incluiu 19 diferenças defasadas.
- (iv) KPSS testa  $H_0$ : série estacionária. Para nível diário utiliza janela de 21 defasagens. Para diferença diária utiliza 20 defasagens.
- (v) \*(\*\*) indica rejeição de  $H_0$  ao nível de significância de 5% (1%). Para o ADF, valor crítico a 5% (1%) igual a -2.86 (-3.44). Para o KPSS, valor crítico a 5% (1%) igual a 0.46 (0.74).

A Tabela 1 apresenta algumas estatísticas descritivas do nível e da variação das taxas de juros à vista anualizadas para o período. Semelhantemente à evidência internacional, observa-se que a estrutura a termo foi em média positivamente inclinada, com maior volatilidade das taxas mais curtas. Ao contrário dos EUA, não se observam efeitos dia da semana, como dias específicos nos quais as taxas são significativamente mais altas ou

mais voláteis. A autocorrelação elevada adverte que a informação disponível na amostra é efetivamente menor do que pode indicar o seu tamanho (1380 observações). A não-estacionariedade das séries foi testada pelo procedimento Augmented Dickey-Fuller (ADF) cuja hipótese nula é não-estacionariedade e pelo procedimento de Kwiatkowski et al. (1992) (KPSS) que assume série estacionária sob a hipótese nula. Ambos os testes oferecem evidências de séries não-estacionárias no nível e estacionárias na diferença. Note-se porém que no teste ADF os prazos de 126 e 252 dias rejeitam a 5% de significância a hipótese nula.

Embora a não-estacionariedade das séries pareça questionável à luz da teoria econômica e a baixa potência dos testes de raiz unitária seja conhecida, muitos trabalhos assumem não-estacionariedade do nível das taxas e optam pela modelagem da diferença das mesmas. Se este é o caso, a cointegração com coeficiente unitário das taxas longas e curtas é condição necessária para que HE valha. A condição não é suficiente porque a cointegração requer apenas que os erros de expectativas e o prêmio pelo prazo sejam estacionários. Ou seja, a cointegração é consistente com um prêmio pelo prazo variável no tempo.

A Tabela 2 apresenta o teste de cointegração de Johansen com intercepto e sem tendência para os vários vértices. Nesta tabela e nas que se seguem, os índices  $m$  e  $n$  representam o prazo, medido em dias úteis, das taxas curta e longa, respectivamente. As 20 diferenças defasadas utilizadas foram suficientes para gerar resíduos não-autocorrelacionados, porém heterocedásticos e leptocúrticos. Observa-se que a cointegração é mais forte entre as taxas mais curtas (direção noroeste) e que perde força a medida que crescem as maturidades (direção sudeste). Exceto para o vértice mais longo ( $n = 252$  e  $m = 126$ ), não se pode rejeitar a hipótese de cointegração. A maior dificuldade em constatar cointegração das taxas mais longas na amostra pode ser devida à menor velocidade de convergência das mesmas relativamente as taxas curtas. De fato, é sempre a taxa mais curta quem maior peso coloca no vetor de correção de erro (não apresentado). Se este é o caso, uma amostra maior permitirá a detecção de cointegração das taxas mais longas inclusive <sup>3</sup>.

A hipótese de coeficiente de cointegração unitário não parece pouco razoável pelo exame das estimativas pontuais. Utilizando os erros-padrão de White (1980) sugeridos por Hansen(1992), obtemos estatísticas  $t$  abaixo de 1,96 para sete vértices ou abaixo de 2,33 para 11 vértices. Se levamos em conta a leptocurtose dos resíduos, fica difícil rejeitar a hipótese da cointegração com coeficiente unitário <sup>4</sup>. A quinta linha de cada vértice apresenta o intercepto do vetor de cointegração, indicando um prêmio não-significativo em todos os casos.

---

<sup>3</sup>Uma hipótese alternativa, sustentável pelo resultado deste teste e do ADF (na Tabela 1) para os vértices  $n = 252$  e  $m = 126$ , é que as taxas podem ser estacionárias no nível.

<sup>4</sup>Uma forma mais rigorosa de proceder a análise de significância dos coeficientes de cointegração seria computar os erros-padrão por um procedimento Monte-Carlo. Contudo, o custo computacional não pareceu compensar a evidência fraca eventualmente oferecida pelo teste de cointegração.

Tabela 2  
Teste de cointegração de Johansen (sem tendência)

n	m				
	1	21	42	63	126
21	37.99**				
	7.18				
	-1.00				
	(0.06)				
	0.00				
	(0.01)				
42	39.56**	53.11**			
	7.49	7.90			
	-0.95	-0.95			
	(0.06)	(0.02)			
	-0.01	-0.01			
	(0.01)	(0.00)			
63	36.86**	42.13**	33.46**		
	7.70	7.99	7.67		
	-0.88	-0.89	-0.93		
	(0.06)	(0.04)	(0.03)		
	-0.03	-0.03	-0.02		
	(0.01)	(0.01)	(0.01)		
126	34.06**	33.42**	29.92**	29.13**	
	7.92	8.20	8.14	9.05	
	-0.79	-0.79	-0.83	-0.91	
	(0.07)	(0.09)	(0.08)	(0.06)	
	-0.05	-0.05	-0.04	-0.02	
	(0.01)	(0.02)	(0.02)	(0.01)	
252	28.02**	26.33**	23.89*	23.42*	20.29*
	7.61	7.94	8.12	9.07	9.43*
	-0.75	-0.75	-0.78	-0.87	-1.00
	(0.10)	(0.14)	(0.15)	(0.12)	(0.05)
	-0.07	-0.06	-0.06	-0.04	-0.01
	(0.02)	(0.03)	(0.03)	(0.02)	(0.01)

Notas:

(i) De cima para baixo: LR(nenhum vetor), LR(máximo 1 vetor), coeficiente de cointegração e erro-padrão de White (1980), segundo Hansen (1992).

(ii) Permite-se intercepto no vetor de cointegração e 20 diferenças defasadas.

(iii) \*(\*\*) indica rejeição da hipótese ao nível de significância de 5%(1%).

(iv) Para LR(nenhum vetor) valor crítico a 5% (1%) igual a 19.96 (24.60).

(v) Para LR(máximo 1 vetor) valor crítico a 5% (1%) igual a 9.24 (12.97).

Nossa interpretação é que a Tabela 2 não fornece evidência suficiente para rejeitar a cointegração com coeficiente unitário e outras implicações de HE merecem exame. Tais implicações são derivadas na seção seguinte e testadas na seção 4, para onde pode se dirigir o leitor familiarizado com a literatura de teste da HE sob expectativas racionais.

### 3 Discussão Teórica

Denominando  $n$  e  $m$  respectivamente por prazo longo e prazo curto e rerepresentando a equação (1) na sua forma geral:

$$R_t^{(n)} = \frac{1}{k} \sum_{i=0}^{k-1} E_t R_{t+mi}^{(m)} + \psi_{k,t}, \quad k = n/m, \quad (2)$$

onde um prêmio pelo prazo longo (*term premium*) dependente do tempo,  $\psi_{k,t}$ , é a média dos próximos  $k$  valores esperados dos prêmios pelo carregamento por prazo curto do título longo (*m-period holding premia*),  $\{E_t \varphi_{k,t+mi}\}_{i=0}^{k-1}$ :

$$\psi_{k,t} = \frac{1}{k} \sum_{i=0}^{k-1} E_t \varphi_{k,t+mi}; \quad (3)$$

observa-se que a taxa de longo prazo é a soma do prêmio pelo prazo com a média das  $k$  taxas de curto prazo esperadas entre a data  $t$  e  $t + n$ .

