

Fundação Getúlio Vargas - RJ
Escola de Pós Graduação em Economia

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

Autor: Alexandre Maia Correia Lima

Orientador: João Victor Issler

Dissertação submetida à congregação da EPGE
em 11 de abril de 2002 para a obtenção do grau de
Mestre em Economia

Rio de Janeiro
Abril de 2002

Agradecimentos

Agradeço a meus pais e irmãos, pelo apoio incondicional e orientação.

Ao meu tio Nelson Bessa Maia, que acompanhou de perto a minha formação como economista, sempre disposto a ajudar e orientar, e a minha avó Itamê, pelos conselhos sempre proveitosos.

Ao professor João Victor Issler, pela orientação e suporte ao longo de todo trabalho, e aos membros da banca, professores Cristiano Augusto Fernandes Coelho e Eduarda Cunha de La Rocque, pelas pertinentes sugestões.

Aos meus colegas de turma da EPGE de 1998, pelo companheirismo e pelas palavras de incentivo quando resolvi retomar a tese. A Guilherme Vilazante, sempre disposto a ajudar na parte computacional.

Aos amigos de longas datas, em especial João Marcelo Albuquerque e Samuel Dias, por sempre terem incentivado as minhas decisões.

A David Beker, pelo companheirismo e pela presteza no envio da base de dados.

Aos professores da EPGE, pela preocupação com a excelência acadêmica, e a todos os alunos e funcionários com quem convivi não apenas em 1998 e 1999, mas também no curto período de 2002 em que me dediquei integralmente à redação da tese.

Finalmente, ao Governo Brasileiro, pelo apoio financeiro recebido.

Resumo

Utilizando dados financeiros brasileiros da BM&F, testa-se a validade do modelo de valor presente (MVP) 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, a previsibilidade na evolução das taxas de juros. Neste artigo é realizada uma análise multivariada em um arcabouço de séries temporais utilizando a técnica de Autorregressões Vetoriais. Os resultados empíricos aceitam apenas parcialmente a Hipótese das Expectativas para a estrutura a termo de juros brasileira.

Sumário

1	Introdução	2
2	Uma Resenha da Literatura sobre a Hipótese das Expectativas	4
2.1	Aspectos Conceituais e Evidências Empíricas	4
3	O Modelo	9
3.1	Os Modelos de Valor Presente - MVP e Hipótese das Expectativas - HE	9
4	Metodologia	12
4.1	Técnica Econométrica	12
5	Base de Dados	18
6	Resultados Empíricos	19
6.1	MVP com Hipótese das Expectativas	19
6.1.1	Testes de Raiz Unitária e Cointegração	19
6.1.2	Estimação do vetor autorregressivo e Teste de Causalidade	23
6.1.3	Teste do MVP e Construção do <i>spread</i> ótimo	25
6.1.4	Comparação com Resultados Internacionais	28
7	Conclusão	30

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, 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. Não obstante a capacidade dos bancos centrais de estabelecer a taxa de juros de curto-prazo, admite-se que é a sua interação com as taxas de juros mais longas que produz os efeitos plenos de uma dada decisão de política monetária sobre a atividade econômica.

A explicação mais antiga e amplamente testada sobre as flutuações na estrutura a termo da taxa de juros 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.

No entanto, os 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 EH para os EUA, atribuindo o fraco poder de previsão do *spread* sobre as mudanças nas taxas longas a erros expectationais sistemáticos.

Mais recentemente, Gerlach e Smets (1997) e Longstaff (2000) trabalharam com o seg-

mento 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 econométricos mais utilizados na literatura é o que foi desenvolvido por Campbell e Shiller (1987,1991), que consiste na estimação de um VAR (autoregressão vetorial) bivariado para a variação em primeiras diferenças nas taxas de juros de curto-prazo e o *spread* entre a taxa longa e a taxa curta e será a metodologia adotada neste trabalho, discutida mais a fundo nas próximas seções. Alguns autores, como Taylor (1992) e Cuthbertson (1996), também seguiram essa metodologia para testar a HE em seus respectivos países.

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, Andrade e Tabak (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. Os autores utilizam duas diferentes metodologias baseadas em regressões de uma equação, 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 na regressão uma variável que seja uma *proxy* do prêmio de risco da economia brasileira e testam se o seu coeficiente é estatisticamente significativo.

O presente trabalho é uma contribuição à esta literatura e busca agregar novos resultados à discussão ainda preliminar no Brasil sobre a importância da estrutura a termo de juros que serve, entre outros aspectos, para a modelagem macroeconômica no âmbito do regime de metas de inflação. Como já citado acima, utiliza-se aqui técnicas econométricas de análise multivariada, aplicando o arcabouço de modelos de Autorregressões Vetoriais (VAR) para testar a validade da HE para o Brasil. Os coeficientes do VAR são utilizados então para construir a série do *spread* ótimo, que será comparada em seguida com a série observada para o *spread*, usando-se, entre outros testes, um teste de Wald para a restrição nos coeficientes do VAR.

Na seção 2 apresenta-se uma breve resenha da literatura, na seção 3 discute-se mais for-

malmente a Hipótese das Expectativas, no bojo dos modelos de valor presente, apresentando a versão desenvolvida por Campbell e Shiller (1987,1991). A seguir, na seção 4 descreve-se a metodologia econométrica a partir do uso de modelos VAR, bem como os outros testes econométricos e estatísticos que serão implementados. A seção 5 faz uma análise da base de dados utilizada, a seção 6 apresenta os resultados empíricos encontrados e a seção 7 apresenta as conclusões.

2 Uma Resenha da Literatura sobre a Hipótese das Expectativas

2.1 Aspectos Conceituais e Evidências Empíricas

A Hipótese das Expectativas (HE) da estrutura a termo de juros estabelece, de uma forma mais ampla, que a inclinação da curva de juros está associada às expectativas sobre as futuras taxas de juros. É possível depreender da HE que a diferença dos retornos esperados entre estratégias de investimentos de curto-prazo e longo-prazo é constante, ainda que possa ser acrescentado a esta constante um prêmio de risco invariante no tempo. Assim, a futura taxa de juros seria formada como uma média da taxa corrente de juros e das taxas esperadas para o futuro, deduzidas a partir da estrutura a termo. De uma forma simplificada, seguindo a definição de Longstaff (2000), isto pode ser escrito como:

$$E [R_{t+n}/H_t] = Y_{t(n)} + a_n, \quad (1)$$

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 no período t , $Y_{t(n)}$ é a taxa de longo-prazo em t e a_n é um prêmio de risco invariante no tempo, ainda que possa diferir em distintos horizontes n . Assim, $Y_{t(n)} + a_n$ é o valor esperado condicional da taxa de juros de curto-prazo média de t até $t + n$. Quando a_n é zero, esta forma de representação da HE é geralmente chamada de *Hipótese Pura das Expectativas* e está associada ao trabalho seminal de Fisher (1896). Nas palavras desse autor, 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. Por trás dos teorizadores iniciais da HE, estava a percepção de que as taxas longas correntes deveriam ser iguais às taxas de juros de curto-prazo no futuro (mais, talvez, uma constante).

