

A Hipótese das Expectativas na Estrutura a Termo de Juros no Brasil: Uma Aplicação de Modelos de Valor Presente*

Alexandre Maia Correia Lima**
João Victor Issler**

Sumário: 1. Introdução; 2. Hipótese das expectativas e modelos de valor presente; 3. Teoria e proposições testáveis; 4. Base de dados; 5. Resultados empíricos; 6. Conclusão.

Palavras-chave: hipótese das expectativas; modelos de valor presente; VAR; cointegração.

Códigos JEL: E32; C32; C53.

Utilizando dados financeiros brasileiros da BM&F, testa-se a validade do modelo de valor presente na estrutura a termo de juros, também conhecido na literatura como Hipótese das Expectativas. Estes modelos relacionam a taxa de juros de longo prazo à uma média das taxas de juros de curto-prazo mais um prêmio de risco, invariante no tempo. Associada a estes modelos está a questão da previsibilidade dos retornos de ativos financeiros ou, mais especificamente, à previsibilidade na evolução das taxas de juros. Neste artigo é realizada uma análise multivariada num arcabouço de séries temporais utilizando a técnica de Auto-Regressão Vetorial. Os resultados empíricos aceitam apenas parcialmente a Hipótese das Expectativas para a estrutura a termo de juros brasileira.

Using Brazilian financial data provided by BMF we test the expectations hypothesis of interest rates present value models (Campbell and Shiller, 1987). The empirical results are ambiguous: we find a common factor for “long” and short interest rates (cointegration), but present-value restrictions are rejected at 10% significance but not at 5%.

1. Introdução

A relação entre as taxas de juros de curto prazo e de longo prazo tem sido um dos temas mais investigados na literatura empírica de macroeconomia e finanças,

* Artigo recebido em jul. 2002 e aprovado em mai. 2003. Esse artigo é uma versão modificada da tese de mestrado de Alexandre Maia Correia Lima, supervisionada por João Victor Issler.

**EPGE/FGV.

tanto pela sua relevância em exercícios de modelagem que envolvam expectativas racionais e mercados eficientes, como pela sua importância para a condução da política monetária (Cox et alii, 1985) para aspectos teóricos da estrutura a termo da taxa de juros.

A explicação mais antiga e amplamente testada sobre as flutuações nas taxas de juros da economia é a Hipótese das Expectativas (HE), que estabelece, em linhas gerais, que a taxa de juros de longo prazo é formada como uma média das taxas de juros de curto prazo esperadas para o futuro mais um prêmio de risco invariante no tempo. Assim, a inclinação da curva de juros refletiria as expectativas do mercado sobre mudanças futuras nas taxas de juros, ou seja, um aumento do *spread* entre a taxa longa e a taxa curta refletiria a expectativa de futuros aumentos nas taxas de curto prazo, provocando então uma correlação positiva entre as variações nas taxas de juros curtas e o referido *spread*. Da mesma forma, um aumento da taxa longa em relação à taxa curta, no período corrente, deveria levar a um aumento da taxa longa no período subsequente.

Numerosos estudos sobre a HE, concentrados na segunda metade da década de 80 e na década de 90, encontram resultados bastante divergentes, a depender do país, do segmento da curva de juros examinado e da técnica econométrica utilizada. Mankiw e Miron (1986), por exemplo, estudam o comportamento das taxas de juros de 3 e 6 meses nos EUA entre 1890-1979 e mostraram que a HE encontra forte suporte empírico apenas até 1915, antes da fundação do Federal Reserve. Os autores argumentam que a partir daí o *spread* entre as taxas de juros passou a ter um poder de previsão bem menos significativo, dado que, com a consolidação do *Federal Reserve*, a taxa curta seria melhor modelada como um passeio aleatório. Hardouvelis (1994), por sua vez, utilizando dados do pós-guerra até 1992 para as taxas de juros de 3 meses e 10 anos dos países do G7, encontra fortes evidências da HE para todos os países, à exceção dos EUA. Froot (1989) também encontra evidências que refutam parcialmente a HE para os EUA, atribuindo o fraco poder de previsão do *spread* sobre as mudanças nas taxas longas a erros expectacionais sistemáticos.

Mais recentemente, Gerlach e Smets (1997) e Longstaff (2000) trabalham com o segmento curto da curva de juros (prazo abaixo de 1 ano) e encontraram evidências adicionais de validade da EH para vários dos países pesquisados. Jondeau e Ricart (1999) chegam a resultados mistos, trabalhando com taxas de juros em diferentes prazos (1, 3, 6 e 12 meses) e cobrindo o período que vai de 1975 a 1997 para quatro países. Em dois deles, França e Inglaterra, a HE é validada pelos resultados empíricos, ao contrário das conclusões obtidas com os dados relativos aos EUA e a Alemanha.

Um dos arcabouços mais utilizados na literatura econométrica para testar a HE é o desenvolvido por Campbell e Shiller (1987, 1991) para investigar relações de valor presente. Este consiste na estimação de um VAR (Auto-Regressão Vetorial) bi-variado para a variação em primeiras diferenças nas taxas de juros de curto prazo e o *spread* entre a taxa longa e a taxa curta. Autores como Taylor (1992) e Cuthbertson (1996) usam essa metodologia para testar a HE em seus respectivos países. Aqui, a metodologia de Campbell e Shiller será utilizada para investigar a HE.

Apesar de existir uma vasta literatura internacional sobre a Hipótese das Expectativas e sua aderência empírica quando aplicada às séries históricas de taxas de juros nos diversos países, a literatura brasileira sobre o assunto ainda é incipiente. Do nosso conhecimento, Tabak e Andrade (2001) foram os primeiros a trabalhar este tema empiricamente no Brasil e refutam a validade da HE para a estrutura a termo brasileira na maioria dos casos por eles estudados; ver também o *survey* de Rossi (1996) e os argumentos que justificavam, naquele momento, a escassa literatura no Brasil sobre o assunto.

Tabak e Andrade utilizaram duas diferentes metodologias baseadas em regressões envolvendo uma equação única (single-equation regression), sendo a primeira delas uma regressão de uma média ponderada das mudanças nas taxas de curto prazo contra o *spread* da taxa longa em relação à taxa curta e uma constante. Em seguida, incluem nessa mesma regressão uma variável *proxy* para o prêmio de risco da economia brasileira, testando se o seu coeficiente é estatisticamente significativo. Além de Tabak e Andrade, e posterior a nosso estudo, há também o estudo recente de Brito et alii (2003), que chega a resultados ambíguos quanto a validade da Hipótese das Expectativas usando uma gama de técnicas econométricas.

Tabak e Andrade não usam métodos multivariados de análise de regressão para implementar seus testes, o que é uma limitação para o conjunto de informação usado. Em modelos de expectativas racionais, estas são usualmente computadas condicionando-se num grande número de variáveis, o que em geral requer modelar simultaneamente mais de uma variável. O presente artigo é uma contribuição à literatura brasileira sobre a HE e sobre a importância da estrutura a termo de juros exatamente porque usa técnicas econométricas de análise multivariada, aplicando o arcabouço de modelos de Auto-Regressões Vetoriais (VAR) para testar a validade da HE para o Brasil em diferentes horizontes – quando se considera o *spread* de 30 e 180 dias e o de 30 e 360 dias.

Na seção 2 apresenta-se a relação entre a Hipótese das Expectativas e modelos de valor presente. A seção 3 discute formalmente a testes para a Hipótese das Expectativas usando modelos de valor presente, apresentando a versão desenvolvida

por Campbell e Shiller (1987, 1991). A seguir, na seção 4 descreve-se os dados, e na seção 5 apresenta-se os resultados empíricos encontrados. A seção 6 apresenta as conclusões.