De forma alternativa, pode-se apresentar (2) para uma taxa de juros de  $(n-m)$  períodos na data  $t + m$ ,

$$R_{t+m}^{(n-m)} = \frac{m}{n-m} \sum_{i=0}^{k-2} E_{t+m} \left[ R_{t+m+mi}^{(m)} + \varphi_{k,t+m+mi} \right], \quad (4)$$

multiplicá-la por  $(n-m)$  e subtrair de (2) multiplicada por  $n$ , para obter o retorno realizado pelo carregamento por prazo curto do título longo (*m-period holding return*):

$$nR_t^{(n)} - (n-m)R_{t+m}^{(n-m)} = m \left( R_t^{(m)} + E_t \varphi_{k,t} - \xi_{t+m} \right), \quad (5)$$

onde:  $\xi_{t+m} = \sum_{i=1}^{k-1} \left[ \left( E_{t+m} R_{t+mi}^{(m)} - E_t R_{t+mi}^{(m)} \right) + \left( E_{t+m} \varphi_{k,t+mi} - E_t \varphi_{k,t+mi} \right) \right]$  representa as perdas (ganhos) oriundas das revisões de expectativas das futuras taxas de curto prazo e dos futuros prêmios de carregamento.

Na equação (5),  $m(E_t \varphi_{k,t} - \xi_{t+m})$  representa o excesso de retorno entre duas estratégias de investimento: na primeira compra-se um título de longo prazo para revendê-lo ao final do prazo curto, e na segunda compra-se um título de curto prazo para carregá-lo até o vencimento<sup>5</sup>.

A equação (5) pode ainda expressar a variação de curto prazo do rendimento do título de longo prazo,  $R_{t+m}^{(n-m)} - R_t^{(n)}$ , subtraindo  $mR_t^{(n)}$  de ambos os lados e multiplicando por  $-1/(n-m)$ :

$$R_{t+m}^{(n-m)} - R_t^{(n)} = \frac{m}{n-m} \left( R_t^{(n)} - R_t^{(m)} - E_t \varphi_{k,t} + \xi_{t+m} \right), \quad (6)$$

ou:

$$R_{t+m}^{(n-m)} - R_t^{(n)} = s_t^{(n,m)} - \frac{m}{n-m} E_t \varphi_{k,t} + \frac{m}{n-m} \xi_{t+m}, \quad (7)$$

onde:  $s_t^{(n,m)} = (m/(n-m)) S_t^{(n,m)} = (m/(n-m)) \left( R_t^{(n)} - R_t^{(m)} \right)$  é um múltiplo do diferencial de rendimento entre os títulos longo e curto  $S_t^{(n,m)}$  (*yield spread*).

A equação (7) permite testar a HE sob expectativas racionais. Uma vez que o prêmio esperado pelo carregamento é constante sob HE,  $E_t \varphi_{k,t} = \varphi_k$ , e as futuras revisões de expectativas são imprevisíveis sob expectativas racionais,  $E_t [\xi_{t+m}] = 0$ , a variação esperada de curto prazo do rendimento do título de longo prazo é dado por:

$$E_t R_{t+m}^{(n-m)} - R_t^{(n)} = s_t^{(n,m)} - \frac{m}{n-m} \varphi_k. \quad (8)$$

Ou seja, o diferencial de rendimento,  $S_t^{(n,m)}$ , prevê a mudança de curto prazo da taxa longa,  $\left( R_{t+m}^{(n-m)} - R_t^{(n)} \right)$ . A intuição é que um aumento de curto prazo do rendimento do título longo provoca perda de capital e o prêmio  $\left( S_t^{(n,m)} - \varphi_k \right)$  é a compensação pela perda esperada.

Se a equação (8) vale, ou seja se HE sob expectativas racionais é válida, a regressão de  $R_{t+m}^{(n-m)} - R_t^{(n)}$  em  $s_t^{(n,m)}$ :

<sup>5</sup>Ou de outras duas estratégias: numa compra-se um título de longo prazo para carregá-lo até o vencimento; e noutra compra-se um título de curto prazo e outro título de prazo  $(n-m)$  após o vencimento do primeiro.

$$R_{t+m}^{(n-m)} - R_t^{(n)} = \alpha + \beta \cdot s_t^{(n,m)} + u_{t+m}, \quad (9)$$

onde  $u_{t+m}$  é  $MA(m-1)$ , deve resultar num coeficiente angular igual a unidade,  $H_0 : \beta = 1$ .

Uma outra implicação de HE sob expectativas racionais sobre (2) é que subtraindo  $R_t^{(m)}$  de ambos os lados e rearranjando:

$$R_t^{(n)} - R_t^{(m)} = E_t \left[ \frac{1}{k} \sum_{i=1}^{k-1} \left[ \sum_{j=1}^i \left( R_{t+mj}^{(m)} - R_{t+m(j-1)}^{(m)} \right) \right] \right] + \psi_k; \quad (10)$$

ou:

$$S_t^{(n,m)*} = E_t S_t^{(n,m)*} + \psi_k, \quad (11)$$

onde,

$$\begin{aligned} S_t^{(n,m)*} &= \frac{1}{k} \sum_{i=1}^{k-1} \left[ \sum_{j=1}^i \left( R_{t+mj}^{(m)} - R_{t+m(j-1)}^{(m)} \right) \right] \\ &= \sum_{i=1}^{k-1} \left( 1 - i/k \right) \left( R_{t+mi}^{(m)} - R_{t+m(i-1)}^{(m)} \right), \end{aligned} \quad (12)$$

é uma média pondera das próximas  $(k-1)$  mudanças de curto prazo da taxa curta. De fato, se é esperada uma alta de longo prazo das taxas curtas, o rendimento atual do título longo deve ser superior ao rendimento atual do título curto, como forma de equalizar o retorno até o vencimento do primeiro com o rendimento da seqüência de  $k$  investimentos em títulos de curto prazo entre as datas  $t$  e  $t+n$ . A variável  $S_t^{(n,m)*}$  pode ser denominado diferencial de rendimento de antecipação perfeita (*perfect foresight spread*), pois, a menos da constante  $\psi_k$ , seria o diferencial de rendimento entre os títulos longo e curto caso a previsão das futuras taxas curtas fosse perfeita.

Se a equação (11) vale, ou seja se HE sob expectativas racionais é válida, a regressão de  $S_t^{(n,m)*}$  em  $S_t^{(n,m)}$ :

$$S_t^{(n,m)*} = \gamma + \beta \cdot S_t^{(n,m)} + v_{t+m}, \quad (13)$$

onde  $v_{t+m}$  é  $MA(n-m-1)$ , deve resultar num coeficiente angular igual a unidade,  $H_0 : \beta = 1$ . Adicionalmente, o intercepto pode ser interpretado como o negativo de prêmio pelo prazo longo  $\gamma = -\psi_k$ .

Apesar de simples, os testes da HE baseados em (8) e (11) têm alguns problemas: (i) possuem erros justapostos difíceis de corrigir quando  $n$  é grande relativamente ao tamanho da amostra (Richardson e Stock (1989)) e (ii) as distribuições das suas estatísticas de testes apresentam viés em pequenas amostras (Bekaert et al. (1997)). Além disso (iii) não permitem comparar os movimentos do diferencial observado com o diferencial implicado pela HE.