A intuição por trás da formulação acima é extremamente simples. Como exemplo, suponhamos que houvesse uma expectativa de que as taxas de curto-prazo fossem bem menores que a taxa longa, em média, ao longo desse período. Assim, os "tomadores" teriam todo o incentivo a não tomar dinheiro a prazos longos e, ao invés disso, tomar emprestado à taxa curta, refinanciando-se repetidamente até a data de vencimento. Se a expectativa fosse exatamente oposta, os "doadores" teriam o incentivo a não emprestar a prazos longos, ou seja, não comprariam títulos de longo-prazo e sim uma série de papéis curtos. Assim, pela teoria embutida na Hipótese das Expectativas, que está relacionada à teoria dos mercados eficientes¹, quaisquer movimentos nas taxas de juros mais longas deveriam estar associados a novas informações sobre a média das taxas de juros no futuro. A curva de juros sinalizaria então as expectativas dos agentes sobre as futuras taxas de juros.

De acordo com Bekahert e Hodrick (2001), há três razões potenciais para a rejeição empírica da Hipótese das Expectativas. Em primeiro lugar, a maioria dos testes empíricos assume que as expectativas são formadas racionalmente; no entanto, pode ser que investidores irracionais façam erros de previsão sistemáticos e a habilidade dos investidores racionais de lucrar esteja limitada pela sua aversão ao risco. Segundo, a presença de um prêmio de risco, que não seja invariante no tempo, significa que os testes padrões da HE omitem as variáveis que explicam esse prêmio de risco. E, por último, os testes em si podem levar a falsas rejeições do modelo, o que é explicado pelas suas fracas propriedades em amostras finitas. Campbell (2000) também corrobora esse último ponto e argumenta que os problemas em pequenas amostras surgem porque os *spreads* são persistentes e suas inovações são correlacionadas com o retorno dos títulos.

Uma das visões alternativas à Hipótese das Expectativas aponta que mudanças na estrutura a termo estariam relacionadas, na verdade, a mudanças nos "retornos em excesso" requeridos dos papéis longos sobre os papéis curtos. Os "retornos em excesso" subiriam por um aumento do risco, manifestado através de um aumento da volatilidade no mercado de títulos, e explicado talvez por uma maior incerteza em relação à política monetária ou até por um aumento na aversão ao risco dos participantes do mercado, provocado, eventualmente, por perdas em operações alavancadas. Dessa forma, uma decisão de política monetária que não afete a expectativa sobre as taxas de juros mais adiante não deveria mover as taxas

¹Na definição de Fama (1970), a eficiência semi-forte dos mercados de ativos ocorre quando os preços revelam toda a informação pública disponível ou, dito de outra forma, os ganhos de arbitragem já foram esgotados.

futuras, ao passo que sob a visão alternativa dos "retornos em excesso" as taxas futuras poderiam sim declinar, caso a decisão tenha reduzido a incerteza sobre a política monetária no futuro.

As evidências empíricas apontam que a inclinação da estrutura a termo muitas vezes fornece sinais ambíguos na validação da Hipótese das Expectativas, configurando o que passou a ser chamado na literatura de *sign puzzle*, destacado em Hardouvelis (1994) e Jondeau e Ricart (1999). De acordo com os resultados obtidos por Shiller (1979), Mankiw (1986), o *spread* entre a taxa longa e a taxa curta não fornece uma boa previsão das mudanças de curto prazo na taxa longa. Como exemplo, se o *spread* entre a taxa de 1 ano e de 3 meses está alto, pela HE a taxa de 1 ano deveria subir ao longo dos próximos 3 meses. Campbell e Shiller (1991) também concluem 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. Este aparente paradoxo é resumido pelos autores da seguinte forma: *"The slope of the term structure almost always gives a forecast in the wrong direction for the short-term change in the yield on the longer bond, but gives a forecast in the right direction for the long-term changes in short rates"*.

Para alguns autores, como coloca Shiller (1991), este fenômeno pode ser explicado pela presença de um ruído aleatório de curto-prazo nas taxas de longo-prazo. Assim, se esse ruído reverte à média, a taxa longa tende, em um primeiro momento, a ser alta em relação à taxa curta pela existência do ruído para, subsequente, cair. Intuitivamente, esta explicação sugere que as taxas longas podem reagir em excesso às futuras taxas de juros, em detrimento das taxas correntes.

Em um artigo recente, Bekahert, Wei e Shing (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 e defendem um argumento interessante: *"One possible reading for these findings is that monetary policy succeeds in keeping short-term interest rates close to random walks and nearly unforecastable in the short run, but at longer horizons larger economic shocks do drive interest rates, and these are more importantly reflected in long-term interest rates"*.

De modo a tornar os pontos acima mais claros, seguindo o exemplo sugerido por Hardouvelis (1994), suponhamos que um anúncio do Banco Central aumente as expectativas do mercado sobre as futuras taxas de juros sem, no entanto, alterar a taxa de juros corrente. De acordo com esta explicação para a falha da HE os mercados iriam reagir em excesso ao

anúncio, aumentando suas expectativas sobre as futuras taxas de juros mais do que deveriam. A taxa longa corrente teria então uma *overreaction*, tornando o *spread* mais alto do que deveria. Ao longo do período seguinte, as taxas longas recuariam corrigindo a reação inicial, gerando assim uma correlação negativa entre a variação na taxa longa e o *spread* e contrariando a HE. Uma outra explicação às evidências contrárias à HE contempla a existência de um prêmio de risco que varia ao longo do tempo, como Tzavali e Wickens (1997).

Por outro lado, as evidências apontam que a inclinação da curva de juros sugere um sinal correto para a trajetória das taxas de curto-prazo ao longo do prazo da taxa longa em questão. Logo, quando a taxa longa sobe em relação à taxa curta, as taxas de juros de curto-prazo mais adiante tenderiam a aumentar. Assim, se o mercado faz, na média, previsões corretas, deveria haver uma correlação positiva entre a variação nas taxas curtas e o *spread* anterior. Campbell e Shiller (1987), Fama e Bliss (1987), Hardouvelis (1994), Gerlach e Smets (1997), Jondeau e Ricart (1999) e Longstaff (2000) obtiveram resultados empíricos nesta linha.