2. Hipótese das Expectativas e Modelos de Valor Presente

Seguindo a definição de Longstaff (2000), a Hipótese das Expectativas (HE) da estrutura a termo de juros estabelece que:

$$E[r_{t+n} | H_t] = R_{t(n)} + a_n \quad (1)$$

onde r_{t+n} é a taxa de juros de curto prazo média do período t até $t + n$, H_t é o conjunto de informação do agente econômico usado no período t , $R_{t(n)}$ é a taxa esperada em t para o período n e a_n é um prêmio de risco invariante no tempo, ainda que possa diferir em distintos horizontes n .

De forma geral, quando a_n é zero, esta representação da HE é geralmente chamada de *Hipótese Pura das Expectativas* e está associada ao trabalho seminal de Fisher (1896). Nas palavras deste, “um investidor que carrega um título por um longo tempo apropria-se de um rendimento que é uma média das taxas oscilantes daqueles que especularam nesse período”.

A intuição por trás de (1) é extremamente simples e se baseia na existência de arbitragem. Suponha que houvesse uma expectativa de que as taxas de curto prazo fossem, em média, exageradamente pequenas (altas) em relação à taxa longa, para um determinado horizonte. Os “tomadores” (emprestadores) teriam todo o incentivo a tomar (emprestar) dinheiro usando títulos de curto prazo, emprestando (tomando emprestado) esses recursos usando títulos de longo prazo, movimento esse que faria a taxa longa cair (subir) e a curta subir (cair) de forma a que o equilíbrio fosse restabelecido por esse mecanismo de arbitragem. Assim, por trás da Hipótese das Expectativas está a teoria dos mercados eficientes, o que impõe que quaisquer movimentos nas taxas de juros mais longas deveriam estar associados a novas informações sobre a média das taxas de juros no futuro.

A Hipótese das Expectativas está intimamente relacionada a modelos de valor presente, o que ficou claro a partir do estudo pioneiro de Shiller (1979). Essa conexão se obtém quando se log-lineariza a equivalência entre o valor-presente esperado do fluxo de caixa de um título de curto prazo, com retorno r_t , replicado por n períodos, com um de longo prazo, com retorno $R_t^{(n)}$, num ambiente de expectativas racionais:¹

¹Numa forma ainda mais simples e intuitiva, considere a equivalência entre os retornos de

$$R_t^{(n)} = \frac{1 - \delta}{1 - \delta^n} \sum_{i=0}^{n-1} \delta^i E_t r_{t+i} + c_n \quad (2)$$

onde (2) foi construída para que a soma dos pesos seja igual a um, e as taxas de juros de curto prazo esperadas para um futuro próximo têm uma ponderação maior na determinação da taxa longa do que as taxas de curto prazo esperadas mais adiante. Estabelecendo que $\delta = 1/(1+r)$, onde r é taxa de desconto, temos que $R_t^{(n)}$, a taxa longa, é descrita na equação (2) como o valor presente das futuras taxas de juros de curto prazo descontadas por r .

A conexão entre (1) e (2) se dá porque podemos escrever a esperança taxa de juros de curto prazo do período t até $t+n$ como a esperança de um valor presente $E_t \left(\frac{1-\delta}{1-\delta^n} \sum_{i=0}^{n-1} \delta^i r_{t+i} \right)$, onde fica óbvio que r_{t+n} será uma média ponderada, $r_{t+n} = \frac{1-\delta}{1-\delta^n} \sum_{i=0}^{n-1} \delta^i r_{t+i}$, para que se obtenha a igualdade entre (1) e (2).

Quando temos n grande o suficiente na equação (2), esta pode ser aproximada tomando-se seu limite quando $n \rightarrow \infty$:

$$R_t = (1 - \delta) \sum_{i=0}^{\infty} \delta^i E_t r_{t+i} + c \quad (3)$$

3. Teoria e Proposições Testáveis

3.1 Testes da hipótese das expectativas (HE) baseados em modelos de valor presente (MVP)

A equação (3) se enquadra perfeitamente ao arcabouço geral de Campbell e Shiller (1987) de valor presente para duas variáveis y_t e Y_t , onde Y_t é uma função linear do valor presente descontado dos valores futuros esperados de y_t :

$$Y_t = \theta(1 - \delta) \sum_{i=0}^{\infty} \delta^i E_t y_{t+i} + c \quad (4)$$

dois títulos, um de curto prazo rendendo r_t em t , com resgate em $t+1$, e outro, de longo prazo, rendendo R_t^n por período. Nesse caso, a equivalência de taxas de retorno para R\$ 1,00 investidos em cada alternativa seria: $(1 + R_t^n)^n = (1 + r_t)(1 + r_{t+1}) \cdots (1 + r_{t+n-1})$. Se log-linearizarmos a equação acima, com $\ln(1+z) \approx z$, considerando um ambiente de expectativas racionais, temos: $R_t^n = \frac{1}{n} E_t (r_t + r_{t+1} + \cdots + r_{t+n-1})$ que é um modelo de valor presente ligando R_t^n e r_t .

Pode-se interpretar θ como o coeficiente de proporcionalidade e δ como o fator de desconto. No caso da estrutura a termo, $Y_t = R_t$, $y_t = r_t$ e $\theta = 1$. Uma interpretação para c seria a de um prêmio de risco constante no tempo, que igualaria o valor presente da taxa de curto prazo com a de longo.

Como apontado por Campbell e Shiller (1987), há alguns problemas para se testar diretamente a equação (4). Primeiro, existem várias formas de testá-la, desde testes de restrições em um processo multivariado a testes de variância, e não parece claro como essas diferentes abordagens se relacionam. Segundo, a rejeição estatística de (4) pode não ter qualquer significado econômico. É perfeitamente possível que o modelo seja rejeitado ao nível de significância de 5% e ainda assim que a aderência do modelo à realidade seja relativamente alta. Finalmente, as variáveis y_t e Y_t usualmente requerem alguma transformação antes que a teoria de processos estocásticos estacionários possa ser aplicada.

Por esta última razão, Campbell e Shiller desenvolveram um teste para a validade da relação de valor presente entre y_t e Y_t quando as variáveis são $I(1)$, ou seja, estacionárias em primeiras diferenças. A parte central de seu procedimento se resume a um teste de restrições nos coeficientes de uma Auto-regressão Vetorial (VAR), onde este modelo é usado para fazer a previsão ótima descontada implícita em (4). A vantagem desta abordagem é que, apesar de não ser possível observar todo o conjunto de informação disponível aos agentes no mercado, a técnica econométrica empregada permite resumir boa parte da informação relevante nas variáveis utilizadas na construção do VAR.

Se as taxas de juros de curto prazo e longo prazo fossem séries “estacionárias”, poderíamos usar a equação (4) para testar o modelo de valor presente no contexto acima. Entretanto, existem evidências de que as taxas de juros seguem um processo com raiz unitária (Campbell e Shiller, 1987). Assim, a equação (4) relacionaria dois processos não estacionários com raízes unitárias. Para que se possa aplicar a teoria usual de séries temporais é necessário que estas sejam estacionárias – livres de raiz unitária.

