

3

Modelo Afim da Estrutura a Termo da Taxa de Juros

Tendo encontrado um relevante conteúdo informacional no *spread* da curva de juros, passaremos agora a uma análise mais estrutural do problema. A primeira questão que surge é qual modelo de estrutura a termo da taxa de juros deve ser utilizado. O modelo selecionado foi um modelo afim da estrutura a termo da taxa de juros, introduzido por Duffie e Kan (1996). Praticamente todas as aplicações de modelos multifatoriais da curva de juros tem focado em casos especiais dessa família, já que esta "*acomoda variáveis de estado com médias e covariâncias variantes no tempo através de uma especificação afim dos coeficientes neutros ao risco da tendência e volatilidade*" (Dai e Singleton (2000)), o que permite que tenhamos um prêmio de risco variante no tempo com especificações relativamente simples da dinâmica do sistema. Esta característica é muito importante, já que existem diversas evidências de que a Hipótese das Expectativas não é válida para o Brasil, sendo portanto essencial que o modelo permita a especificação de um prêmio de risco variante no tempo. Além disso, essa classe de modelos nos fornece soluções fechadas, o que facilita a sua aplicação econométrica.

Após a estimação de um modelo afim robusto para a estrutura a termo da taxa de juros brasileira, podemos caminhar para o objetivo central desta dissertação, que é a interpretação econômica da dinâmica da curva de juros. Para estudar o caso brasileiro, torna-se necessário expandir o modelo utilizado para a economia americana incluindo variáveis referentes ao setor externo, especialmente se levarmos em consideração a grande quantidade de choques externos sofridos pela economia no período recente. Assim, tentou-se selecionar as variáveis da maneira mais parcimoniosa possível, de modo que tivéssemos um modelo que nos fornecesse uma boa descrição da dinâmica macroeconômica ao mesmo tempo que fosse tratável, permitindo uma estimação confiável.

Nas próximas subseções serão introduzidas as hipóteses principais do modelo e suas consequências para a precificação dos títulos¹. Avaliaremos também a adequação deste modelo ao caso brasileiro utilizando somente variáveis latentes. Finalmente, apresentaremos a seleção das variáveis macroeconômicas e

¹A derivação do modelo baseia-se na estrutura utilizada em Ang e Piazzesi (2003).

sua inserção no modelo afim para em seguida discutirmos os resultados e obtermos dessa forma a interpretação macroeconômica da dinâmica da estrutura a termo no Brasil.

3.1

Estrutura Geral do Modelo

3.1.1

Hipóteses sobre a dinâmica das variáveis

No modelo afim, o vetor de variáveis de estado² F_t de dimensão $n \times 1$ segue um processo gaussiano VAR(1):

$$F_t = \Phi_0 + \Phi F_{t-1} + \Sigma \varepsilon_t \quad (3-1)$$

onde $\varepsilon_t \sim N(0, I_n)$, Σ é diagonal e $\Phi_{n \times n}$ é uma matriz triangular inferior.

A taxa de juros de curto-prazo é definida como uma função linear das variáveis latentes:

$$i_t = \delta_0 + \delta_1' F_t \quad (3-2)$$

Finalmente, os preços de risco associados aos choques ε_t são definidos como uma função linear do vetor de estados:

$$\lambda_t = \lambda_0 + \lambda_1 F_t \quad (3-3)$$

A equação de transição de estados 3-1, a equação da taxa de curto-prazo 3-2 e os preços de risco 3-3 formam um modelo de estrutura a termo gaussiano "essencialmente afim" em tempo discreto com imposição de ausência de arbitragem, ou um modelo $A_0(N)$ como definido em Dai e Singleton (2000). Este modelo contém uma volatilidade constante dos fatores da curva de juros mas a precificação do risco é estado-dependente, o que implica em heterocedasticidade condicional nos prêmios de risco. Segundo a comparação de Dai e Singleton (2002), este tipo de especificação tem melhor performance na estimação de variações das taxas longas sobre o slope da curva normalizado em toda a amostra.