Para avaliar (iii) a capacidade da HE de explicar o formato observado da estrutura a termo, Campbell e Shiller (1991) propõe uma abordagem VAR, projetando  $S_t^{(n,m)*}$  num subconjunto de informação dos agentes. A abordagem VAR (i) contorna o problema de estimações com erros justapostos e, segundo Bekaert et al. (1997), (ii) minoram o problema de viés da distribuição. Resumidamente, assumindo o VAR(p) de componentes estacionários  $x_t \equiv \left[ \Delta R_t^{(m)}, S_t^{(n,m)} \right]$  como o subconjunto de informação dos agentes, ou mais compactamente o VAR(1):



$$z_t = Az_{t-1} + u_t,$$

onde  $z_t = [\Delta R_t^{(m)}, \Delta R_{t-1}^{(m)}, \dots, \Delta R_{t-p+1}^{(m)}, S_t^{(n,m)}, \dots, S_{t-p+1}^{(n,m)}]'$ , definem-se os vetores canônicos  $g$  e  $h$  tais que  $g'z_t = S_t^{(n,m)}$  e  $h'z_t = (R_{t+1}^{(m)} - R_t^{(m)})$ , e computa-se a previsão VAR,  $S_t^{(n,m)'}$ , do diferencial de antecipação perfeita,  $S_t^{(n,m)*}$ :

$$S_t^{(n,m)'} \equiv h'A \left[ I - (m/n)(I - A^n)(I - A^m)^{-1} \right] (I - A)^{-1} z_t, \quad (14)$$

que é o diferencial teórico sob expectativas racionais.

Se a HE sob expectativas racionais vale, a projeção de (11) num subconjunto de informação  $z_t$  implica:

$$S_t^{(n,m)} = g'z_t = S_t^{(n,m)'}, \quad (15)$$

qualquer que seja o conjunto de informação contendo  $S_t^{(n,m)}$ , significando que o diferencial observado coincide com a melhor previsão de crescimento de longo prazo das taxas de curto prazo. A intuição é que se o prêmio pelo prazo é constante, toda a informação relevante para o mercado está contida em  $S_t^{(n,m)}$ , o qual foi incluído no VAR.

Pode-se testar (15) de várias maneiras. Por exemplo sob HE, nem a correlação, nem a razão dos desvios-padrão entre  $S_t^{(n,m)'}$  e  $S_t^{(n,m)}$  deve ser diferente de 1. Outra implicação testável de (15) é que:

$$g'(I - A) \equiv h'A \left[ I - (m/n)(I - A^n)(I - A^m)^{-1} \right]. \quad (16)$$

Ainda sob HE, a regressão:

$$S_t^{(n,m)'} = \gamma + \beta S_t^{(n,m)} + \zeta \Omega_t + w_t, \quad (17)$$

onde  $\Omega_t$  é algum elemento do conjunto de informação do agente na data  $t$  e  $w_t$  é um termo de erro, deve resultar numa estimativa  $\zeta$  não-significativamente diferente de zero <sup>6</sup>.

Na seção seguinte, efetuamos os testes sugeridos de (9), (13) e (17) além de alguns outros motivados na mesma.

## 4 Testando a Hipótese das Expectativas

Na Tabela 3.a, as duas primeiras linhas de cada par  $(m, n)$  apresentam a estimativa de Mínimos Quadrados Ordinários (MQO) do coeficiente  $\beta$  da equação (9) e o erro-padrão assintótico, corrigido para heterocedasticidade e erros justapostos <sup>7</sup>. Para algumas combinações de  $m$  e  $n$  em que não estão disponíveis as taxas de prazo  $(n - m)$ , adotou-se a aproximação  $R_{t+m}^{n-m} = R_{t+m}^n$ .

A análise das estimativas de MQO evidencia que (i) o valor pontual do  $\beta_{mqo}$  é negativo para algumas combinações de prazos ( $n = 21, m = 1$ ) e ( $n = 42, m = 1$ ); e (ii)

<sup>6</sup>Complementarmente, se HE não vale, o diferencial observado  $S_t^{(n,m)}$  não resume adequadamente a informação contida em  $S_t^{(n,m)'}$ . Como consequência, devemos rejeitar as hipóteses sugeridas nos três testes acima.

<sup>7</sup>Os erros-padrão foram computados de acordo com Hansen e Hodrick (1980), considerando que o erro da equação (9) segue um processo  $MA(m - 1)$ .

vista a magnitude dos erros-padrão, não é possível rejeitar a hipótese de  $\beta = 1$ , nem a hipótese de  $\beta = 0$ . Nossos resultados são semelhantes aos de Campbell & Shiller(1991) e Hardouvelis(1994), entre outros, que caracterizaram as estimativas de  $\beta$ 's negativos como um enigma para a HE sob expectativas racionais (*puzzle*).

Tabela 3.a  
Coeficiente angular da regressão:  
 $R_{t+m}^{n-m} - R_t^n = \alpha + \beta(m/(n-m))(R_t^n - R_{t-1}^n) + u_{t+m}$   
Mínimos Quadrados Ordinários (MQO) e Variáveis Instrumentais (VI)

n		m				
		1	21	42	63	126
21	$\beta_{mqo}$	-0.02 (1.45)				
	$\beta_{vi}$	0.03 (0.39)				
42	$\beta_{mqo}$	-0.06 (1.60)	0.85 (0.50)			
	$\beta_{vi}$	-0.16 (0.63)	0.92 (0.31)			
63	$\beta_{mqo}$	0.22 (1.52)	0.58 (0.51)	0.71 (0.53)		
	$\beta_{vi}$	-0.12 (0.71)	0.60 (0.30)	0.71 (0.20)		
126	$\beta_{mqo}$	1.21 (1.51)	0.76 (0.74)	0.95 (0.68)	0.68 (1.06)	
	$\beta_{vi}$	0.28 (0.83)	0.70 (0.50)	0.94 (0.31)	0.68 (0.26)	
252	$\beta_{mqo}$	1.37 (2.15)	0.22 (1.23)	0.58 (0.99)	0.60 (1.92)	1.29 (1.22)
	$\beta_{vi}$	-0.28 (1.28)	0.12 (0.87)	0.55 (0.63)	0.41 (0.53)	1.35 (0.50)

Notas: (i) Erros-padrão, reportados entre parênteses, computados conforme Hansen-Hodrick (1980), supondo que os erros seguem um processo MA(m-1). (ii) A estimação por variáveis instrumentais utiliza vinte defasagens do spread ( $R_t^n - R_{t-1}^n$ ) e vinte defasagens da variação da taxa curta em m períodos ( $R_{t+m}^m - R_t^m$ ).

O baixo poder preditivo do diferencial,  $S_t^{(n,m)}$ , sobre a variação de curto prazo da taxa longa, levanta a dúvida se o primeiro tem alguma capacidade de explicar a variação futura da taxa de curto prazo. Indagação que pode ser repondida pela regressão auxiliar:

$$R_{t+m}^{(m)} - R_t^{(m)} = \chi + \delta \cdot S_t^{(n,m)} + \nu_{t+m}, \quad (18)$$

onde  $\nu_{t+m}$  é  $MA(m-1)$ , que projeta a variação de curto prazo da taxa curta,  $R_{t+m}^{(m)} - R_t^{(m)}$ , sobre o diferencial, como em Hardouvelis (1994). Observe, contudo, que o tamanho de  $\delta$  e sua distância da unidade não informam quão bem o comportamento das taxas curtas se adequa à HE, mas apenas se  $S_t^{(n,m)}$  tem algum poder preditivo sobre elas <sup>8</sup>.