Para contornarmos esse problema, pode-se aplicar uma transformação na equação (3) de forma a obter um relação entre séries estacionárias. Em primeiro lugar, definamos o *spread* entre a taxa longa e a taxa curta como:

$$S_t = R_t - r_t \quad (5)$$

Subtraindo-se r_t de ambos os lados da equação (3) temos:

$$S_t = -r_t + (1 - \delta) \sum_{i=0}^{\infty} \delta^i E_t r_{t+i} + c \quad (6)$$

Desmembrando os termos da equação acima,

$$S_t = E_t(-r_t + r_t - \delta r_t + \delta r_{t+1} - \delta^2 r_{t+1} + \delta^2 r_{t+2} - \delta^3 r_{t+2} + \delta^3 r_{t+3} - \dots) + c \quad (7)$$

$$= E_t [\delta(r_{t+1} - r_t) + \delta^2(r_{t+2} - r_{t+1}) + \delta^3(r_{t+3} - r_{t+2}) + \dots] + c \quad (8)$$

$$= \sum_{i=1}^{\infty} \delta^i E_t \Delta r_{t+i} + c \quad (9)$$

Também podemos expressar S_t como função das variações em R_t . De (3),

$$E_t(R_{t+1} - R_t) = E_t \{ (1 - \delta) [\delta(r_{t+1} - r_t) + \delta^2(r_{t+2} - r_{t+1}) + \dots] \} \quad (10)$$

$$= (1 - \delta) \sum_{i=1}^{\infty} \delta^i \Delta r_{t+i} \quad (11)$$

Logo,

$$\begin{aligned} S_t &= \frac{\delta}{1 - \delta} E_t \Delta R_{t+1} \\ &= \frac{\delta}{1 - \delta} \Delta R_{t+1} + \mu_{t+1} \end{aligned} \quad (12)$$

onde μ_{t+1} é um erro imprevisível do ponto de vista do conjunto de informação do agente em t . O *spread* entre a taxa curta e a taxa longa é dado por uma média ponderada de mudanças futuras nas taxas de curto prazo. A equação (13), por sua vez, diz que o *spread* é linear em relação à previsão ótima de ΔR_{t+1} .

As equações (9) e (13) ajudam a resolver alguns dos problemas discutidos anteriormente sobre a equação (3) e sua aplicação empírica. Em primeiro lugar, se Δr_t é estacionário, o que será verificado mais adiante, então, segue de (9) que S_t também o será, o que implica em cointegração entre R_t e r_t . Por (13), segue-se também que ΔR_t também é uma série estacionária. Assim, é possível utilizar S_t e Δr_t , ou S_t e ΔR_t , como séries estacionárias que resumem a história bi-variada de R_t e r_t para um teste econométrico do modelo de valor presente ou, de forma mais específica, para testar a Hipótese das Expectativas aplicada a estrutura a termo de juros no Brasil. Como optaremos pela combinação (S_t e Δr_t), o modelo a ser testado é originado da equação (9).

3.2 Testes do modelo e técnicas econométricas

A abordagem econométrica utilizada por Campbell e Shiller (1987, 1991) para testar os modelos discutidos na seção anterior é baseada em estimativas de modelos VAR. Para que se possa aplicar essa metodologia de séries temporais é necessário que as séries utilizadas no VAR sejam estacionárias. Pode-se observar, como já demonstramos acima, que a equação (3) relacionaria dois processos não estacionários com raízes unitárias, mas existiria um combinação linear entre estas duas séries que é estacionária. Em outras palavras, as taxas de juros de curto prazo e longo prazo seriam cointegradas, o que verificaremos mais adiante.

Deste modo, a primeira etapa da análise empírica é o estudo das propriedades estocásticas das séries usadas no trabalho, o que está vinculado ao estudo das raízes do polinômio auto-regressivo destas. Essa questão será investigada usando os testes de raiz unitária de Dickey e Fuller (1979) aumentado (ADF), e de Phillips e Perron (1988). Se as séries de taxa longa e taxa curta forem realmente $I(1)$, realizaremos testes de cointegração entre R_t e r_t , usando a técnica de Johansen (1988, 1991). Se confirmada a cointegração entre essas variáveis, com coeficiente $(1, -1)$, será usado um modelo VAR bi-variado para modelar S_t e Δr_t . Dado que, nesse caso, tanto S_t quanto Δr_t serão séries estacionárias, esse VAR não terá nenhuma raiz do determinante do polinômio auto-regressivo sobre o círculo unitário, sendo possível adotar o procedimento de Campbell e Shiller (1987) para testar a igualdade entre o valor presente esperado de Δr_t e S_t .

O procedimento proposto por Campbell e Shiller (1987) utiliza a seguinte representação VAR para S_t e Δr_t :

$$\begin{bmatrix} \Delta r_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} \Delta r_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (13)$$

onde os polinômios $a(L)$, $b(L)$, $c(L)$, $d(L)$ são todos de ordem $p-1$ no operador de defasagens L , e as raízes de $\left| I_2 - \begin{bmatrix} a(\lambda) & b(\lambda) \\ c(\lambda) & d(\lambda) \end{bmatrix} \lambda \right| = 0$ estão todas fora do círculo unitário.² Este VAR pode ser usado para se fazer uma previsão em múltiplos períodos de Δr_t e, além disso, também inclui-se S_t , que pode ser entendido como a previsão ótima do valor presente dos futuros Δr_t . Qualquer modelo $VAR(p)$ pode ser escrito na forma de um $VAR(1)$, conforme observamos abaixo:

²O VAR em (13) é de ordem p . Note que os polinômios $a(L)$, $b(L)$, $c(L)$, $d(L)$ são todos de ordem $p-1$ pois estão sendo aplicados sobre variáveis datadas em $t-1$.

$$\begin{bmatrix} \Delta r_t \\ \cdot \\ \cdot \\ \Delta r_{t-p+1} \\ S_t \\ \cdot \\ \cdot \\ S_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \dots & a_p & b_1 & \dots & b_p \\ 1 & & & & & \\ & \cdot & & 0 & & \\ & & 1 & & & \\ c_1 & \dots & c_p & d_1 & \dots & d_p \\ & & & 1 & & \\ & 0 & & & \cdot & \\ & & & & & 1 \end{bmatrix} \times \begin{bmatrix} \Delta r_{t-1} \\ \cdot \\ \cdot \\ \Delta r_{t-p} \\ S_{t-1} \\ \cdot \\ \cdot \\ S_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ 0 \\ \cdot \\ 0 \\ \varepsilon_{2t} \\ 0 \\ \cdot \\ 0 \end{bmatrix} \quad (14)$$

ou, de forma compacta:

$$Z_t = AZ_{t-1} + \mu_t \quad (15)$$

onde A é a matriz de coeficientes do VAR e μ_t é o vetor de choques.

Com esta formulação é possível fazer previsão para quaisquer i períodos à frente:

$$E [Z_{t+i}|H_t] = A^i Z_t \quad (16)$$

onde H_t é o conjunto de informação contendo valores correntes e defasados de z_t , ou seja, valores correntes e defasados de Δr_t e S_t .

A primeira implicação testável do modelo (mais fraca) é que S_t causa, no sentido de Granger, Δr_t , ou seja, S_t ajuda a prever Δr_t . A intuição por trás dessa afirmação é que, conforme visto anteriormente, S_t é a previsão ótima para uma soma ponderada dos valores *futuros* de Δr_t , condicionada ao conjunto de informação do agente. Naturalmente, pode-se realizar um teste de Causalidade de Granger para testar tal implicação.