Conforme comentado anteriormente, Andrade e Tabak (2001) encontraram evidências desfavoráveis à aceitação da Hipótese das Expectativas com dados brasileiros. Utilizando a regressão de uma média ponderada das mudanças nas taxas de curto-prazo contra o *spread* e uma constante, concluem que os resultados tendem a corroborar a HE apenas para prazos mais curtos. Em seguida, incluem na regressão uma variável que seja uma *proxy* do prêmio de risco da economia brasileira e concluem que os resultados sugerem uma forte rejeição da Hipótese das Expectativas, ao indicarem que o prêmio de risco na curva de juros não seria invariante ao longo do tempo.

Vale salientar também que existe um contraste na literatura entre os resultados obtidos para os EUA e outros países da G7. Nos EUA, o *puzzle* é mais evidente, o que torna-se um fenômeno particularmente interessante porque os mercados financeiros americanos têm sido os mais líquidos no mundo do pós-guerra. Se os mercados são dominados por *traders* racionais, então seria exatamente nos mercados americanos onde menos deveria haver qualquer sinal de racionalidade limitada ou *overreaction*. Se, por outro lado, os mercados são dominados por *noise traders*², o volume mais alto de operações nos EUA implicaria em uma presença mais forte desses *traders*, o que poderia justificar em parte o *puzzle* obtido

²De acordo com Schmidt (2001), o exemplo clássico dos *noise traders* são os indivíduos que acreditam que possuem informações privilegiadas mas na verdade só possuem ruído. Na definição de Black (1986): "noise trading is trading on noise as it were information".

com os dados dos EUA.

As anomalias na estrutura a termo de juros também são estudadas por Bekahert, Hodrick e Marshall (2001), que investigam se esses desvios observados nos dados americanos podem ser atribuídos a um *peso problem*³, onde os investidores antecipam a possibilidade de uma mudança de um regime nas taxas de juros que não se confirma adiante. Os autores formalizam essa idéia estimando um modelo de mudança de regime das taxas de juros de curto-prazo em sete países, mas concluem que o modelo não é capaz de explicar totalmente as anomalias na estrutura a termo de juros do EUA.

Antes de entrar na discussão analítica do modelo, vale destacar um exemplo citado por Campbell (1995) sobre a história recente da política monetária nos EUA e que ilustra alguns dos pontos aqui discutidos. Entre 1990 e 1992, o *Federal Reserve (FED)* havia baixado repetidamente as taxas de juros de curto-prazo (*federal fund rates*) e em 1993 manteve-as baixas, até que na primavera de 1994 começou um ciclo altista que iria acumular um aumento de 150 pontos-base em 4 meses. Em seu primeiro movimento, o *FED* aumentou a taxa de curto-prazo em 25 pontos-base, enquanto as taxas de 6 meses, 5 anos e 10 anos aumentaram em 50 pontos-base, ao passo que a taxa de 30 anos aumentou na mesma magnitude da taxa curta. A resposta da estrutura a termo ao segundo movimento do *FED* foi similar. O último aumento do *FED*, por sua vez, foi seguido por um declínio nas taxas de juros em todos os prazos. O que explicaria então esse movimentos na curva de juros, no arcabouço da discussão acima? Uma das interpretações enfatiza o papel das mudanças nas expectativas dos agentes participantes do mercado de títulos. Ao longo dos anos 90, o *FED* estabeleceu uma reputação de gradualismo na condução da política monetária. Assim, o primeiro aumento nos *Fed Funds* levou o mercado a esperar movimentos adicionais de alta nos juros, o que foi refletido nas taxas de prazo mais curto. A dúvida que resta é como explicar que as taxas tenham aumentado não só para prazos curtos, como também para 5, 10 ou 30 anos. Se a atuação anti-inflacionária do *FED* fosse crível, para uma expectativa de estabilidade na taxa de juros real, seria razoável imaginar que as taxas longas nominais deveriam inclusive diminuir. Uma visão alternativa a esta afirma que o aumento nas taxas longas é explicado,

³Um *peso problem* surge quando as frequências *ex-post* de estados dentro de uma amostra diferem substancialmente de suas probabilidades *ex-ante*, de modo que esses desvios venham a distorcer a inferência econométrica. Quando um *peso problem* está presente, os momentos amostrais calculados a partir dos dados disponíveis não coincidem com os momentos populacionais que agentes racionais usariam para tomar suas decisões.

na verdade, por mudanças nos retornos requeridos em excesso dos títulos longos. Como já abordado acima, estas mudanças devem-se a aumentos no risco, através de uma maior volatilidade no mercado de títulos, e pode ser explicado por uma maior incerteza em relação à política do *FED* ou por um aumento na aversão ao risco de agentes participantes do mercado. Assim, esse relato de um breve período de condução de política monetária nos EUA reforça alguns argumentos contrários à Hipótese das Expectativas.

3 O Modelo

3.1 Os Modelos de Valor Presente - MVP e Hipótese das Expectativas - HE

Um modelo de valor presente geral é apresentado em Campbell e Shiller (1987) para duas variáveis, y_t e Y_t , mostrando que a variável Y_t é uma função linear do valor presente descontado dos valores futuros esperados de y_t :

$$Y_t = \theta(1 - \delta) \sum_{i=0}^{\infty} \delta^i E_t y_{t+i} + c, \quad (2)$$

onde c é uma constante, θ é o coeficiente de proporcionalidade e δ é o fator de desconto. A aplicação para a estrutura a termo é um caso específico da equação (2), onde R_t é a taxa de juros de longo-prazo, r_t é a taxa de juros de curto-prazo e θ é igual a 1, conforme vemos abaixo:

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

Note que a equação (3) acima pode ser facilmente associada à equação (1), visto que ambas descrevem o processo de formação da taxa de juros no longo-prazo, associando esta a uma média das taxas de juros de curto-prazo. A equação (3) atribui, todavia, uma ponderação desigual às taxas de curto-prazo esperadas para o futuro.

A equação (3) acima também pode ser descrita como um caso particular do modelo linearizado descrito por Shiller (1979), para $n = \infty$:

$$R_t^{(n)} = \frac{1 - \delta}{1 - \delta^n} \sum_{i=0}^{n-1} \delta^i E_t r_{t+i} + c_n. \quad (4)$$

A equação (4), na verdade, foi construída para que a soma dos pesos seja igual a um, de modo que as taxas de juros de curto-prazo esperadas para um futuro próximo tenham 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 (4) como o valor presente das futuras taxas de juros de curto-prazo descontadas por r .