Na Tabela 3.b, as duas primeiras linhas de cada par  $(m, n)$  apresentam as estimativas de MQO do coeficiente  $\delta$  da equação (18) e o erro-padrão assintótico, corrigido para heterocedasticidade e erros justapostos <sup>9</sup>. Ao contrário do  $\beta$  estimado na equação (9), o MQO produz estimativas de  $\delta$  positivas e significativas ao nível de 1% para 12 dos 15 vértices estudados, mostrando que, em geral, o diferencial tem poder preditivo sobre a variação de curto prazo da taxa curta.

<sup>8</sup>Para estudar esta questão é necessário examinar a evolução da taxa de curto prazo ao longo de vários períodos como em (13), o que será efetuado abaixo.

<sup>9</sup>Os erros-padrão foram computados de acordo com Hansen e Hodrick (1980), considerando que o erro da equação (9) segue um processo  $MA(m-1)$ .

Um explicação aventada para o desvio do  $\beta_{mqo}$  em relação a unidade é que as taxas são observadas com erro (Mankiw (1986)). Por exemplo, se a taxa longa é observada com um ruído branco aditivo:

$$R_t^{(n)} = R_t^{(n)*} + \varepsilon_t, \quad (19)$$

onde  $R_t^{(n)}$  é a taxa longa observada,  $R_t^{(n)*}$  é a verdadeira taxa longa implicada por HE e  $\varepsilon_t$  é o ruído, tem-se para a variação da taxa longa e para o diferencial, respectivamente:

$$R_{t+m}^{(n-m)} - R_t^{(n)} = R_{t+m}^{(n-m)*} - R_t^{(n)*} + (\varepsilon_{t+m} - \varepsilon_t),$$

e

$$R_t^{(n)} - R_t^{(m)} = R_t^{(n)*} - R_t^{(m)} + \varepsilon_t;$$

o que implica que o  $\beta_{mqo}$  possui viés negativo igual a (derivado no Apêndice C):

$$Viés(\beta_{mqo}) = -\frac{n}{m} \cdot \frac{\sigma^2(\varepsilon_t)}{\sigma^2(R_t^{(n)} - R_t^{(m)})}. \quad (20)$$

Por outro lado, se a taxa curta é observada com um ruído branco aditivo:

$$R_t^{(m)} = R_t^{(m)*} + \varepsilon_t, \quad (21)$$

onde  $R_t^{(m)*}$  é a verdadeira taxa curta e  $\varepsilon_t$  é o ruído, tem-se para a variação da taxa curta e para o diferencial, respectivamente:

$$R_{t+m}^{(m)} - R_t^{(m)} = R_{t+m}^{(m)*} - R_t^{(m)*} + (\varepsilon_{t+m} - \varepsilon_t),$$

e

$$R_t^{(n)} - R_t^{(m)} = R_t^{(n)*} - R_t^{(m)*} - \varepsilon_t,$$

o que implica que o  $\delta_{mqo}$  possui viés igual a (derivado no Apêndice C):

$$Viés(\delta_{mqo}) = (-\delta^* + 1) \cdot \frac{\sigma^2(\varepsilon_t)}{\sigma^2(R_t^{(n)} - R_t^{(m)})}. \quad (22)$$

Como as estimativas de MQO das equações (9) e (18) podem ter sido viesadas pela presença de ruído branco nas taxas observadas, reestimamos as duas equações por Variáveis Instrumentais (VI).

De (20), espera-se que o  $\beta$  estimado por VI cresça se a taxa longa é observada com ruído e convirja para a unidade. As duas últimas linhas de cada vértice  $(n, m)$  da Tabela 3.a apresentam os resultados da estimação por VI da equação (9). Usamos como instrumentos vinte defasagens do diferencial,  $R_t^{(n)} - R_t^{(m)}$ , e vinte defasagens da variação da taxa curta,  $R_t^{(m)} - R_{t-1}^{(m)}$ . As estimativas por VI não convergem para 1 como previsto pela hipótese do ruído. Contrário ao racionalizado pela hipótese do ruído branco, em cada linha referente a uma taxa longa, observam-se  $\beta'_{vi}$ s menores que  $\beta'_{mqo}$ s para todos os prazos, exceto  $n = 21$ , que pode estar sendo observada com ruído. Portanto, a explicação simples de ruído branco aditivo na taxa longa não parece suficiente para conciliar os baixos  $\beta'_{mqo}$ s com a HE.

Ainda insistindo na hipótese de ruído branco na taxa curta, de (22) espera-se que o  $\delta$  estimado por VI decresça se  $\delta^* < 1$  ou cresça se  $\delta^* > 1$ . As duas últimas linhas de cada

vértice  $(n, m)$  da Tabela 3.b apresentam os resultados da estimação por VI da equação (18). Usamos como instrumentos vinte defasagens do diferencial,  $R_t^{(n)} - R_t^{(m)}$ , e vinte defasagens da variação da taxa curta,  $R_t^{(m)} - R_{t-1}^{(m)}$ . Analisando a coluna referente à taxa de um dia ( $m = 1$ ), para a qual  $\delta$  parece pequeno, observam-se todos os  $\delta'_{vi}$ s menores ou iguais que  $\delta'_{mqo}$ s, evidência a favor da suspeita de taxa de um dia observada com ruído. Na coluna  $m = 21$ , para  $n = 42$  e  $63$ ,  $\delta'$ s aparentemente maiores que 1 e  $\delta'_{vi}$ s maiores que  $\delta'_{mqo}$ s também acordam com a suspeita de ruído na taxa de 21 dias. Esta suspeita sobrevive para  $m = 21$ , quando para  $\delta'$ s aparentemente menores que 1, os  $\delta'_{vi}$ s são menores que os  $\delta'_{mqo}$ s para  $n = 126$  e  $252$ . Como o exame das outras taxas curtas é desnecessário<sup>10</sup>, conclui-se das reestimações por VI que a hipótese de ruído branco só é razoável para as taxas curtas mais curtas ( $m = 1$  e  $21$ ), ao contrário do concluído por Hardouvelis (1994).

Os baixos  $\beta'_{mqo}$ s obtidos na Tabela 3.a são a regra na literatura de teste da HE, mas não parecem suficientes para rejeitá-la, dados os problemas de amostra finita da equação (9), já listados na última seção. Portanto, assim como na literatura, seguem-se outros testes em busca de evidências mais confiáveis.

Tabela 3.b  
Coeficiente angular da regressão:  
 $R_{t+m}^m - R_t^m = \chi + \delta(R_t^m - R_{t-1}^m) + v_{t+1}$   
Mínimos quadrados ordinários e variáveis instrumentais

n		m				
		1	21	42	63	126
21	$\delta_{mqo}$	0.13 (0.05)				
	$\delta_{vi}$	0.09 (0.01)				
42	$\delta_{mqo}$	0.10 (0.03)	1.85 (0.50)			
	$\delta_{vi}$	0.08 (0.01)	1.92 (0.31)			
63	$\delta_{mqo}$	0.08 (0.03)	1.11 (0.26)	2.57 (1.01)		
	$\delta_{vi}$	0.07 (0.01)	1.12 (0.15)	2.59 (0.40)		
126	$\delta_{mqo}$	0.06 (0.02)	0.67 (0.16)	1.13 (0.40)	1.68 (1.06)	
	$\delta_{vi}$	0.05 (0.00)	0.66 (0.08)	1.13 (0.17)	1.68 (0.26)	
252	$\delta_{mqo}$	0.04 (0.01)	0.45 (0.13)	0.72 (0.26)	1.38 (0.81)	2.29 (1.22)
	$\delta_{vi}$	0.04 (0.00)	0.45 (0.07)	0.71 (0.15)	0.89 (0.20)	2.35 (0.50)

Notas: (i) Erros-padrão, reportados entre parênteses, computados conforme Hansen-Hodrick (1980), supondo que os erros seguem um processo MA(m-1). (ii) A estimação por variáveis instrumentais utiliza vinte defasagens do spread  $((R_t^m - R_{t-1}^m))$  e vinte defasagens da variação da taxa curta em m períodos  $(R_{t+m}^m - R_t^m)$ .