A segunda implicação do modelo é que este impõe um conjunto de restrições no sistema dinâmico (14). Para derivar tais restrições, toma-se a esperança da equação (9) condicionada ao conjunto de informação H_t . O lado esquerdo da equação não será afetado já que S_t pertence ao conjunto H_t . Já o lado direito da equação torna-se um valor esperado descontado condicionado a H_t . Então, tem-se:

$$S_t = E \left[\sum_{i=1}^{\infty} \delta^i \Delta r_{t+i} \middle| H_t \right] \equiv S_t^* \quad (17)$$

A equação (17) nos diz que S_t deve ser igual à previsão ótima irrestrita das variações futuras de Δr_{t+i} , denotada por S_t^* . Essa igualdade será usada mais

adiante para verificar, de forma informal, a aderência do modelo de valor presente às observações de S_t .³

Além do teste informal, podemos também testar formalmente o modelo de valor presente e, por conseqüência, a Hipótese das Expectativas. Definindo h' e g' como vetores canônicos com $2p$ elementos, todos os quais são nulos à exceção do primeiro para h' e à exceção do elemento na linha $p + 1$ para g' , temos que $h'Z_t \equiv \Delta r_t$ e $g'Z_t \equiv S_t$. Assim, podemos reescrever (17) como:

$$S_t^* = \sum_{i=1}^{\infty} \delta^i h' [A^i Z_t] \quad (18)$$

Dado que assume-se que S_t e Δr_t são estacionários, a soma infinita acima deve convergir, o que resulta em:

$$S_t^* = h' \delta A (I - A \delta)^{-1} Z_t \quad (19)$$

Note que a equação (19) nos dá a série do *spread* ótimo. Como Z_t empilha o presente e o passado de S_t e Δr_t , fica claro que S_t^* será uma combinação linear do presente e o passado de S_t e Δr_t . Num ambiente de expectativas racionais, os agentes não cometem erros sistemáticos, mas isso não implica que não cometam erros de previsão. Logo, S_t^* e S_t não serão idênticos, embora não se deva observar diferenças sistemáticas se o modelo de valor presente for verdadeiro.

Para comparar formalmente os valores estimados com os valores observados do *spread*, i.e., comparar formalmente S_t^* com S_t , notamos que $S_t = g'Z_t$, de modo tal que, quando temos $S_t^* = S_t$, segue que:⁴

$$g'Z_t = h' \delta A (I - \delta A)^{-1} Z_t \quad (20)$$

Como a equação acima tem que valer para todas as realizações de Z_t , isso requer que,

$$g' = h' \delta A (I - \delta A)^{-1} \quad (21)$$

Pós-multiplicando a equação (21) acima por $(I - A \rho)^{-1}$, chegamos a:

$$g' (I - \delta A) = h' \delta A \quad (22)$$

³É nessa passagem que fica clara a vantagem de usar técnicas multivariadas, pois, ao usarmos técnicas de regressão em equação única restringimos o conjunto de informação do agente.

⁴Na verdade, iremos comparar formalmente os valores estimados com os valores observados do *spread* em relação à sua média, i.e., comparar $S_t^* - E(S_t^*)$ com $S_t - E(S_t)$.

A equação (22) define um conjunto de $2p$ restrições nos coeficientes do sistema VAR. Da estrutura da matriz A , as restrições impostas por (22) nos coeficientes individuais, a serem testadas por meio de um teste de Wald, são:

$$\begin{aligned} c_i &= -a_i, i = 1, \dots, p \\ d_1 &= (1/\delta) - b_1, \\ d_i &= -b_i, i = 2, \dots, p. \end{aligned} \quad (23)$$

Como pode ser notado, há duas formas de se testar (23). A primeira usa (21) e a segunda usa (22). Como é sabido, o teste de Wald não é invariante a transformações não lineares como a que se operou para transformar (21) em (22), e os resultados podem, potencialmente, ser diferentes usando (21) ou (22). Seguindo a literatura sobre o assunto, nos limitamos a implementar o teste baseado em (22), nos abstendo de testar hipóteses usando (21) pela complexidade envolvida na sua implementação.

Em se tratando de dados financeiros, há usualmente forte evidências de heterocedasticidade condicional dos dados⁵, o que requer o uso de estimativas de desvio-padrão robustas à heterocedasticidade, de modo a permitir a realização de testes de hipótese válidos – como o teste de Wald para as restrições descritas em (23). Sabemos, sob homocedasticidade dos erros do VAR, que a fórmula usual da matriz de variância-covariância dos estimadores dos coeficientes do VAR é,

$$\Gamma = \Sigma(X'X)^{-1} \quad (24)$$

onde X inclui os regressores do VAR e Σ é a matriz de covariância dos erros associados a este. Quando há heterocedasticidade, sabemos que $\widehat{\Sigma}(X'X)^{-1}$, $\widehat{\Sigma} = 1/T \sum_{t=1}^T \widehat{\varepsilon}_t \widehat{\varepsilon}_t'$, onde $\widehat{\varepsilon}_t$ é um vetor 2×1 empilhando os resíduos do VAR, não é um estimador consistente da matriz de variância-covariância dos estimadores dos coeficientes do VAR. Entretanto, seguindo a sugestão de White (1980), podemos usar o estimador robusto à heterocedasticidade,

$$\widehat{\Gamma} = (X'X)^{-1} X \widehat{V} X (X'X)^{-1} \quad (25)$$

onde define-se \widehat{V} como uma matriz diagonal contendo resíduos quadrados em sua diagonal, que é a alternativa usada aqui para testar hipóteses.

Além do teste de restrição nos coeficientes do VAR, como forma de testar se os desvios em relação à média do *spread* observado e do *spread* estimado são

⁵De fato, constatamos que os modelos estimados para os VARs, tanto para a taxa de 180 dias como de 360 dias, rejeitam fortemente a hipótese de homocedasticidade dos erros.

estatisticamente equivalentes, também podemos testar a validade do modelo de valor presente por meio do teste de razão de variâncias. Constrói-se uma estatística $F(\cdot)$, para testar a seguinte restrição:

$$\frac{Var(S_t^*)}{Var(S_t)} = 1$$

de igualdade entre a variância incondicional de S_t^* e de S_t . Sob o modelo de valor presente, a razão $Var(S_t^*)/Var(S_t) = 1$, mas será menor que um se S_t for muito volátil relativamente à informação sobre o futuro de Δy_t , por exemplo. Uma estatística complementar é a correlação entre S_t e S_t^* , que também deveria ser igual a um, já que se razão da variância e a correlação forem iguais a um, então $S_t = S_t^*$, como prevê o modelo.

Em suma, as proposições testáveis do modelo de valor presente para a estrutura a termo da taxa de juros, apropriado à Hipótese das Expectativas, a serem examinadas nas próximas seções são:

- Confirmar a existência de raiz unitária nas séries das taxas de juros de curto prazo e de longo prazo observadas para o Brasil, bem como testar a cointegração entre essas séries.
- Analisar se o *spread* entre as taxas de juros longa e curta ajuda a prever variações nas taxas curtas: S_t Granger-causa Δr_t ?
- Testar formalmente a igualdade das séries observada e estimada do *spread*, de modo a validar o modelo aqui descrito. Conforme descrito anteriormente, será utilizado o teste de Wald para verificar a validade das restrições impostas aos coeficientes do VAR pela equação (22).
- Comparação da volatilidade das séries S_t e S_t^* e a estimação da correlação entre as séries.

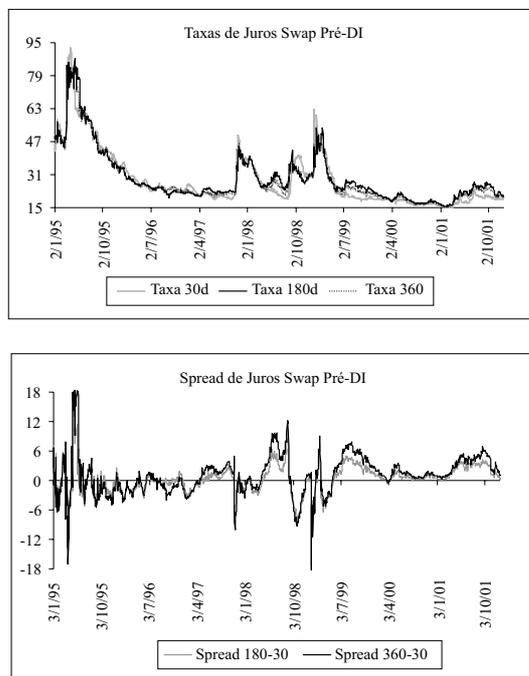
4. Base de Dados

A construção das séries r_t e S_t , definidas anteriormente como a taxa de juros de curto prazo e o *spread* entre a taxa de juros de longo prazo e de curto prazo, foi feita a partir das séries das taxas de swap pré-DI obtidas junto à BM&F, com frequência mensal (taxa do último dia útil do mês) e cobrindo o período de janeiro de 1995 a dezembro de 2001, perfazendo um total de 84 observações. Como *proxy* da taxa de curto prazo, usamos a taxa de juros nos contratos de swap pré-DI de

1 mês, ao passo que para as taxas de longo prazo utilizamos, em um primeiro momento, a taxa de swap pré-DI de 6 meses e, em seguida, a taxa de swap pré-DI de 1 ano. As séries de *spread* são construídas a partir da subtração das taxas nos dois períodos. A única transformação nos dados é que todas as séries serão trabalhadas como desvios em relação à sua média.