Tomando como exemplo $n = 3$, facilmente percebe-se a intuição associada à equação (4) acima:

$$R_t^{(3)} = \frac{1-\delta}{1-\delta^3}(E_t r_t + \delta E_t r_{t+1} + \delta^2 E_t r_{t+2}) + c_3. \quad (5)$$

Note que as taxas de curto-prazo esperada mais adiante tem um peso menor da determinação da taxa longa e que a soma das ponderações é igual à unidade, conforme podemos observar abaixo:

$$\frac{1-\delta}{1-\delta^3}(1 + \delta + \delta^2) = 1. \quad (6)$$

Como apontado por Campbell e Shiller (1987), há alguns problemas para se testar a equação (2). 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 (2) 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 bastante 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 Autorregressão Vetorial (VAR), onde este modelo é usado para fazer a previsão ótima descontada implícita em (2). 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 (3) 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; ver Campbell e Shiller (1987). Assim, a equação (3) 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.

No entanto, ao aplicar uma transformação na equação (3), pode-se obter um relação entre séries estacionárias. Em primeiro lugar, definamos o *spread* entre a taxa longa e a taxa curta como:

$$S_t = R_t - r_t. \quad (7)$$

Agora subtrai-se r_t de ambos os lados da equação(3), de modo que temos:

$$S_t = -r_t + (1 - \delta) \sum_{i=0}^{\infty} \delta^i E_t r_{t+i} + c. \quad (8)$$

Desmembrando os termos da equação acima, temos que:

$$S_t = E(-r_t + r_t - \delta r_t + \delta r_{t+1} - \delta^2 r_{t+1} + \delta^2 r_{t+2} - \delta^3 r_{t+2} + \delta^3 r_{t+3} - \dots) + c. \quad (9)$$

Rearranjando os termos, temos:

$$S_t = E [\delta(r_{t+1} - r_t) + \delta^2(r_{t+2} - r_{t+1}) + \delta^3(r_{t+3} - r_{t+2}) + \dots] + c. \quad (10)$$

Podemos então reescrever a equação(3) da seguinte forma:

$$S_t = E_t S_t^* + c, \quad (11)$$

onde

$$S_t^* = \sum_{i=1}^{\infty} \delta^i \Delta r_{t+i}. \quad (12)$$

Também podemos colocar S_t como função das variações em R_t . Para isso devemos fazer algumas manipulações algébricas. De (3), temos:

$$R_{t+1} - R_t = E \{ (1 - \delta) [\delta(r_{t+1} - r_t) + \delta^2(r_{t+2} - r_{t+1}) + \delta^3(r_{t+3} - r_{t+2}) + \dots] \} + c. \quad (13)$$

Logo, temos que:

$$E_t \Delta R_{t+1} = (1 - \delta) \sum_{i=1}^{\infty} \delta^i \Delta r_{t+i} + c, \quad (14)$$

e

$$S_t = \frac{\delta}{1 - \delta} E_t \Delta R_{t+i} + c. \quad (15)$$

A equação (12) refere-se ao *spread* como uma constante mais a previsão ótima de S_t^* , uma média ponderada de mudanças futuras em r_t . De outra forma, o *spread* entre a taxa curta e a taxa longa é dado por uma constante (*term premium*) mais uma média ponderada de mudanças futuras nas taxas de curto-prazo. A equação (15), por sua vez, diz que o *spread* é linear em relação à previsão ótima da variação em R .

As equações (12) e (15) 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 (11) e (12) que S_t também o será. Por (15), segue-se também que nessas circunstâncias 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 bivariada 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 das equações (11) e (12).

4 Metodologia

4.1 Técnica Econométrica

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.

Definition 1 (Engle e Granger, 1987) *Um vetor x_t é dito ser cointegrado de ordem d , b , denotado por $x_t \sim CI(d, b)$, se: i) todas as componentes de x_t são integradas de ordem d (estacionárias na d -ésima diferença), e, ii) existe um vetor $\alpha \neq 0$, tal que $W_t = \alpha'x_t$ é integrado de ordem $d - b$, $b > 0$. O vetor α é chamado de vetor de cointegração.*

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 autorregressivo 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 elas, usando a técnica de Johansen (1988, 1991). Se confirmada a cointegração entre essas variáveis, será usado um modelo VAR para modelar as séries em (3).

Discutiremos, a seguir, o procedimento proposto por Campbell e Shiller(1987), utilizando as variáveis acima definidas $S_t = R_t - r_t$ e Δr_t , como variáveis estacionárias que resumem a dinâmica das séries R_t e r_t . Considere a seguinte representação VAR para S_t e Δr_t :

$$\begin{bmatrix} \Delta r_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} \Delta r_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}, \quad (16)$$

onde os polinômios $a(L)$, $b(L)$, $c(L)$, $d(L)$ são todos de ordem p no operador de defasagens L . 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:

$$\begin{bmatrix} \Delta r_t \\ \cdot \\ \cdot \\ \Delta r_{t-p+1} \\ S_t \\ \cdot \\ \cdot \\ S_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \dots & a_p & b_1 & \dots & b_p \\ 1 & & & & & \\ & \cdot & & & 0 & \\ & & 1 & & & \\ c_1 & \dots & c_p & d_1 & \dots & d_p \\ & & & 1 & & \\ & 0 & & & \cdot & \\ & & & & & 1 \end{bmatrix} \times \begin{bmatrix} \Delta r_{t-1} \\ \cdot \\ \cdot \\ \Delta r_{t-p} \\ S_{t-1} \\ \cdot \\ \cdot \\ S_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ 0 \\ \cdot \\ 0 \\ \varepsilon_{2t} \\ 0 \\ \cdot \\ 0 \end{bmatrix} \quad (17)$$

ou, de forma compacta:

$$z_t = Az_{t-1} + \epsilon_t, \quad (18)$$

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

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

$$E[(z_{t+i}/H_t)] = A^i(z_t), \quad (19)$$

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 (mais fraca) do modelo 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. Em outras palavras, S_t incorpora toda a informação do mercado sobre o vetor de variáveis de estado z_t . 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 (17). Para derivar tais restrições, toma-se a esperança da equação (12) 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:

$$S_t = E \left[\sum_{i=1}^{\infty} \delta^i [\Delta r_{t+i}] \middle| H_t \right] \equiv S_t^*. \quad (20)$$

A equação (20) nos diz que S_t deve ser igual à previsão ótima irrestrita das variações futuras de Δr_{t+i} , denotada por S_t^* .