Para examinar quão bem o comportamento das taxas curtas se adequa à HE é necessário examinar a evolução de longo prazo da taxa curta como em (13). A Tabela 4 exibe coeficientes angulares ( $\beta'$ s) positivos e significativos, indicando que o diferencial antecipa a variação de longo prazo da taxa curta em acordo com a HE. Confirmando a magnitude da reação prevista pela HE, nenhum coeficiente é significativamente diferente de 1<sup>11</sup>..Apesar

<sup>10</sup>Pois foram analisadas como taxas longas na Tabela 3.a, onde concluiu-se pela inexistência do ruído branco aditivo nas mesmas.

<sup>11</sup>14 dos 15 coeficientes não são significativamente diferentes de 1 a 5%, e nenhum deles o é a 1% de significância.

de não-significativos na maioria dos casos, os prêmios pelo prazo positivos ( $-\gamma$ , menos o intercepto) tem o sinal esperado de um mercado com agentes neutros ao risco.

Tabela 4  
Estimativas de MQO de:  
 $S_t^{(n,m)*} = \gamma + \beta S_t^{(n,m)} + v_{t+m}$

n	m				
	1	21	42	63	126
21	0.60 (0.19) -0.002 (0.001)				
42	0.74 (0.18) -0.003 (0.003)	0.92 (0.25) -0.001 (0.002)			
63	0.77 (0.18) -0.005 (0.005)	0.84 (0.27) -0.002 (0.004)			
126	0.80 (0.23) -0.008 (0.008)	0.98 (0.29) -0.006 (0.008)	0.94 (0.35) -0.005 (0.007)	0.84 (0.53) -0.004 (0.005)	
252	0.74 (0.19) -0.014 (0.012)	1.02 (0.31) -0.013 (0.012)	1.05 (0.38) -0.012 (0.011)	1.09 (0.46) -0.011 (0.010)	1.14 (0.61) -0.008 (0.006)

Notas: (i) De cima para baixo, coeficiente angular e erro-padrão, intercepto e erro-padrão. (ii) Erros-padrão, reportados entre parênteses, computados conforme Hansen-Hodrick (1980), supondo que os erros seguem um processo MA(n-m-1).

O melhor ajuste de (13) relativamente a (13) foi obtido anteriormente por Campbell e Shiller (1991), embora estes rejeitem a HE para maturidades inferiores a 4 anos.

Passando aos procedimentos VAR, da equação (15), a HE prevê uma correlação perfeita entre  $S_t^{(n,m)}$  e  $S_t^{(n,m)'}$ . A Tabela 5 abaixo apresenta as correlações estimadas e os seus respectivos erros-padrão computados pelo método delta <sup>12</sup>, para  $S_t^{(n,m)'}$  calculado por um VAR(20) de componentes  $x_t \equiv [\Delta R_t^{(m)}, S_t^{(n,m)}]$ . Todas as correlações são altas, embora significativamente diferentes da unidade, devido aos pequenos erros-padrão computados pelo método delta.

<sup>12</sup>Método delta: Considere o estimador  $\hat{\beta}$  que satisfaz:

$$\sqrt{T} (\hat{\beta} - \beta_0) \rightarrow^d N(0, V),$$

e seja  $f(\beta)$  uma sua transformação com primeira derivada contínua. Então:

$$\sqrt{T} (f(\hat{\beta}) - f(\beta_0)) \rightarrow^d N(0, \Psi),$$

onde:

$$\Psi = \frac{\partial f(\beta_0)}{\partial \beta} V \frac{\partial f(\beta_0)}{\partial \beta}.$$

Veja Cochrane (2001) para uma exposição detalhada.

Tabela 5:  
Correlação de  $S_t^{(n,m)}$  em  $S_t^{(n,m)}$

n	m				
	1	21	42	63	126
21	0.9442 0.0145				
42	0.9753 0.0054	0.9250 0.0294			
63	0.9824 0.0037	0.9575 0.0177	0.9233 0.0251		
126	0.9865 0.0028	0.9674 0.0115	0.9585 0.0130	0.9472 0.0165	
252	0.9858 0.0029	0.9710 0.0096	0.9654 0.0111	0.9537 0.0156	0.8933 0.0399

Nota: Correlação e erro padrão calculado pelo método delta.

Ainda da equação (15), a HE prevê que a razão dos desvios-padrão de  $S_t^{(n,m)}$  e  $S_t^{(n,m)}$  é igual a unidade. A Tabela 6 abaixo apresenta estas razões e os seus respectivos erros-padrão estimados pelo método delta.

Tabela 6:  
Razão  $\sigma(S^{(n,m)}_t) / \sigma(S^{(n,m)}_t)$

n	m				
	1	21	42	63	126
21	0.6322				
	0.0289				
42	0.7520	0.9928			
	0.0182	0.0187			
63	0.8275	0.9343	0.7323		
	0.0155	0.0221	0.0262		
126	0.9043	0.9296	0.8543	0.9540	
	0.0187	0.0276	0.0319	0.0422	
252	0.8874	0.8757	0.8298	0.8973	0.8126
	0.0200	0.0250	0.0244	0.0285	0.0288

Nota: Razão dos desvios-padrão e erro-padrão calculado pelo m. delta.

Embora positivos, todos os coeficientes são menores que a unidade, indicando que o diferencial observado é mais volátil que o diferencial teórico.

Como as evidências das Tabelas 4, 5 e 6 indicam que o coeficiente de reação da previsão das variações das taxas curtas ao diferencial observado é inferior a 1, somos levados a cogitar a hipótese que há uma reação exagerada (*overreaction*) do diferencial observado em relação a expectativa das variações das taxas curtas. A reação exagerada pode ser incorporada na forma:

$$(R_t^{(n)} - R_t^{(m)}) = (1 + d) \cdot E_t(R_t^{(n)*} - R_t^{(m)}), \quad (23)$$

onde  $R_t^{(n)} - R_t^{(m)}$  é o diferencial observado,  $R_t^{(n)*} - R_t^{(m)}$  é o diferencial teórico implicado pela HE e  $d$  representa o grau de reação exagerada.

Além de poder resultar num coeficiente angular negativo para a equação (9), a hipótese (23) implica um coeficiente de  $1/(1 + d)$  para a regressão (13), uma correlação unitária entre  $S_t^{(n,m)}$  e  $S_t^{(n,m)*} = E_t S_t^{(n,m)*}$  e uma razão de desvios padrão:

$$\frac{\sigma(E_t R_t^{(n)*} - R_t^{(m)})}{\sigma(R_t^{(n)} - R_t^{(m)})} = \frac{1}{1 + d}. \quad (24)$$

De (23), (6) e (9), deriva-se ainda a seguinte relação entre o grau de reação exagerada  $d$  e os estimadores de MQO  $\beta_{mqo}$  e  $\delta_{mqo}$  (derivada no Apêndice B):

$$d = \left[ \frac{1 - \beta_{mqo}}{\delta_{mqo}} \right] \cdot \frac{m}{(n - m)}. \quad (25)$$

Na Tabela 7 apresentamos os valores de  $d$  calculados pela expressão (25) e os respectivos valores da razão de desvios-padrão (24) implicados. Comparando a segunda linha de vértice da Tabela 7 com a primeira linha de cada vértice da Tabela 4, observa-se que não são significativamente diferentes e que, portanto, a hipótese de reação exagerada pode fazer sentido.