²Variáveis de estado são aquelas que explicam a dinâmica das diferentes taxas de juros

3.1.2

Taxa de desconto estocástica

A equação básica para precificação de ativos é derivada da decisão dos agentes de quanto consumir, quanto poupar e qual portfolio de ativos possuir a cada instante do tempo. Dessa forma, podemos derivá-la à partir da condição de primeira ordem do consumidor, obtendo:

$$P_t = E_t \left[\beta \frac{u'(c_{t+1})}{u'(c_t)} P_{t+1} \right] \quad (3-4)$$

onde P_t é o preço do ativo em t , $u'(c_t)$ é a utilidade marginal de se consumir uma unidade adicional do bem de consumo e β é o fator de desconto que representa o grau de impaciência do consumidor. Esta equação sumariza a condição de otimalidade: $P_t u'(c_t)$ é a perda de utilidade do investidor ao poupar investindo uma unidade no ativo financeiro enquanto $E_t [\beta P_{t+1} u'(c_{t+1})]$ é o aumento esperado da utilidade decorrente da venda do ativo em $t + 1$.

Podemos definir a taxa de desconto estocástica como:

$$m_{t+1} \equiv \beta \frac{u'(c_{t+1})}{u'(c_t)}$$

Por outro lado, o retorno bruto do ativo é definido como:

$$R_{t+1} \equiv \frac{P_{t+1}}{P_t}$$

Assim podemos reescrever a equação básica de precificação da seguinte forma:

$$E_t [m_{t+1} R_{t+1}] = 1 \quad (3-5)$$

Se temos ausência de arbitragem, existe uma medida martingal equivalente (ou medida neutra ao risco) Q de modo que o preço de qualquer ativo P_t que não paga dividendos em $t+1$ satisfaz:

$$P_t = E_t^Q [\exp(-i_t) P_{t+1}] \quad (3-6)$$

Além disso, pode-se afirmar que existe a derivada de Radon-Nikodyn³, que converte a medida neutra ao risco à medida geradora dos dados de modo que:

$$E_t^Q(P_{t+1}) = E_t \left(\frac{\xi_{t+1}}{\xi_t} P_{t+1} \right) \quad (3-7)$$

Unindo as equações 3-6 e 3-7 temos então:

$$P_t = E_t \left[\frac{\xi_{t+1}}{\xi_t} \exp(-r_t) P_{t+1} \right] \quad (3-8)$$

e unindo as equações 3-4 e 3-8 obtemos finalmente que:

$$m_{t+1} = \exp(-r_t) \frac{\xi_{t+1}}{\xi_t} \quad (3-9)$$

A derivada de Radon-Nikodyn é um martingal e, sob as hipóteses iniciais, segue um processo log-normal da forma:

$$\xi_{t+1} = \xi_t \exp \left(-\frac{1}{2} \lambda_t \lambda_t' - \lambda_t' \varepsilon_{t+1} \right) \quad (3-10)$$

Substituindo as equações 3-3 e 3-10 em 3-9 temos que a taxa de desconto estocástica segue o seguinte processo:

$$m_{t+1} = \exp \left(-\frac{1}{2} \lambda_t \lambda_t' - \delta_0 - \delta_1' F_t - \lambda_t' \varepsilon_{t+1} \right) \quad (3-11)$$

3.1.3

Preços e Taxas de Juros dos Títulos

Queremos relacionar as taxas de juros dos títulos com as variáveis de estado. A taxa composta $i_t^{(n)}$ de um título de maturidade n é dada por:

$$i_t^{(n)} = -\frac{\log P_t^{(n)}}{n} \quad (3-12)$$

Sob as hipóteses utilizadas, Duffie e Kan (1996) mostram que o preço de um título de maturidade n pode ser descrito como:

³Para uma demonstração formal da relação entre a existência da derivada de Radon-Nikodyn e ausência de arbitragem, ver Duffie (2001).