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 (20) como:

$$S_t^* = \sum_{i=1}^{\infty} \delta^i h' [A^i(Z_t)]. \quad (21)$$

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

$$S_t^* = h' \delta A (I - \delta A)^{-1} (Z_t). \quad (22)$$

Note que a equação acima nos dá a série do *spread* ótimo. Para comparar formalmente os valores estimados com os valores observados do *spread* em relação à sua média, definimos $S_t^* = g'Z_t$, de modo tal que podemos reescrever a equação acima como:

$$g'Z_t = h' \delta A (I - \delta A)^{-1} (Z_t), \quad (23)$$

ou

$$g' = h' \delta A (I - \delta A)^{-1}. \quad (24)$$

Pós-multiplicando a equação(24) acima por $(I - \delta A)^{-1}$, chegamos a :

$$g'(I - \delta A) = h' \delta A. \quad (25)$$

A equação (25) define um conjunto de $2p$ restrições nos coeficientes do sistema VAR. Da estrutura da matriz A , as restrições impostas por (25) nos coeficientes individuais, a serem testadas por meio de um teste de Wald, são:

$$c_i = -a_i, i = 1, \dots, p \quad (26)$$

$$d_1 = (1/\delta) - b_1, i = 1 \quad (27)$$

$$d_i = -b_i, i = 2, \dots, p. \quad (28)$$

Nas seções seguintes, vamos constatar que os modelos estimados para os vetores autoregressivos estimados, tanto para a taxa de 180 dias como de 360 dias, rejeitam fortemente a hipótese de homocedasticidade dos erros. Assim, conforme abordado anteriormente, fez-se necessário a construção de estimadores robustos à heterocedasticidade, de modo a permitir a realização de testes mais acurados sobre a validade do modelo.

Em primeiro lugar, sabemos que a fórmula usual da matriz de variância-covariância dos estimadores dos coeficientes no VAR qualquer é dada por:

$$\Gamma = \Sigma(X'X)^{-1}. \quad (29)$$

Quando há heterocedasticidade, esse estimador não é mais consistente e, seguindo a sugestão de White (1984), deve ser substituído por:

$$\Gamma = (X'X)^{-1}X'VX(X'X)^{-1}, \quad (30)$$

onde define-se V como uma matriz diagonal com resíduos quadrados na diagonal.

Em seguida, devemos explicar a construção geral de um teste de Wald a partir de um conjunto de J restrições lineares da forma:

$$H_0 : RB = q. \quad (31)$$

Cada linha de R é uma restrição linear única sobre o vetor de coeficientes B , onde R é uma matriz $J \times N$, J é o número de restrições lineares e N é o número de coeficientes estimados, B é uma matriz $N \times 1$ e q é uma matriz $J \times 1$. Pode-se então definir um vetor m , onde $m = RB - q$, de modo tal que a essência do teste de Wald é medir se o desvio de m em relação a um vetor de zeros é atribuídos a erros amostrais ou se é estatisticamente significativo. Assim, pode-se definir um teste de hipóteses em Wald como:

$$W = m'\{Var[m]\}^{-1}m. \quad (32)$$

Assim, a estatística W tem uma distribuição qui-quadrado com J graus de liberdade. Intuitivamente, quanto maior for m maior a falha dos estimadores em satisfazer as restrições e, assim, maior será a estatística calculada qui-quadrado, o que levaria a rejeição da hipótese de que as restrições lineares são válidas. Implementado o método de White de correção de heterocedasticidade, descrito acima, obteve-se uma nova matriz de covariância dos coeficientes, de modo que definindo essa matriz como VB , temos que:

$$Var[m] = R[VB]R'. \quad (33)$$

Assim, a estatística de Wald a ser construída para testar formalmente a validade do modelo é dada por:

$$W = [RB - q]' \{ [R(VB)R'] \}^{-1} [RB - q], \quad (34)$$

onde R , V e q são definidas como em (31).

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 estatisticamente equivalentes, também podemos testar a validade das equações desenvolvidas na seção anterior por meio do teste de razão de variância. Constrói-se uma estatística para testar a seguinte restrição:

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

onde S_t^* é a previsão ótima do *spread*. Sob o modelo de valor presente, a razão $Var(S_t^*)/Var(S_t)$ deve ser igual a um, 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 a razão das variâncias e a correlação forem iguais a um, então $S_t = S_t^*$, como prevê o modelo.

As implicações do modelo estudado a serem examinadas nas próximas seções e os principais testes aqui feitos podem ser resumidos nos pontos abaixo:

1. 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.
2. 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 ?
3. 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 checar a validade das restrições impostas aos coeficientes do VAR pela equação (25).

4. Comparação da volatilidade das séries S_t e S_t^* e a estimação da correlação entre as séries.

5 Base de Dados

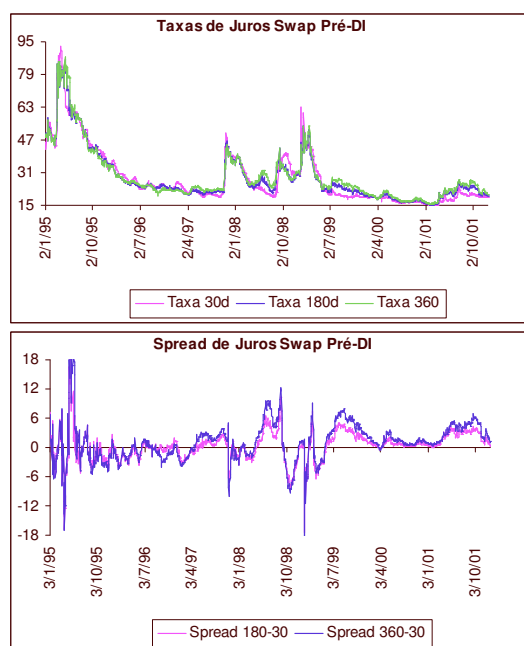
A construção das séries r_t e S_t , definidas anteriormente como a taxa de juros de curto-prazo e o *spread* entre a taxa de juros de longo-prazo e de curto-prazo, foi feita a partir das séries das taxas de swap pré-DI obtidas junto a BM&F, com frequência mensal (taxa do último dia útil do mês) e cobrindo 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 trabalhamos com duas séries - a taxa de swap pré-DI de 6 meses e 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.

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 indí-

cios 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.



6 Resultados Empíricos

6.1 MVP com Hipótese das Expectativas

6.1.1 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 I, 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-Perron.