No quadro abaixo são apresentados, respectivamente, os gráficos das séries das taxas de 1 mês, de 6 meses, de 12 meses, bem como dos *spreads* entre as taxas de 6 meses de 1 mês e entre as taxas de 12 meses e de 1 mês. É interessante fazer uma análise do comportamento das séries antes de efetuarmos os testes de raiz unitária, a serem mostrados na próxima seção. Para os gráficos, apresentamos todos as séries sem qualquer transformação nos dados originais.

Figura 1



Note que os picos nas séries estão associados às diversas crises financeiras, desde à crise mexicana até a crise que desencadeou a mudança no regime cambial, passando pela crises da Ásia e da Rússia. Desde o último grande choque de juros, que se sucedeu à desvalorização cambial em 1999, o comportamento das taxas de

juros tem sido bem menos volátil. Vale notar também que a mera inspeção visual do gráfico nos permite fazer uma leitura superficial dos movimentos na estrutura a termo de juros. No segundo gráfico abaixo, cabe observar que dois dos principais vales estão associados a uma expectativa de redução agressiva dos juros que se sucederia à implementação de um choque de juros pela autoridade monetária.

No que diz respeito ao comportamento das séries de tempo em questão, há fortes indícios de que as séries de taxas de juros possuem uma raiz unitária, sendo então integradas de ordem um. Por outro lado, as séries de *spreads* apresentam claramente um comportamento estacionário. Para confirmar as suspeitas geradas pela análise gráfica relacionadas à existência ou não de uma raiz unitária, realizaremos na seção seguinte os testes apropriados.

5. Resultados Empíricos

5.1 Testes de raiz unitária e cointegração

A primeira etapa da análise empírica é o estudo das propriedades estocásticas das séries envolvidas. Neste contexto, o primeiro teste a ser realizado é o teste de raiz unitária; ver tabela 1, que apresenta as estatísticas ADF sem tendência, onde as suas respectivas defasagens foram escolhidas a partir da sugestão de Perron (1995), e também a estatística do teste Phillips e Perron (1988).

Tabela 1
Teste de Raiz Unitária

Variável	Defasagens	Teste ADF	Teste Philips-Perron
r_t	1	-2,19	-2,88
$S_{t(180)}$	4	-4,23**	-9,50**
$S_{t(360)}$	3	-4,21**	-8,23**
Variável	Defasagens	Teste ADF	Teste Philips-Perron
$R_{t(180)}$	1	-2,64	-2,38
$R_{t(360)}$	1	-2,71	-2,40

1) *: rejeita-se a 5%, **: rejeita-se a 1%

2) Todos os testes incluem intercepto.

Como é possível observar, há forte evidência da existência de uma raiz unitária nas séries de taxas de juros, tanto na taxa curta de 1 mês como nas taxas longas, tanto de 6 meses como de 12 meses. Os resultados dos testes obtidos para os *spreads* das taxas de juros, tanto na relação 180-30 dias como na relação 360-30 dias, sugerem fortemente que não há raiz unitária. A ausência de raiz unitária nas

séries de *spread* são um indício que as taxas curtas e longas cointegram. Porém testes formais de cointegração serão implementados a seguir.

Para confirmar a relação de longo prazo entre as variáveis R_t e r_t , nos prazos aqui referidos, usaremos análise de cointegração, baseada em verossimilhança, conforme propõe Johansen (1988, 1991). Previamente à realização do teste em questão, faz-se necessária a escolha da ordem do VAR. Para escolher a ordem do VAR contendo $R_{t(i)}$, $i = 180, 360$ e r_t , utilizou-se critérios de informação usuais (Hannan-Quinn, Schwarz e Akaike) e testes de diagnóstico. Levando-se em conta os testes de diagnóstico, optou-se pela escolha de 2 defasagens em nível.

Segundo os testes de autocorrelação serial, heterocedasticidade e normalidade, usando-se os resíduos, não encontramos autocorrelação serial, mas rejeitamos a hipótese de homocedasticidade dos erros. O teste de normalidade dos erros não é reportado aqui, mas não corroborou a hipótese de normalidade dos resíduos. Desvios da normalidade e da homocedasticidade dos erros podem ser um problema quando se faz um teste de cointegração baseado em verossimilhança – com o teste de Johansen. Felizmente, estudos Monte-Carlo como o de Gonzalo (1994) não detectam que esses desvios causem grandes problemas para o teste de Johansen.

Tabela 2
Testes de Diagnóstico nos Resíduos do VAR

Modelo Contendo	Autocorrelação Serial	
	$p = 8$	
	LM	valor-p
$R_{t(180)}$	0,35	0,986
$R_{t(360)}$	0,34	0,987

Modelo Contendo	Heterocedasticidade	
	χ^2	valor-p
$R_{t(180)}$	75,97	0,000
$R_{t(360)}$	78,66	0,000

- 1) o teste de heterocedasticidade foi realizado com quadrados e produtos cruzados.
- 2) no teste de correlação serial p corresponde a ordem a ser testada.

Dada a escolha do número ótimo de defasagens, o próximo passo é verificar as possíveis relações de longo prazo entre R_t e r_t através da análise de cointegração, baseada em verossimilhança, conforme propõe Johansen (1988, 1991). Em um primeiro momento, analisaremos a relação de longo prazo entre a taxa de 180 dias e de 30 dias para, em seguida, fazermos o mesmo com a taxa de 360 dias:

Tabela 3
Teste de Johansen

H_0 : posto= p	Est. λ_{\max}	95%	Est. do Traço	95%	Defasagens em nível no VAR
$p = 0$	27,62	15,67	33,97	19,96	2
$p \leq 1$	6,34	9,24	6,34	9,24	2

Coeficientes de cointegração normalizados: 1 equação de cointegração

$R_{t(180)}$	r_t
1	-0,937

O resultado do teste de cointegração sugere que as séries $R_{t(180)}$ e r_t cointegram com apenas um vetor de cointegração, tanto se analisado pela estatística do λ_{\max} quanto pela estatística do traço. A hipótese nula de que as séries não cointegram é rejeitada ao nível de significância de 5% e 1% para as duas estatísticas. Já a hipótese nula de que o número de vetores de cointegração é menor ou igual a um não pode ser rejeitada tanto ao nível de significância de 1% como de 5%. O vetor de cointegração sugerido é $(1 - 0,937)$. Na tabela abaixo, testamos se esse vetor é estatisticamente igual a $(1 - 1)$, pelo teste de cointegração restrita. Note que não é rejeitada a hipótese nula de que o vetor de cointegração é estatisticamente igual a $(1 - 1)$.