$$P_t^{(n)} = \exp(A_n + B_n' F_t) \quad (3-13)$$

Pela equação fundamental de preços sabemos que:

$$P_t^{(n+1)} = E_t(m_{t+1} P_{t+1}^{(n)}) \quad (3-14)$$

Substituindo as equações 3-11 e 3-13 em 3-15 temos então que:

$$P_t^{(n+1)} = E_t \left[\exp \left\{ -\frac{1}{2} \lambda_t \lambda_t' - \delta_0 - \delta_1' F_t - \lambda_t' \varepsilon_{t+1} + A_n + B_n' F_{t+1} \right\} \right] \quad (3-15)$$

Substituindo a equação 3-1 obtemos:

$$\begin{aligned} P_t^{(n+1)} &= E_t \left[\exp \left\{ -\frac{1}{2} \lambda_t \lambda_t' - \delta_0 - \delta_1' F_t - \lambda_t' \varepsilon_{t+1} + A_n + B_n' (\Phi_0 + \Phi F_t + \Sigma \varepsilon_{t+1}) \right\} \right] = \\ &= \exp \left\{ -\delta_0 + A_n + B_n' \Phi_0 + (B_n' \Phi - \delta_1) F_t - \frac{1}{2} \lambda_t \lambda_t' \right\} E_t [\exp \{ (-\lambda_t' + B_n' \Sigma) \varepsilon_{t+1} \}] = \end{aligned}$$

Como ε_t é gaussiano temos que:

$$\begin{aligned} E_t [\exp \{ (-\lambda_t' + B_n' \Sigma) \varepsilon_{t+1} \}] &= \exp \left\{ E_t [(-\lambda_t' + B_n' \Sigma) \varepsilon_{t+1}] + \frac{1}{2} \text{Var}_t [(-\lambda_t' + B_n' \Sigma) \varepsilon_{t+1}] \right\} = \\ &= \exp \left\{ \frac{1}{2} \text{Var}_t [(-\lambda_t' + B_n' \Sigma) \varepsilon_{t+1}] \right\} = \\ &= \exp \left\{ \frac{1}{2} [(-\lambda_t' + B_n' \Sigma) \text{Var}_t \varepsilon_{t+1} (-\lambda_t + \Sigma' B_n)] \right\} = \\ &= \exp \left\{ \frac{1}{2} [-\lambda_t' \lambda_t + B_n' \Sigma \Sigma' B_n - 2 B_n' \Sigma \lambda_t] \right\} \end{aligned}$$

Substituindo na expressão original obtemos:

$$P_t^{(n+1)} = \exp \left\{ -\delta_0 + A_n + B_n' \Phi_0 + (B_n' \Phi - \delta_1) F_t + \frac{B_n' \Sigma \Sigma' B_n}{2} - B_n' \Sigma \lambda_t \right\}$$

Utilizando a equação 3-3 podemos então escrever:

$$P_t^{(n+1)} = \exp \left\{ -\delta_0 + A_n + B_n' (\Phi_0 - \Sigma \lambda_0) + \frac{B_n' \Sigma \Sigma' B_n}{2} + (-\delta_1 - B_n' (\Phi - \Sigma \lambda_1)) F_t \right\}$$

Assim, A_n e B_n são definidos recursivamente por:

$$A_{n+1} = A_n + B_n' (\Phi_0 - \Sigma \lambda_0) + \frac{1}{2} B_n' \Sigma \Sigma' B_n - \delta_0 \quad (3-16)$$

$$B_{n+1} = B_n' (\Phi - \Sigma \lambda_1) - \delta_1 \quad (3-17)$$

Falta encontrarmos os valores iniciais A_1 e B_1 . Temos que:

$$\begin{aligned} P_t^{(1)} &= E_t[m_{t+1}] = E_t \left[\exp \left(-\frac{1}{2} \lambda_t' \lambda_t - \delta_0 - \delta_1' F_t - \lambda_t' \varepsilon_{t+1} \right) \right] = \\ &= \exp \left(-\frac{1}{2} \lambda_t' \lambda_t - \delta_0 - \delta_1' F_t \right) E_t [\exp(-\lambda_t' \varepsilon_{t+1})] \end{aligned}$$