Tabela 7  
Grau de reação exagerada ( $d$ ) implicado por  $\beta_{ols}$  e  $\delta_{ols}$ .

n		m				
		1	21	42	63	126
21	d	0.41				
	$\sigma(R^n, R^m)/\sigma(R^n, R^m)$	0.71				
42	d	0.27	0.08			
	$\sigma(R^n, R^m)/\sigma(R^n, R^m)$	0.79	0.92			
63	d	0.15	0.19	0.23		
	$\sigma(R^n, R^m)/\sigma(R^n, R^m)$	0.87	0.84	0.81		
126	d	-0.03	0.07	0.02	0.19	
	$\sigma(R^n, R^m)/\sigma(R^n, R^m)$	1.03	0.93	0.98	0.84	
252	d	-0.03	0.16	0.12	0.10	-0.13
	$\sigma(R^n, R^m)/\sigma(R^n, R^m)$	1.03	0.86	0.89	0.91	1.14

$d = [(1 - \beta_{ols})\delta_{ols}] / [(n/m) - 1]$  é o grau de *overreaction* necessário para explicar a diferença entre o parâmetro  $\beta$  estimado e a unidade.  $\sigma(R^n, R^m)/\sigma(R^n, R^m) = 1/(1+d)$  é a razão entre o desvio padrão teórico e realizado implicado por  $d$ .

Os conjuntos de razões apresentados nas Tabelas 7 e Tabela 6 são qualitativamente parecidos, embora significativamente diferentes. A divergência entre os valores das Tabelas 7 e Tabela 6 não é gritante uma vez que os erros-padrão estimados pelo método delta são pequenos.

Em suma, a explicação de reação exagerada parece mais razoável.

## 5 Conclusão

Este artigo constrói a estrutura a termo diária das taxas de juros brasileira para o período compreendido entre Julho de 1996 a Dezembro de 2002 utilizando a metodologia sugerida na Nota Técnica a Circular  $n^o$  2972 do Banco Central do Brasil. Estudamos a HE sob expectativas racionais segundo as abordagens de Campbell e Shiller (1991) e Hardouvelis (1994). (i) Os coeficientes estimados do diferencial na equação de mudança de curto prazo da taxa de longo prazo são bastante imprecisos e não permitem conclusão. (ii) A estimação da equação de mudança de curto prazo da taxa de curto prazo, por MQO e por VI indica que um ruído branco aditivo só é plausível na taxa observada de curto prazo. (iii) Os coeficientes de MQO do diferencial na equação de mudança de longo prazo da taxa de curto prazo são mais precisamente estimados e não são significativamente diferentes da unidade. (iv) A previsão de expectativas racionais da mudança de longo prazo da taxa de curto prazo é altamente correlacionada com o diferencial, e (iv) o diferencial observado é mais volátil que o diferencial teórico.



Em suma, os resultados dos testes rejeitam a hipótese particular coeficiente unitário, porém reconhecem o poder preditivo do diferencial. Preliminarmente, a hipótese de reação exagerada parece ser uma explicação razoável para as divergências em relação a HE.

A validade da HE para o Brasil continua uma questão em aberto, bem como o é para o resto do mundo.

## References

- [1] **Banco Central do Brasil, 2000**, Nota Técnica sobre a Circular *n*<sup>o</sup> 2.972, de 23 de março de 2000.
- [2] **Bekaert, G., Hodrick, R.J., Marshall, D.A., 1997**, “On biases in test of the expectations hypothesis of the term structure of interest rates.” *Journal of Financial Economics*, 44, 309-348.
- [3] **Campbell, J.Y. e Shiller, R.J., 1991**, “Yield Spreads and Interest Rate Movements: A Bird’s Eye View.” *Review of Economics Studies*, 58, 495-514.
- [4] **Campbell, J.Y. e Shiller, R.J., 1987**, “Cointegration and Tests of Present Value Models.” *Journal of Political Economy*, 95, 5, 1062-1088.
- [5] **Cochrane, J., 2000**, *Asset Pricing*, Princeton University Press.
- [6] **Engle, R.F., Lilien, D.M. e Robins, R.P., 1987**, “Estimating Time Varying Risk Premia in the Term Structure: the ARCH-M Model.” *Econometrica*, 55/2: 391-407.
- [7] **Fama, E.F., 1984**, “The Information in the Term Structure.” *Journal of Financial Economics*, 13, 509-528.
- [8] **Fama, E.F. e Bliss, R.R., 1987**, “The Information in Long-Maturity Forward Rates.” *The American Economic Review*, Vol. 77, No. 4, 680-692.
- [9] **Froot, K., 1989**, “New hope for the expectations hypothesis of the term structure.” *The Journal of Finance*, Vol. XLIV, No. 2, 283-305.
- [10] **Hansen, B.E., 1992**, “Heteroskedastic Cointegration.” *Journal of Econometrics*, 54, 139-158.
- [11] **Hansen, L.P. e Hodrick, R.J., 1980**, “Forward Rates as Optimal Predictors of Future Spot Rates.” *Journal of Political Economy*, 88, 829-853.
- [12] **Hardouvelis, G.A., 1994**, “The term structure spread and futures changes in long and short rates in the G7 countries.” *Journal of Monetary Economics*, 33, 255-283.
- [13] **Johansen, S., 1991**, “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models.” *Econometrica*, Vol. LIX, No. 6, 1551-1580.
- [14] **Kwiatkowski, D., Phillips, P. C., Schmidt, P. e Shin, Y. 1992**, “Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root.” *Journal of Econometrics*, ??, 159-178.

- [15] **Lima, A.M. e Issler, J. V., 2002**, “A Hipótese das Expectativas na Estrutura a Termo de Juros no Brasil: uma Aplicação de Modelos de Valor Presente.” Anais II Encontro Brasileiro de Finanças.
- [16] **Longstaff, F.A., 2000**, “The term structure of very short-term rates: New evidence for the expectations hypothesis.” *Journal of Financial Economics*, 58, 397-415.
- [17] **Mankiw, N.G. e Summers, L.H., 1984**, “Do Long-Term Interest Rates Overreact to Short-Term Interest Rates.” *Brooking Papers on Economic Activity*, 1:1984, 223-242.
- [18] **Mankiw, N.G., 1986**, “The Term Structure of Interest Rates Revisited.” *Brooking Papers on Economic Activity*, 1:1986, 61-96.
- [19] **Modigliani F. e Sutch, R., 1969** “The Term Structure of Interest Rates: a re-examination of evidence.” *Journal of Money Credit and Banking*, , 112-20.
- [20] **Richardson, M. e Stock, J., 1989** “Drawing Inferences from Statistics Based on Multi-Year Asset Returns.” *Journal of Financial Economics*, 25, 323-348.
- [21] **Shiller, R.J., Campbell, J.Y. e Schoenholtz, K.L., 1983**, “Foward Rates and Future Policy: Interpreting the Term Structure of Interest Rates.” *Brookings Papers on Economic Activity*, 1:1983, 173-217.
- [22] **Tabak, B.M., and Andrade, S.C., 2001**, “Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates.” *Working Paper Series*, n. 30, Banco Central do Brasil.
- [23] **White, H., 1980**, “A Heteroskedasticity-Consistent Covariance Matrix and a Direct Test for Heteroskedasticity.” *Econometrica*, Vol. XLVIII, No. ??, 817-838.