Tabela 4
Teste de Cointegração Restrita

Restrição	Resultado
$R_t = 1; r_t = -1,$	valor-p = 0,127

No que diz respeito à cointegração entre a taxa de 360 dias e a taxa de 30 dias, obtemos resultados semelhantes, como é evidenciado nos quadros abaixo:

Tabela 5
Teste de Johansen

H_0 : posto= p	Est. λ_{\max}	95%	Est. do Traço	95%	Defasagens em nível no VAR
$p = 0$	21,58	15,67	27,63	19,96	2
$p \leq 1$	6,05	9,24	6,05	9,24	2

Coeficientes de cointegração normalizados: 1 equação de cointegração

$R_{t(360)}$	r_t
1	-0,902

O resultado do teste de cointegração também sugere que as séries $R_{t(360)}$ e r_t cointegram com um vetor de cointegração, tanto se analisado pela estatística do λ_{\max} quanto pela estatística do traço. O vetor de cointegração sugerido é $(1 - 0,902)$. Na tabela abaixo, realizamos mais uma vez o teste de cointegração restrita, de modo a checar se esse vetor é estatisticamente igual a $(1 - 1)$. Note

que é bastante difícil rejeitar a hipótese nula de que o vetor de cointegração é estatisticamente igual a $(1 - 1)$.

Tabela 6
Teste de Cointegração Restrita

Restrição	Resultado
$R_t = 1; r_t = -1,$	p-valor = 0,123

Deste modo, parece claro que as séries de taxas de juros de curto prazo e longo prazo cointegram e têm um vetor de cointegração estatisticamente igual a $(1 - 1)$. Assim, conforme verificamos anteriormente, a estacionariedade do *spread* entre a taxa longa e a taxa curta indicava fortemente que as séries de taxas de juros cointegravam, dado que o *spread* nada mais é do que uma combinação linear das duas séries em questão. Isto foi ratificado pelos testes de cointegração acima.

Em que pese a evidência dos desvios da normalidade e da homocedasticidade dos erros, os resultados dos testes de Johansen para ambos os *spreads*, aliados à evidência de que os *spreads* são $I(0)$, obtidas nos testes de raiz unitária de Phillips e Perron (1988), que não fazem hipótese alguma sobre normalidade e homocedasticidade dos erros, nos dão confiança de seguir adiante presumindo que os *spreads* sejam $I(0)$.

5.2 Estimação do VAR e teste de causalidade

De acordo com o método econométrico desenvolvido na seção anterior, uma vez verificada a cointegração entre a taxa curta e a taxa longa, estimaremos um VAR irrestrito com Δr_t e S_t . A escolha da ordem do VAR é feita a partir da estrutura do VAR original em nível – o que implicaria em escolher também duas defasagens. De qualquer forma, apenas para confirmar nossa escolha, usamos, mais uma vez, critérios de informação e testes de diagnóstico, corroborando nossa escolha original de 2 defasagens. Há evidência de não autocorrelação serial nos erros do VAR, embora claramente haja heterocedasticidade nos mesmos, que, posteriormente, será levada em conta ao usarmos estimativas robustas para a matriz de variância-covariância dos estimadores, a partir do método de White (1980) descrito anteriormente.

Tabela 7
Testes de Diagnóstico nos Resíduos do VAR

Autocorrelação Serial		
$p=8$		
Modelo Contendo	LM	p-valor
$S_{t(180)}$	4,55	0,335
$S_{t(360)}$	4,51	0,341

Heterocedasticidade		
Modelo Contendo	χ^2	p-valor
$S_{t(180)}$	10,32	0,00
$S_{t(360)}$	96,15	0,00

- 1) o teste de heterocedasticidade foi realizado utilizando quadrados e produtos cruzados dos regressores
 2) no teste de correlação serial p corresponde a ordem a ser testada.

Testamos, na tabela abaixo, a causalidade de Granger entre os *spreads* e as taxas curtas. Como podemos observar há causalidade bi-direcional e não só dos *spreads* para as taxas curtas.

Tabela 8
Testes de Causalidade Granger

Teste de Causalidade de Granger	Estatísticas Descritivas			
	$H_0 :$	χ^2	p-value	Equação no Sistema R^2
a) Sistema contendo $S_{t(180)}$ e Δr_t				
$S_{t(180)}$ não causa granger Δr_t		17,47	0,000	Equação Δr_t 0,32
Δr_t não causa granger $S_{t(180)}$		27,39	0,000	Equação $S_{t(180)}$ 0,29
b) Sistema contendo $S_{t(360)}$ e Δr_t				
$S_{t(360)}$ não causa granger Δr_t		12,86	0,000	Equação Δr_t 0,28
Δr_t não causa granger $S_{t(360)}$		29,77	0,000	Equação $S_{t(360)}$ 0,34

As estimativas do VAR sugerem que as variações nas taxas de curto prazo são razoavelmente previsíveis, tendo em vista o R^2 de aproximadamente 30% para ambos os sistemas na equação de variações nas taxas curtas. O R^2 da equação para S_t também é razoável (29% e 34% para ambos os modelos, respectivamente).

5.3 Construção do *spread* ótimo e teste do MVP

O passo seguintes é verificar formalmente a validade da Hipótese das Expectativas – ou dos Modelos de Valor Presente – para a estrutura a termo de juros no Brasil. Para isso, em primeiro lugar devemos comparar a volatilidade das séries

S_t e S_t^* , computando o seu coeficiente de correlação. Em seguida, testaremos as restrições impostas nos coeficientes do VAR pela equação (22) usando o teste de Wald.

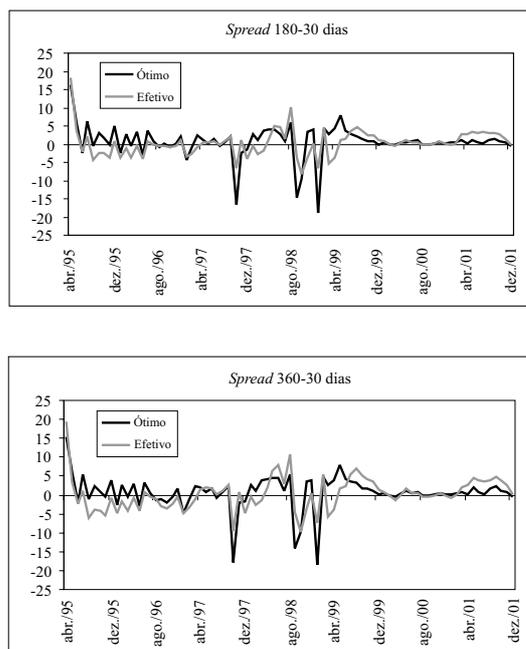
A partir da equação (19) e supondo que a taxa de desconto é igual a 2,5%, por razões aqui já explicitadas, chegamos às seguintes equações, para os dois modelos em discussão:

$$S_{t(180)}^* = -0,1596 \times \Delta r_t - 0,307 \times \Delta r_{t-1} + 0,7378 \times S_{t(180)} - 0,3623 \times S_{t-1(180)} \quad (26)$$

$$S_{t(360)}^* = -0,1574 \times \Delta r_t - 0,294 \times \Delta r_{t-1} + 0,6826 \times S_{t(360)} - 0,2968 \times S_{t-1(360)} \quad (27)$$

Antes de explicitarmos os resultados obtidos nos testes formais, fazemos uma análise informal através da inspeção dos gráficos das séries do *spread* ótimo e do *spread* observado, tanto para o prazo de 180 dias como de 360 dias, sempre lembrando que as séries são expressas como desvios em torno da sua média.

Figura 2



Analisando a figura 2, percebe-se que as séries guardam uma correlação positiva, bem como suas variâncias não parecem divergir consideravelmente, conforme sugerido pela teoria.