Tabela I
Teste de Raiz Unitária

Variável	Defasagens	Teste ADF	Teste Philips-Perron
r_t	1	-2,19	-2,88
Δr_t	2	-10,43**	-13,50**
$S_{t(180)}$	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
$\Delta R_{t(180)}$	2	-10,73**	-13,51**
$R_{t(360)}$	1	-2,71	-2,40
$\Delta R_{t(360)}$	3	-5,75**	-7,95**

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

2) Δ corresponde a 1a. Diferença

3) 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. Por outro lado, os testes acima rejeitam a hipótese nula de duas raízes unitárias, sugerindo que as taxas de juros, em seus diferentes prazos, são integradas de ordem 1 - $I(1)$, ou seja, é necessário diferenciá-las uma vez para que se tornem séries estacionárias. 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 e, assim, as séries de *spread* em nível não requereriam nenhuma diferenciação para tornarem-se estacionárias. A verificação da estacionaridade das séries de *spread* também indica que as taxas curtas e longas cointegram, dado que o *spread* nada mais é do que uma combinação linear da taxa curta com a taxa longa, conforme definido na equação (7).

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 $R_{(i)}$, $i = 180, 360$ e r_t , utilizou-se os critérios de informação de Hannan-Quinn, Schwarz e Akaike. Tanto para o sistema con-

struído com $R_{t(180)}$ e r_t como também para $R_{t(360)}$ e r_t , houve divergência quanto à escolha da defasagem ótima. Levando-se em conta também os testes de diagnósticos, optou-se pela escolha de 1 defasagem. Nas tabelas abaixo, são mostrados a escolha ótima do número de defasagens de acordo com os diferentes métodos, bem como os testes de diagnóstico.

Tabela II

Ordem ótima do VAR

Modelo VAR Contendo	Critério de Informação		
	SC	HQ	AIC
$R_{T(180)}$	1	3	5
$R_{t(360)}$	1	3	5

Note que 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.

Tabela III

Testes de Diagnóstico nos Resíduos do VAR

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

Modelo Contendo	Heterocedasticidade	
	Qui-q	p-value
$R_{t(180)}$	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 IV
Teste de Johansen

H ₀ : posto= p	Est. λ_{\max}	95%	Est. do Traço	95%	defasagens
$p = 0$	27,62	15,67	33,97	19,96	1
$p \leq 1$	6,34	9,24	6,34	9,24	1

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 não pode ser rejeitada tanto ao nível de significância de 1% como de 5%. O vetor de cointegração sugerido é $\begin{pmatrix} 1 & -0,937 \end{pmatrix}$. Na tabela abaixo, testamos se esse vetor é estatisticamente igual a $\begin{pmatrix} 1 & -1, \end{pmatrix}$, pelo teste de cointegração restrita. Note que não é rejeitada a hipótese nula de que o vetor de cointegração é estatisticamente igual a $\begin{pmatrix} 1 & -1, \end{pmatrix}$.

Tabela V
Teste de Cointegração Restrita

Restrição	Resultado
$R_t = 1; r_t = -1, \quad p - value = 0.127$	

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

Tabela VI

Teste de Johansen

$H_0: \text{posto}=p$	Est. λ_{\max}	95%	Est. do Traço	95%	defasagens
$p = 0$	21,58	15,67	27,63	19,96	1
$p \leq 1$	6,05	9,24	6,05	9,24	1

Coefficientes 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 $\begin{pmatrix} 1 & -0,902 \end{pmatrix}$. Na tabela abaixo, realizamos mais uma vez o teste de cointegração restrita, de modo a checar se esse vetor é estatisticamente igual a $\begin{pmatrix} 1 & -1, \end{pmatrix}$. Note que é bastante difícil rejeitar a hipótese nula de que o vetor de cointegração é estatisticamente igual a $\begin{pmatrix} 1 & -1, \end{pmatrix}$.

Tabela VII

Teste de Cointegração Restrita

Restrição	Resultado
$R_t = 1; r_t = -1,$	$p - \text{value} = 0.123$

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 $\begin{pmatrix} 1 & -1, \end{pmatrix}$. 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.

6.1.2 Estimação do vetor autorregressivo e Teste de Causalidade

Seguindo 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 . Novamente a escolha da ordem do VAR é feita a partir dos critérios de informação de

Hannan-Quinn, Schwarz e Akaike. ⁴ Tanto para o sistema construído com $S_{t(180)}$ e Δr_t como também para $S_{t(360)}$ e Δr_t , há divergência na escolha das defasagens pelos diferentes métodos:

Tabela VIII
Ordem ótima do VAR

Modelo VAR Contendo	Métodos		
	SC	HQ	AIC
$S_{t(180)}$	2	2	6
$S_{t(360)}$	2	2	6

Posteriormente, como ilustrado na tabela VII, realizaram-se testes de diagnóstico para verificar se o sistema está bem especificado e confirmar a ordem escolhida para a construção do VAR. Pela combinação dos critérios acima com os testes de diagnóstico, optamos pela escolha de 2 defasagens. Nota-se que há evidência de que os modelos não detectaram a presença de autocorrelação serial nos erros do VAR, embora claramente haja heterocedasticidade nos erros do VAR, que posteriormente será corrigida utilizando o método de White descrito anteriormente.

Tabela IX

Testes de Diagnóstico nos Resíduos do VAR

Modelo Contendo	Autocorrelação Serial	
	$p=8$	
	LM	p-value
$S_{t(180)}$	4,55	0,335
$S_{t(360)}$	4,51	0,341
Modelo Contendo	Heterocedasticidade	
	Qui-q	p-value
	$S_{t(180)}$	10,32
$S_{t(360)}$	96,15	0,00

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

⁴Como foi dito anteriormente, o VAR foi construído com as séries extraídas de suas médias não fazendo-se necessária a inclusão de um vetor de constantes no sistema.

Verificamos abaixo a comprovação da causalidade bilateral de Granger, conforme apontado na tabela abaixo, de modo a comprovar se o *spread* da taxa longa e da taxa curta Granger-causa as variações nas taxas de juros, conforme previsto pela Hipótese das Expectativas e já discutido anteriormente:

Tabela X
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. Deste modo, existe forte evidência de que o *spread* entre a taxa longa e a taxa curta causa no sentido de Granger as variações nas taxas curtas. Entretanto, o R^2 da equação para S_{it} também é razoável (29% e 34% para ambos os modelos, respectivamente), apontando uma possível causalidade bidirecional entre as variáveis S_{it} e Δr_t , o que é confirmado pelos resultados do teste. Em ambos os casos, rejeita-se a hipótese de que S_{it} não Granger causa Δr_t e vice-versa aos níveis de significância usuais. Note que para ambas as combinações de taxa curta e taxa longa, não podemos rejeitar a hipótese de que tanto o *spread* Granger-causa as variações na taxa curta, como também que as variações na taxa curta precedem temporalmente o *spread* entre a taxa curta e a taxa longa. Vale ressaltar que pela Hipótese das Expectativas não seria esperada uma causalidade bidirecional. Assim, os resultados dos testes de causalidade corroboram apenas parcialmente a HE.