## A Construção da Estrutura a Termo

A construção da estrutura a termo seguiu a metodologia sugerida na Nota Técnica a Circular n<sup>o</sup> 2972 do Banco Central do Brasil, na qual são adotados os prazos (eixo temporal da estrutura a termo) de 21, 42, 63, 126 e 252 dias úteis. Os dados brutos são compostos por dados diários, a partir de 04/07/1994, da taxa DI (*over*), das cotações de ajuste dos PUs dos primeiro, segundo e terceiro vencimentos de contratos futuros de DI de um dia da BM&F, e das taxas médias de swap pré x DI de 6, 12, 24 e 36 meses apuradas pela BM&F.

Após a correção e atualização da base de dados das cotações dos diferentes instrumentos financeiros mencionados na seção 2, calcula-se a taxa a termo implícita anual entre diferentes instrumentos financeiros. Esse procedimento tem o objetivo de uniformizar prazos e cotações dos diferentes instrumentos. Inicialmente calcula-se a taxa a termo implícita entre o CDI e o primeiro vencimento do futuro de DI de um dia mediante a fórmula,

$$Termo_{CDI,DI_1} = \left[ \left( \frac{100.000}{PU_{DI_1} \cdot \left\{ 1 + \frac{CDI}{100} \right\}^{\frac{1}{252}}} \right)^{\frac{252}{TDU_{DI_1} - 1}} - 1 \right] \cdot 100 \quad (26)$$

onde  $PU_{DI_1}$  é a cotação de ajuste do primeiro vencimento de futuros de DI e  $TDU_{DI_1}$  é o total de dias úteis até o vencimento do contrato futuro. O segundo passo é calcular a taxa a termo implícita entre dois vencimentos de futuros de DI de um dia, este cálculo é efetuado mediante a fórmula,

$$Termo_{DI_i,DI_j} = \left[ \left( \frac{PU_{DI_i}}{PU_{DI_j}} \right)^{\frac{252}{TDU_{DI_j} - TDU_{DI_i}}} - 1 \right] \cdot 100 \quad (27)$$

onde  $i$  é menor que  $j$ ,  $PU_{DI_k}$  ( $k = i, j$ ) são as cotações de ajuste dos vencimentos de futuros de DI  $i$  e  $j$  respectivamente, e  $TDU_{DI_k}$  ( $k = i, j$ ) são os totais de dias úteis até o vencimento de cada contrato. Em seguida calculamos a taxa a termo implícita entre um contrato futuro de DI de um dia e um swap de  $j$  dias utilizando a fórmula,

$$Termo_{DI_i,Swap_j} = \left[ \left( \frac{\left\{ 1 + \frac{TxSwap_j}{100} \right\} \frac{TDC_{swap_j}}{360}}{\frac{100.000}{PU_{DI_i}}} \right)^{\frac{252}{TDC_{swap_j} - TDU_{DI_i}}} - 1 \right] \cdot 100 \quad (28)$$

onde,  $TxSwap_j$  é a cotação média do swap de  $j$  dias,  $TDC_{swap_j}$  é o total de dias corridos deste swap,  $PU_{DI_i}$  é o preço de ajuste do  $i$ -ésimo vencimento de futuro de DI de um dia e  $TDU_{DI_i}$  é o total de dias úteis deste futuro. Por último, para finalizar o cálculo das taxas a termo implícitas entre instrumentos financeiros, devemos calcular a taxa implícita entre dois swap,

$$Termo_{Swap_i,Swap_j} = \left[ \left( \frac{\left\{ 1 + \frac{TxSwap_j}{100} \right\} \frac{TDC_{swap_j}}{360}}{\left\{ 1 + \frac{TxSwap_i}{100} \right\} \frac{TDC_{swap_i}}{360}} \right)^{\frac{252}{TDC_{swap_j} - TDC_{swap_i}}} - 1 \right] \cdot 100 \quad (29)$$

onde onde,  $TxSwap_k$  ( $k = i, j$ ) é a cotação média do swap de  $i$  e  $j$  dias,  $TDC_{swap_k}$  ( $k = i, j$ ) é o total de dias corridos destes swaps.

De posse das taxas a termo implícitas entre os vários instrumentos utilizados, construímos as taxas spot mediante a composição das taxas a termo para cada prazo escolhido. A taxa spot de  $T$  dias úteis será construída pela fórmula:

$$S_T = \left\{ \left[ VCDI \cdot \prod_{j=1}^{T-1} \left( 1 + \frac{T_{j,j+1}}{100} \right)^{g(i,j)} \right] \cdot \left( 1 + \frac{T_{N-1,N}}{100} \right)^{g(T_N)} \right]^{\frac{1}{T}} - 1 \right\} \cdot 100 \quad (30)$$

onde:  $VCDI = (1 + CDI/100)$ ,  $g(i, j) = \max[0, \min[T - T_j, T_{j+1} - T_j]]$ ,  $g(T_N) = \max[0, T - T_N]$  e  $T_{i,j}$  é a taxa a termo entre os instrumentos financeiros  $i$  e  $j$ .

## B Derivação da Equação (25)

Adiantando a equação (23) em  $m$  períodos e subtraindo dela mesmo, obtemos:

$$(1 + d) \cdot (R_{t+m}^{(n-m)*} - R_t^{(n)*}) = (R_{t+m}^{(n-m)} - R_t^{(n)}) + d \cdot (R_{t+m}^{(m)} - R_t^{(m)}), \quad (31)$$

aplicando a equação (6) para as taxas teóricas e fazendo  $E_t \varphi_{k,t} = 0$ , obtemos:

$$R_{t+m}^{(n-m)*} - R_t^{(n)*} = \frac{m}{n-m} \left( R_t^{(n)*} - R_t^{(m)} + \xi_{t+m} \right), \quad (32)$$

substituindo a equação (32) na equação (31), temos:

$$\frac{(1+d) \cdot m}{n-m} \left( R_t^{(n)*} - R_t^{(m)} + \xi_{t+m} \right) = (R_{t+m}^{(n-m)} - R_t^{(n)}) + d \cdot (R_{t+m}^{(m)} - R_t^{(m)}), \quad (33)$$

substituindo a equação (23) no lado esquerdo da equação (33), e reescrevendo-a, chegamos a:

$$(R_{t+m}^{(n-m)} - R_t^{(n)}) = \frac{m}{n-m} \left( R_t^{(n)} - R_t^{(m)} + (1+d) \cdot \xi_{t+m} \right) - d \cdot (R_{t+m}^{(m)} - R_t^{(m)}). \quad (34)$$

Da equação (9) temos:

$$\beta_{mqo} = \frac{n-m}{m} \cdot \frac{\text{cov} \left( R_{t+m}^{(n-m)} - R_t^{(n)}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}, \quad (35)$$

substituindo (34) em (35),

$$\beta_{mqo} = 1 + \frac{(1+d) \cdot \text{cov} \left( \xi_{t+m}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} - \frac{n-m}{m} \cdot \frac{d \cdot \text{cov} \left( R_{t+m}^{(m)} - R_t^{(m)}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}, \quad (36)$$

supondo  $\text{cov} \left( \xi_{t+m}, R_t^{(n)} - R_t^{(m)} \right) = 0$  e utilizando o estimador de MQO  $\delta$  da equação (18),

$$\beta_{mqo} = 1 - \frac{n-m}{m} \cdot d \cdot \delta_{mqo} \text{ ou } d = \left[ \frac{1 - \beta_{mqo}}{\delta_{mqo}} \right] \cdot \frac{m}{(n-m)} \quad (37)$$