Novamente, como ε_{t+1} é gaussiano, podemos escrever:

$$\begin{aligned} E_t[\exp(-\lambda_t' \varepsilon_{t+1})] &= \exp \left\{ E_t(-\lambda_t' \varepsilon_{t+1}) + \frac{1}{2} \text{Var}_t(-\lambda_t' \varepsilon_{t+1}) \right\} = \\ &= \exp \left(\frac{1}{2} \lambda_t' \lambda_t \right) \end{aligned}$$

Voltando à expressão original:

$$P_t^{(1)} = \exp(-\delta_0 - \delta_1' F_t)$$

e assim $A_1 = -\delta_0$ e $B_1 = -\delta_1$.

Para obtermos as taxas de juros, utilizamos a equação 3-12 e temos portanto:

$$i_t^{(n)} = -\frac{A_n}{n} - \frac{B_n}{n} F_t = a_n + b_n F_t \quad (3-18)$$

onde $a_n = -\frac{A_n}{n}$ e $b_n = -\frac{B_n}{n}$.

Podemos também calcular o prêmio de risco presente na estrutura a termo. Para isso, temos que focar no excesso de retornos de 1 período (*one period excess holding period return*), onde comparamos o retorno obtido ao comprarmos um título de longo-prazo e mantê-lo por um período ao retorno obtido com a taxa de curto-prazo. Dessa forma, o excesso de retornos de 1 período pode ser obtido através da seguinte expressão:

$$rx_{t+1}^{(n)} = \log \left(\frac{P_{t+1}^{(n-1)}}{P_t^{(n)}} \right) - i_t = ni_t^{(n)} - (n-1)i_{t+1}^{(n-1)} - i_t^{(1)} \quad (3-19)$$

O excesso de retornos esperado condicionalmente é portanto:

$$\begin{aligned} E_t(rx_{t+1}^{(n)}) &= -(A_n + B'_n F_t) + A_{n-1} + B'_{n-1} E_t(F_{t+1}) - (\delta_0 + \delta'_1 F_t) = & (3-20) \\ &= -[A_{n-1} + B'_{n-1} (\Phi_0 - \Sigma \lambda_0) + \frac{1}{2} B'_{n-1} \Sigma \Sigma' B_{n-1} - \delta_0] - [B'_{n-1} (\Phi - \Sigma \lambda_1) - \delta'_1] F_t \\ &\quad + A_{n-1} + B'_{n-1} (\Phi_0 + \Phi F_t) - (\delta_0 + \delta'_1 F_t) = \\ &= B'_{n-1} \Sigma \lambda_0 + B'_{n-1} \Sigma \lambda_1 F_t - \frac{1}{2} B'_{n-1} \Sigma \Sigma' B_{n-1} \end{aligned}$$

Desta equação, podemos ver que temos um componente de desigualdade de Jensen, um prêmio de risco constante no tempo e um prêmio de risco variante no tempo. É fácil ver que se $\lambda_1 = 0$, os excessos de retorno esperados são constantes no tempo e assim a Hipótese das Expectativas é válida. Para o caso brasileiro, uma série de estudos como Tabak e Andrade (2001), Lima e Issler (2003), Brito, Duarte e Guillén (2003) e Almeida (2004) rejeitam a validade da Hipótese das Expectativas ao menos parcialmente e portanto temos que especificar $\lambda_1 \neq 0$.

3.2 Modelo com Variáveis Latentes

3.2.1 Especificação do Modelo

Tendo introduzido o modelo afim da curva de juros, gostaríamos agora de verificar a sua aplicabilidade para o caso brasileiro. Para isso, estimamos inicialmente um modelo afim somente com variáveis latentes que, apesar de

não possuir *insights* econômicos, pode servir como base de comparação para o modelo completo.