6.1.3 Teste do MVP e Construção do *spread* ótimo

Dada a estimação do vetor autoregressivo e da verificação dos testes de causalidade, 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 testar a restrição imposta nos coeficientes do VAR, dada pela equação (25). Em seguida, conforme apresentado anteriormente, realizaremos a comparação da variância das séries S_t e S_t^* e a estimação do seu coeficiente de correlação.

Dadas as restrições obtidas a partir do modelo em (25), foram calculadas as estatísticas de Wald, descritas em (34) a partir de diversas taxas de desconto (de 0% a 10%). Utilizando a sugestão de Campbell e Shiller (1987)⁵, 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.

Para o complemento da análise sobre o MVP para a estrutura a termo de juros no Brasil, além do teste de Wald para as restrições, devemos construir a série do *spread* ótimo conforme definido em (22) para daí, conforme comentado anteriormente, realizarmos a comparação da volatilidade das séries S_t e S_t^* e a estimação do seu coeficiente de correlação.

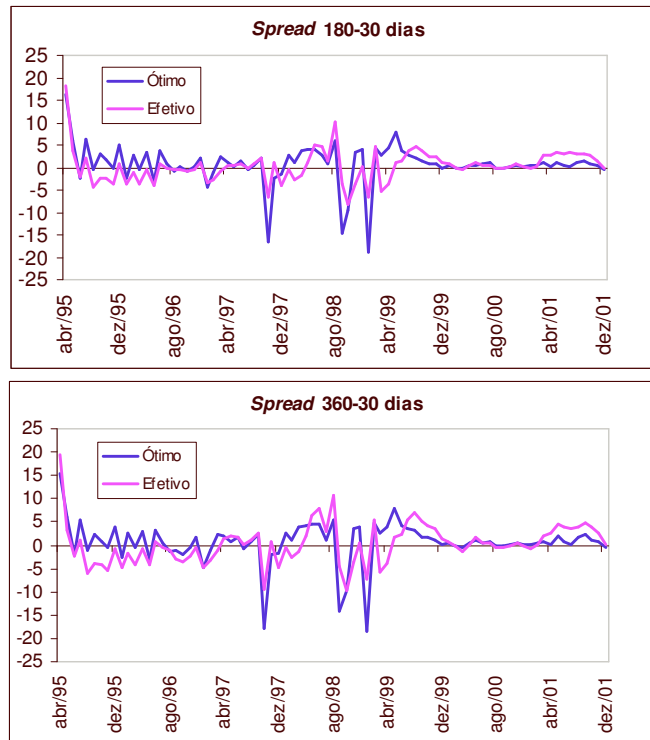
A partir da equação (22) e supondo que a taxa de desconto é igual a 2,5%, por razões aqui já explicitadas, chegamos às seguintes equações, para os dois modelos em discussão:

$$S_{t(180)}^* = -0,1596 * \Delta r_t - 0,307 * \Delta r_{t-1} + 0,7378 * S_{t(180)} - 0,3623 * S_{t-1(180)}, \quad (35)$$

$$S_{t(360)}^* = -0,1574 * \Delta r_t - 0,294 * \Delta r_{t-1} + 0,6826 * S_{t(360)} - 0,2968 * S_{t-1(360)}. \quad (36)$$

Antes de explicitarmos os resultados obtidos, fazemos uma análise menos formal e mais intuitiva, 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.

⁵O 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.



Analisando o gráfico acima, percebe-se que as séries guardam uma correlação positiva, bem como suas variâncias não parecem divergir consideravelmente, conforme sugerido pela teoria. Fazendo uma análise mais formal, ao se testar a restrição imposta nos coeficientes do VAR, dada pela equação (25), os *p-values* para ambos os modelos sugerem que o modelo não pode ser rejeitado para os dados brasileiros, a um nível de significância de 5%, o que não é uma evidência das mais robustas sobre a real validade da Hipótese das Expectativas para a curva de juros.

Tabela XI

Testes do Modelo de Valor Presente

Teste de Wald para restrições conjuntas (eq. (25))		
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_{it} e da prevista pelo modelo S_{it}^* , também reportados na tabela acima, 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-value* 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 Expectativas para a estrutura a termo de juros no Brasil, principalmente quando trabalhamos com a taxa de juros de de 1 mês e de 12 meses como aproximações, respectivamente, da taxa de juros de curto-prazo e de longo-prazo da economia brasileira.

6.1.4 Comparação com Resultados Internacionais

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, inclusive no arcabouço de modelos de otimização intertemporal (ver Brito(2001)), a evidência empírica contra a HE está longe de ser conclusiva.

Em um de seus artigos clássicos, diversas vezes citado neste trabalho, 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, que nada mais é que a previsão irrestrita do valor presente das mudanças futuras nas taxas de juros, tem uma correlação positiva bastante alta com o *spread* observado, o que poderia ser interpretado como uma evidência que os desvios em relação ao modelo são transitórios. Os testes de razão de variância também não conseguem rejeitar a hipótese de que a variância das duas séries são estatisticamente iguais. Campbell e Shiller (1987) também encontraram outras evidências que corroboram os nossos resultados - 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. Assim como o nosso trabalho, Campbell e Shiller (1987) também chegam a uma forte evidência que os *spreads* Granger-causam as variações nas taxas de curto-prazo.

Nos demais artigos que, assim como este trabalho, 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. Através da estimação de um modelo de segmentação do mercado, onde a política de gestão da dívida pública influencia a estrutura a termo, o autor conclui que a política de *buyback* de dívida da Inglaterra na segunda metade da década de 80 foi em grande parte responsável pela inversão da curva de juros inglesa ao longo do período amostral de seu trabalho, sobrepondo-se a qualquer efeito que as políticas governamentais pudessem vir a ter sobre as taxas de juros esperadas para o futuro.

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

7 Conclusão

Depois de mais um século, desde Fischer (1896), a Hipótese das Expectativas continua sendo a teoria mais conhecida e mais intuitiva sobre as flutuações na estrutura a termo de juros e, conforme já discutido ao longo deste trabalho, 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 autoregressivo 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 verificadas 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 Expectativas 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 Andrade e Tabak (2001), examinar o papel do gerenciamento da dívida pública pode ser um dos caminhos a ser trilhado.