## C Viés dos estimadores de MQO devido ao ruído aditivo ( $\varepsilon_t$ )

O  $\beta$  estimado por MQO da equação (9) tem a forma:

$$\beta_{mqo} = \frac{n-m}{m} \cdot \frac{\text{cov} \left( R_{t+m}^{(n-m)} - R_t^{(n)}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}, \quad (38)$$

que, dado o ruído branco na taxa longa (equação (19)):

$$R_{t+m}^{(n-m)} - R_t^{(n)} = R_{t+m}^{(n-m)*} - R_t^{(n)*} + (\varepsilon_{t+m} - \varepsilon_t), \quad (39)$$

e

$$R_t^{(n)} - R_t^{(m)} = R_t^{(n)*} - R_t^{(m)} + \varepsilon_t, \quad (40)$$

implica:

$$\begin{aligned}\beta_{mqo} &= \frac{n-m}{m} \cdot \frac{\text{cov}\left(R_{t+m}^{(n-m)*} - R_t^{(n)*} + (\varepsilon_{t+m} - \varepsilon_t), R_t^{(n)*} - R_t^{(m)} + \varepsilon_t\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)} \\ &= \frac{n-m}{m} \cdot \left[ \frac{\text{cov}\left(R_{t+m}^{(n-m)*} - R_t^{(n)*}, R_t^{(n)*} - R_t^{(m)}\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)} + \frac{\text{cov}\left((\varepsilon_{t+m} - \varepsilon_t), \varepsilon_t\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)} \right],\end{aligned}\quad (41)$$

ou,

$$\beta_{mqo} = \frac{n-m}{m} \cdot \left\{ \frac{\frac{\text{cov}\left(R_{t+m}^{(n-m)*} - R_t^{(n)*}, R_t^{(n)*} - R_t^{(m)}\right)}{\text{var}\left(R_t^{(n)*} - R_t^{(m)}\right)}}{\frac{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)}{\text{var}\left(R_t^{(n)*} - R_t^{(m)}\right)}} - \frac{\text{var}\left(\varepsilon_t\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)} \right\}.$$

Como o coeficiente angular da regressão (9) para taxas sem ruído é dada por:

$$\beta^* = \frac{n-m}{m} \cdot \frac{\text{cov}\left(R_{t+m}^{(n-m)*} - R_t^{(n)*}, R_t^{(n)*} - R_t^{(m)}\right)}{\text{var}\left(R_t^{(n)*} - R_t^{(m)}\right)},\quad (43)$$

tem-se que:

$$\beta_{mqo} = \beta^* \frac{\text{var}\left(R_t^{(n)*} - R_t^{(m)}\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)} - \frac{n-m}{m} \cdot \frac{\text{var}\left(\varepsilon_t\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)}.$$

Lembrando que:

$$\text{Var}\left(R_t^{(n)} - R_t^{(m)}\right) = \text{Var}\left(R_t^{(n)*} - R_t^{(m)}\right) + \text{Var}\left(\varepsilon_t\right),\quad (45)$$

então:

$$\begin{aligned}\beta_{mqo} &= \beta^* \left[ \frac{\text{Var}\left(R_t^{(n)} - R_t^{(m)}\right) - \text{Var}\left(\varepsilon_t\right)}{\text{Var}\left(R_t^{(n)} - R_t^{(m)}\right)} \right] - \frac{n-m}{m} \cdot \frac{\text{var}\left(\varepsilon_t\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)} \\ &= \beta^* - \left( \beta^* + \frac{n-m}{m} \right) \cdot \frac{\text{var}\left(\varepsilon_t\right)}{\text{var}\left(R_t^{(n)} - R_t^{(m)}\right)}.\end{aligned}\quad (46)$$

Se de HE espera-se  $\beta^* = 1$ , resulta que:

$$\beta_{mqo} = 1 - \frac{n}{m} \cdot \frac{\sigma^2\left(\varepsilon_t\right)}{\sigma^2\left(R_t^{(n)} - R_t^{(m)}\right)}\quad (47)$$

e:

$$\sigma\left(\varepsilon_t\right) = \left[ \frac{(1 - \beta_{mqo})}{(n/m)} \right]^{0.5} \cdot \sigma\left(R_t^{(n)} - R_t^{(m)}\right),\quad (48)$$

diferente do apresentado em Hardouvelis (1994).

O  $\delta$  estimado por MQO da equação (18) tem a forma:

$$\delta_{mqo} = \frac{\text{cov} \left( R_{t+m}^{(m)} - R_t^{(m)}, R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}, \quad (49)$$

que, dado o ruído branco na taxa curta (equação (21)):

$$R_{t+m}^{(m)} - R_t^{(m)} = R_{t+m}^{(m)*} - R_t^{(m)*} + (\varepsilon_{t+m} - \varepsilon_t), \quad (50)$$

e

$$R_t^{(n)} - R_t^{(m)} = R_t^{(n)*} - R_t^{(m)} - \varepsilon_t, \quad (51)$$

implica:

$$\begin{aligned} \delta_{mqo} &= \frac{\text{cov} \left( R_{t+m}^{(m)*} - R_t^{(m)*} + (\varepsilon_{t+m} - \varepsilon_t), R_t^{(n)*} - R_t^{(m)} - \varepsilon_t \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} \\ &= \left[ \frac{\text{cov} \left( R_{t+m}^{(m)*} - R_t^{(m)*}, R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} + \frac{\text{cov} \left( (\varepsilon_{t+m} - \varepsilon_t), -\varepsilon_t \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} \right], \end{aligned} \quad (52)$$

ou,

$$\delta_{mqo} = \left\{ \frac{\frac{\text{cov} \left( R_{t+m}^{(m)*} - R_t^{(m)*}, R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)*} - R_t^{(m)} \right)}}{\frac{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)*} - R_t^{(m)} \right)}} + \frac{\text{var} \left( \varepsilon_t \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} \right\}.$$

Como o coeficiente angular da regressão (9) para taxas sem ruído é dada por:

$$\delta^* = \frac{\text{cov} \left( R_{t+m}^{(m)*} - R_t^{(m)*}, R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)*} - R_t^{(m)} \right)}, \quad (54)$$

tem-se que:

$$\delta_{mqo} = \delta^* \frac{\text{var} \left( R_t^{(n)*} - R_t^{(m)} \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} + \frac{\text{var} \left( \varepsilon_t \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}.$$

Lembrando que:

$$\text{Var} \left( R_t^{(n)} - R_t^{(m)} \right) = \text{Var} \left( R_t^{(n)*} - R_t^{(m)} \right) + \text{Var} \left( \varepsilon_t \right), \quad (56)$$

então:

$$\begin{aligned} \delta_{mqo} &= \delta^* \left[ \frac{\text{Var} \left( R_t^{(n)} - R_t^{(m)} \right) - \text{Var} \left( \varepsilon_t \right)}{\text{Var} \left( R_t^{(n)} - R_t^{(m)} \right)} \right] + \frac{\text{var} \left( \varepsilon_t \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)} \\ &= \delta^* + (-\delta^* + 1) \cdot \frac{\text{var} \left( \varepsilon_t \right)}{\text{var} \left( R_t^{(n)} - R_t^{(m)} \right)}. \end{aligned} \quad (57)$$