Dadas as restrições obtidas a partir do modelo em (22), foram calculadas as estatísticas de Wald, a partir de diversas taxas de desconto (de 0% a 10%). Estas são apresentadas na tabela 9. Utilizando a sugestão de Campbell e Shiller (1987),⁶ exibiremos aqui apenas os resultados associados à taxa de desconto de 2,5%. Vale ressaltar que não é relevante a magnitude com que os resultados se alteram para as demais taxas de desconto. Os p-valores dos testes de Wald, usando uma estimativa da matriz de variância-covariância robusta à heterocedasticidade para ambos os VARs, sugerem que o modelo de valor presente não pode ser rejeitado para os dados brasileiros, a um nível de significância de 5%. Entretanto, ao nível de significância de 10% rejeitamos a validade dessas restrições.

Tabela 9
Testes do Modelo de Valor Presente

Teste de Wald para restrições conjuntas (eq. (22))		
Modelo Contendo	χ^2	p-value
$S_{t(180)}$	9,141	0,0577
$S_{t(360)}$	8,531	0,0740
Teste de Razão de Variância		
$H_0 :$	F	p-value
$Var(S_{t(180)}^*)/Var(S_{t(180)}) = 1$	1,59	0,0199
$Var(S_{t(360)}^*)/Var(S_{t(360)}) = 1$	1,11	0,3209

1) $S_{t(180)}^*$ e $S_{t(360)}^*$ são os *spreads* ótimos impostas pelo modelo.

Finalmente, os testes de razão de variância da série observada S_t e da prevista pelo modelo S_t^* , também reportados na tabela 9, fornecem sinais ambíguos. Por um lado, a hipótese de razão unitária entre a variância estimada e a observada é rejeitada fortemente para o *spread* de 180 dias, enquanto o contrário acontece para o *spread* de 360 dias, onde o p-valor de 0.3209 indica que não se pode rejeitar a hipótese de que as variâncias das séries do *spread* ótimo e do *spread* observado são iguais. O coeficiente de correlação entre $S_{t(180)}^*$ e $S_{t(180)}$ é de 0,56 e de 0,68 entre $S_{t(360)}^*$ e $S_{t(360)}$. Note que os resultados do teste de Wald também sugerem que é mais difícil rejeitar a validade do modelo quando estamos trabalhando com o prazo de 360 dias para a taxa longa. Apesar das rejeições de algumas hipóteses da teoria, pode-se concluir que os dados são parcialmente favoráveis à Hipótese das

⁶O fator de desconto δ é igual a $1/(1+r)$, onde r é a taxa de desconto que, segundo os autores, pode ser definida como a taxa de juros de longo-prazo média expressa em taxa mensal. Assim, utilizamos 2,5% como uma proxy da taxa longa média em bases mensais.

Expectativas para a estrutura a termo de juros no Brasil, principalmente quando trabalhamos com a taxa de juros de 1 mês e de 12 meses como aproximações, respectivamente, da taxa de juros de curto prazo e de longo prazo da economia brasileira.

5.4 Comparação com resultados internacionais e resultados para o Brasil

Conforme abordado anteriormente, a Hipótese das Expectativas tem sido amplamente testada e discutida na literatura. Embora existam modelos bem mais sofisticados sobre a estrutura a termo de juros (Brito, 2001), a evidência empírica contra a HE está longe de ser conclusiva e oferece diversos *puzzles* interessantes.

Em um artigo clássico, Campbell e Shiller (1987) encontram evidências parciais sobre a validade da Hipótese das Expectativas, em linha com os resultados obtidos neste trabalho. Se, por um lado, o modelo é formalmente rejeitado pelo teste de Wald, por outro, o *spread* ótimo tem uma correlação positiva bastante alta com o *spread* observado. Os testes de razão de variância também não conseguem rejeitar a hipótese de que a variância das duas séries sejam estatisticamente iguais. Campbell e Shiller (1987) também encontraram os resultados de que as variáveis defasadas no VAR do *spread* e das variações na taxa curta têm um poder explicativo razoável sobre as mudanças nas taxas de curto prazo (R^2 de cerca 20%, para um R^2 em torno de 30% no nosso modelo) o que, para esses autores, seria uma evidência contrária à visão, por exemplo, de Mankiw e Miron (1986), para os quais as mudanças nas taxas de curto prazo são essencialmente imprevisíveis no período do pós-guerra dos EUA. Campbell e Shiller (1987) também chegam a uma forte evidência que os *spreads* Granger-causam as variações nas taxas de curto prazo.

Num outro trabalho clássico sobre o tema, Campbell e Shiller (1991) chegam a diferentes resultados a depender da especificação das regressões e dos prazos das taxas de juros analisadas, concluindo que a taxa de juros de longo prazo da economia americana não se comporta como prevê a teoria, ao passo que as futuras taxas de juros de curto prazo movem-se no sentido previsto.

Nos demais artigos que seguem a metodologia descrita por Campbell e Shiller (1987, 1991), os resultados são divergentes. Cuthbertson (1996), trabalhando com taxas de juros em diferentes prazos para dados da Inglaterra, conclui que os resultados de cointegração e do teste de Wald para a validade das restrições lineares são consistentes à Hipótese das Expectativas apenas em prazos curtos (até 6 meses). Taylor (1992), também trabalhando com dados ingleses, conclui pela

rejeição formal da Hipótese das Expectativas, não só pelos resultados obtidos a partir do teste de Wald, como também pelos testes de causalidade de Granger e de razão de variância. Hardouvelis (1994), também trabalhando com a estimação do spread *ótimo* e sua comparação com o *spread* efetivo, chega a resultados bastante animadores para a Hipótese das Expectativas. O resultados do teste de Wald não permitem rejeitar a HE em nenhum dos sete países examinados pelo autor. Com exceção dos Estados Unidos, o coeficiente de correlação entre os *spreads* é muito próximo da unidade.

Em um artigo recente, Bekaert et alii (2002), por sua vez, defendem que o poder de previsão dos *spreads* parece muito fraco para as taxas de curto prazo em horizontes curtos de tempo, mas é significativamente melhor em horizontes mais longos.

Com relação aos resultados obtidos para testes da Hipótese das Expectativas, com dados brasileiros, há pelo menos dois artigos a considerar: Tabak e Andrade (2001) e Brito et alii (2003). Além disso, há o *survey* de Rossi (1996), discutindo diversos testes para a Hipótese das Expectativas e as inerentes dificuldades de se implementá-los ao caso brasileiro. Os resultados de Tabak e Andrade rejeitam a Hipótese das Expectativas (conjuntamente com expectativas racionais) em testes diretos usando regressões individuais. Estas representam a relação da média ponderada da taxa curta com o *spread*, para diferentes horizontes, incluindo ou não *proxies* para prêmios de risco. Já Brito, Duarte e Guillén usam uma gama de métodos para testar a Hipótese das Expectativas, chegando a resultados ambíguos, parecidos com os aqui obtidos. É oportuna a colocação destes últimos ao final de seu artigo: “*A validade da Hipótese das Expectativas para o Brasil continua uma questão em aberto, assim como o é para o resto do mundo.*”

Por fim, vale fazer uma conjectura a respeito dos resultados empíricos brasileiros. Em nossa opinião, talvez a rejeição da Hipótese das Expectativas esteja intimamente ligada ao fato de que se usou, até o momento, um prêmio de risco constante ao longo do tempo nos modelos econométricos que serviram de base para os testes empíricos. Numa economia tão sujeita a choques como a brasileira, talvez esses modelos devessem incluir prêmios de risco variantes no tempo, quiçá baseados em modelos estruturais CCAPM, como os discutidos e testados por Issler e Piqueira (2000).