Referências

- [1] Anchite, Claudine F. e Issler, J.V. (2001). “Racionalidade e Previsibilidade no Mercado Brasileiro de Ações: Uma Aplicação de Modelos de Valor Presente.” *Ensaio Econômicos* 415, EPGE, FGV.
- [2] Andrade, S.C. e Tabak, B.M. (2001). “Testing the Expectation Hypothesis in the Brazilian Term Structure of Interest Rates.” *Working Paper Series, no.18*, Banco Central do Brasil.
- [3] Bekaert, G. e Hodrick, R. (2001). “Expectation Hypotheses Tests.” *Journal of Finance, vol 56, no.4*, pp. 1357-1371.
- [4] Bekaert, G., Hodrick, R. e Marshall, D. (2001). “Peso problem explanation for term structure anomalies.” *Journal of Monetary Economics, vol 48, no.4*, pp. 241-270.
- [5] Bekaert, G., Wei, M. e Xing, Y. (2002). “Uncovered Interest Rate Parity and the Term Structure.” *NBER Working Papers, no.8795*.
- [6] Black, F. (1986). “Noise”. *Journal of Finance, vol 41*, pp. 529-543.

- [7] Brito, R. (2001). "Essays on the Monetary Aspects of the Term Structure of Nominal Interest Rates". *Tese de Doutorado em Economia- EPGE-FGV*.
- [8] Campbell, J. (1995) "Some Lessons from the Yield Curve." *Journal of Economic Perspectives*, vol 9, no.3, pp. 129-152.
- [9] Campbell, J. (2000) "Asset Pricing at the Millenium." *Journal of Finance*, vol 50, no.4, pp. 1515-1540.
- [10] Campbell, J., Lo, W. e MacKinlay, A.C. (1997) "*The Econometrics of Financial Markets*." Princeton University Press.
- [11] Campbell, J. e Shiller, R. (1987). "Cointegration and Tests of Present Value Models." *Journal of Political Economy*, vol 95, no.5, pp. 1062-1088.
- [12] Campbell, J. e Shiller, R. (1991). "Yield Spread and Interest Rate Movements: A Bird's Eye View." *Review of Economic Studies*, 58, pp. 495-514.
- [13] Cuthbertson, K. (1996). "The Expectations Hypothesis of the Term Structure: The UK Interbank Market." *The Economic Journal*, 106, pp. 578-592.
- [14] Dickey D. e Fuller, W. (1979) "Distribution of the Estimates for Autoregressive Time Series with a Unit Root." *Journal of the American Statistical Association*, 74, pp. 427-31.
- [15] Engle, Robert F. and Clive W.J. Granger (1987). "Cointegration and Error-Correction: Representation, Estimation and Testing." *Econometrica*, 55, pp. 251-276.
- [16] Evans, M. e Lewis, K. (1994). "Do stationary risk premia explain it all?" *Journal of Monetary Economics*, 33, pp. 285-318.
- [17] Fabossi, F. (1995). "*Bond Markets, Analysis and Strategies*," Prentice Hall, Inc., Upper Saddle River, NJ.
- [18] Fama, E. (1970). "Efficiency Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance*, 25, pp. 383-417.
- [19] Fama, E. e Bliss, R. (1987). "The Information in Long-Maturity Forward Rates." *American Economic Review*, 77, pp. 680-692.

- [20] Fischer, I. (1896). "Appreciation and interest." *Publications of the American Economic Association*, 11, pp. 21-29.
- [21] Froot, K. (1989). "New Hope for the Expectation Hypothesis of the Term Structure of Interest Rates." *The Journal of Finance*, vol 44, no.2, pp. 283-305.
- [22] Gerlach, S. e Smets, F. (1997). "The term structure of Euro-rates: some evidence in support of the expectations hypothesis." vol 16, no.2, pp. 305-321.
- [23] Greene, W. (2000). "*Econometric Analysis*," Prentice Hall, Inc., Upper Saddle River, NJ.
- [24] Hamilton, J. (1994). "*Time Series Analysis*," Princeton University Press, Princeton, NJ.
- [25] Hardouvelis, G. (1994). "The term structure spread and future changes in long and short rates in the G7 countries." *Journal of Monetary Economics*, 33, pp. 255-283.
- [26] Johansen, S.(1988). "Statistical Analysis of Cointegrating Vectors." *Journal of Economic Dynamics and Control*, vol. 12, pp. 231-254.
- [27] Johansen, S.(1991). "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica*, vol. 59, pp. 1551-1580.
- [28] Jondeau, E. 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, pp. 725-750.
- [29] Longstaff, Francis (2000). "The term structure of very short-term rates: New evidence for the expectation hypothesis". *Journal of Financial Economics*, 58, pp. 397-415.
- [30] Mankiw, N.G. e Miron, J. (1986). "The Changing Behaviour of the Term Structure of Interest Rates." *Quarterly Journal of Economics*, 101, pp. 211-228.
- [31] McCallum, B. (1994). "Monetary Policy and the Term Structure of Interest Rates". *NBER Working Papers*, no. 4938.
- [32] Perron, P.(1995). "Lecture Notes," Mimeo: Université de Montréal.

- [33] Phillips, P. e Perron, P. (1988). "Testing for a Unit Root in Time Series Regression." *Biometrika* 75, pp. 335-46.
- [34] Schmidt, B. (2001). "Expectativas Heterogêneas, Análises Técnicas e os Preços de Equilíbrio". *Dissertação de Mestrado em Economia- EPGE-FGV*.
- [35] Senna, F. e Issler, J.V. (2000). "Mobilidade de Capitais e Movimentos da Conta-Corrente do Brasil: 1947-1997". *Ensaio Econômicos* 379, EPGE, FGV.
- [36] Shiller, R., (1979). "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure", *Journal of Political Economy*, 87, pp. 1190-1219.
- [37] Shiller, R., (1990). "The Term Structure of Interest Rates." Em Friedman, B. e Hahn, F., eds., *Handbook of Monetary Economics, vol.1*. Amsterdam, North-Holland, 1990, pp. 627-772.
- [38] Shiller, R., (1991). "*Market Volatility*", The Mit Press, Cambridge, MA.
- [39] Taylor, M. (1992). "Modelling the yield curve.", *Economic Journal*, vol 102, no. 412, pp.524-537.
- [40] Tzavalis, E. and Wickens, M. (1997). "Explaining the Failures of the Term Spread Models of the Rational Expectation Hypothesis of the Term Structure." *Journal of Money, Credit and Banking*, 29, no.3, pp. 364-380.
- [41] White, Halbert (1984). "*Asymptotic Theory for Econometricians*", Academic Press, New York, NY.