6. Conclusão

Depois de mais um século, desde Fisher (1896), a Hipótese das Expectativas continua sendo a teoria mais conhecida e mais intuitiva sobre as flutuações na

estrutura a termo de juros. Numerosos testes empíricos sobre o tema já foram realizados, com diferentes metodologias, países e segmentos da estrutura a termo de juros. Os resultados sugerem a existência de alguns *puzzles* não explicados teoricamente e diversas têm sido as tentativas não só de contornar a aparente ambiguidade empírica da HE, como também de demonstrar a incapacidade da HE em explicar alguns fatos amplamente percebidos nos mercados financeiros mundiais.

Neste trabalho usou-se o arcabouço de modelos de valor presente (MVP), descrito em Campbell e Shiller (1987, 1991), para testar as implicações da Hipótese das Expectativas em um modelo auto-regressivo vetorial envolvendo o *spread* entre a taxa longa e a taxa curta e a primeira diferença da taxa de juros de curto prazo. Em outras palavras, testou-se aqui a racionalidade e a previsibilidade na estrutura a termos de juros do Brasil.

A partir de uma gama de testes empíricos – teste de Wald nos coeficientes do VAR, testes de causalidade de Granger, testes de razão unitária de variâncias – pôde-se concluir que as evidências empíricas são apenas parcialmente favoráveis à Hipótese das Expectativas com dados brasileiros. Como exemplo, se tomarmos um nível de significância de 10%, rejeita-se a hipótese que as restrições lineares derivadas do modelo são válidas, por meio de um teste de Wald. Isto implicaria a rejeição formal dos modelos, tanto para a taxa de 180 dias como de 360 dias. Os testes de razão de variância entre o *spread* teórico e o *spread* efetivo também não fornecem evidências conclusivas. No modelo onde a taxa de 12 meses representa a taxa de juros de longo prazo, não podemos rejeitar a hipótese que as variâncias do *spread* ótimo e do *spread* efetivo são estatisticamente iguais. Por outro lado, os resultados para o modelo com a taxa de 6 meses revelam o contrário. Nos testes de causalidade de Granger, foi apresentada causalidade bi-direcional entre S_t e Δr_t para ambos os modelos.

O presente esforço empírico pode ser entendido como uma contribuição a um campo de pesquisa ainda em fase preliminar no Brasil. Como exemplo da relevância do tema para o caso brasileiro, o Banco Central do Brasil, em sua modelagem macroeconômica, tem destacado o papel da estrutura a termo de juros e desde 2001 introduziu uma hipótese exógena sobre a inclinação da curva de juros para os seus exercícios de simulação. Ainda há muito que ser entendido não só sobre o seu processo de formação, como também sobre os efeitos que as movimentações na estrutura a termo produzem sobre o nível de atividade econômica.

Em linha com as conclusões empíricas verificados em outros países, os resultados obtidos neste trabalho não chegam a evidências conclusivas e permitem apenas admitir que não é possível rejeitar completamente a Hipótese das Expec-

tativas para os dados brasileiros. Há um campo aberto de pesquisa para testar teorias alternativas sobre a estrutura a termo de juros no Brasil e, seguindo a linha de Taylor (1992) e a sugestão de Tabak e Andrade (2001), examinar o papel do gerenciamento da dívida pública pode ser um dos caminhos a ser trilhado. Adicionaríamos, à lista de pesquisas futuras, tentar modelar o prêmio de risco variante no tempo, conjuntamente ao spread e ao valor presente das taxas de curto prazo. Os primeiros poderiam ser baseados em modelos estruturais CCAPM, como os discutidos e testados por Issler e Piqueira (2000).

Referências

- Bekaert, G., Wei, M., & Xing, Y. (2002). Uncovered interest rate parity and the term structure. NBER Working Papers, 8795.
- Brito, R. (2001). Essays on the monetary aspects of the term structure of nominal interest rates. Tese de Doutorado em Economia, EPGE-FGV.
- Brito, R., Duarte, A. J. M., & Guillén, O. C. T. (2003). O prêmio pela maturidade na estrutura a termo das taxas de juros brasileira. Anais do XXIV Encontro Brasileiro de Econometria.
- Campbell, J. (1995). Some lessons from the yield curve. *Journal of Economic Perspectives*, 9:129–152.
- Campbell, J. (2000). Asset pricing at the millenium. *Journal of Finance*, 55:1515–1540.
- Campbell, J., Lo, W., & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Campbell, J. & Shiller, R. (1987). Cointegration and tests of present value models. *Journal of Political Economy*, 95:1062–1088.
- Campbell, J. & Shiller, R. (1991). Yield spread and interest rate movements: A bird's eye view. *Review of Economic Studies*, 58:495–514.
- Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53:385–408.
- Cuthbertson, K. (1996). The expectations hypothesis of the term structure: The UK interbank market. *The Economic Journal*, 106:578–592.

- Dickey, D. & Fuller, W. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–31.
- Engle, R. F. & Granger, C. W. J. (1987). Cointegration and error-correction: Representation, estimation and testing. *Econometrica*, 55:251–276.
- Fama, E. (1970). Efficiency capital markets: A review of theory and empirical work. *Journal of Finance*, 25:383–417.
- Fama, E. & Bliss, R. (1987). The information in long-maturity forward rates. *American Economic Review*, 77:680–692.
- Fisher, I. (1896). Appreciation and interest. *Publications of the American Economic Association*, 11:21–29.
- Froot, K. (1989). New hope for the expectation hypothesis of the term structure of interest rates. *The Journal of Finance*, 44:283–305.
- Gerlach, S. & Smets, F. (1997). The term structure of Euro-rates: Some evidence in support of the expectations hypothesis. *Journal of International Money and Finance*, 16:305–321.
- Gonzalo, J. (1994). Five alternative methods of estimating long run relationships. *Journal of Econometrics*, 60:203–233.
- Hardouvelis, G. (1994). The term structure spread and future changes in long and short rates in the G7 countries. *Journal of Monetary Economics*, 33:255–283.
- Issler, J. V. & Piqueira, N. S. (2000). Estimating relative risk aversion, the discount rate, and the intertemporal elasticity of substitution in consumption for Brazil using three types of utility function. *Brazilian Review of Econometrics*, 20:200–238.
- Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control*, 12:231–254.
- Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59:1551–1580.
- Jondeau, E. & Ricart, R. (1999). The expectation hypothesis of the term structure: Tests on US, German, French, and UK Euro-rates. *Journal of International Money and Finance*, 18:725–750.

- Longstaff, F. (2000). The term structure of very short-term rates: New evidence for the expectation hypothesis. *Journal of Financial Economics*, 58:397–415.
- Mankiw, N. G. & Miron, J. (1986). The changing behaviour of the term structure of interest rates. *Quarterly Journal of Economics*, 101:211–228.
- Perron, P. (1995). Lecture notes. Université de Montréal. mimeo.
- Phillips, P. C. & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75:335–46.
- Rossi, J. W. (1996). A estrutura a termo de taxa de juros: Uma síntese. *Pesquisa e Planejamento Econômico*, 26:521–548.
- Senna, F. & Issler, J. V. (2000). Mobilidade de capitais e movimentos da conta-corrente do Brasil: 1947-1997. *Estudos Econômicos*, 30:493–523.
- Shiller, R. (1979). The volatility of long-term interest rates and expectations models of the term structure. *Journal of Political Economy*, 87:1190–1219.
- Shiller, R. (1990). The term structure of interest rates. In Friedman, B. & Hahn, F., editor, *Handbook of Monetary Economics*, pages 627–772. Amsterdam, North-Holland.
- Tabak, B. M. & Andrade, S. C. (2001). Testing the expectation hypothesis in the brazilian term structure of interest rates. Working Paper Series, 18. Banco Central do Brasil.
- Taylor, M. (1992). Modelling the yield curve. *Economic Journal*, 102:524–537.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, 48:817–838.
- White, H. (1984). *Asymptotic Theory for Econometricians*. Academic Press, New York, NY.