Inicialmente, seguindo Cochrane (2001), realizamos uma análise de componentes principais para identificar o número de fatores necessários para descrever a dinâmica da curva de juros. A análise de componentes principais consiste na diagonalização da matriz de covariância das taxas, que representa uma medida do risco associado aos movimentos da curva de juros. Através desta técnica, obtemos uma matriz de autovetores e um vetor de autovalores com a matriz de covariância das taxas V satisfazendo a seguinte equação:

$$VA = A\Lambda$$

onde A é a matriz de autovetores e Λ é uma matriz diagonal com o vetor de autovalores na diagonal principal. Utilizando esta relação e sabendo que A é uma matriz ortonormal, podemos escrever:

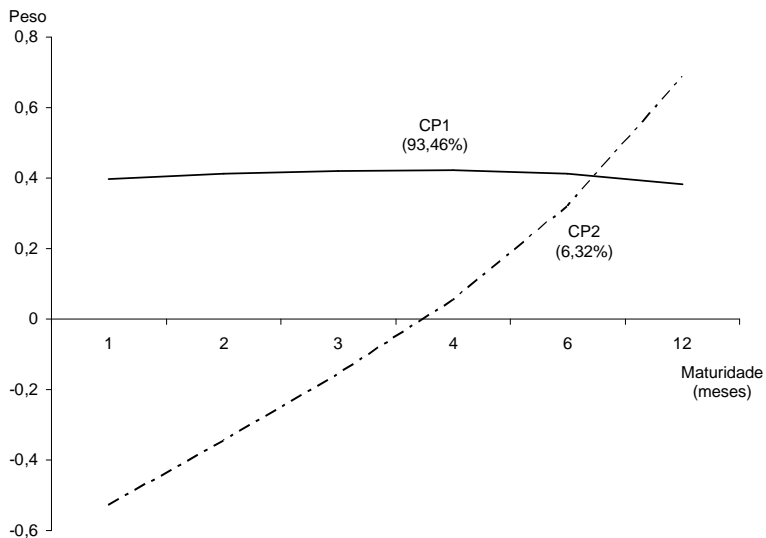
$$V = A\Lambda A'$$

A proporção de risco atribuída a cada fator é obtida através da normalização dos autovalores associados a cada fator obtidos desta decomposição. A tabela 3.1 mostra a variância explicada por cada um dos fatores. Pode-se observar que os dois primeiros componentes principais são responsáveis por 99,8% da variação das taxas. Os componentes obtidos possuem uma interpretação bem clara, similar àquela proposta por Litterman e Scheinkman (1991), com um componente de nível e outro de inclinação. A figura 3.1 mostra os pesos de cada componente nas diferentes taxas.

Tabela 3.1: Variância explicada pelos componentes principais

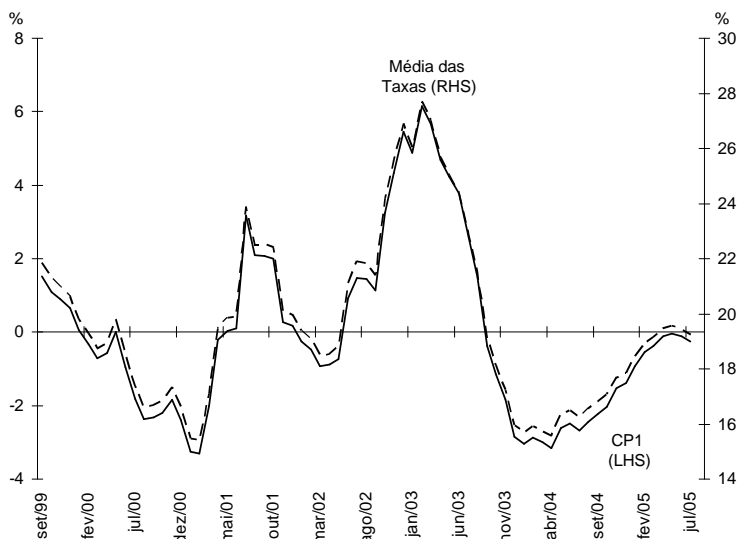
	CP1	CP2	CP3	CP4	CP5	CP6
Autovalor	5,6077	0,3792	0,0116	0,0014	0,0002	4,20E-05
Variância Explicada	0,9346	0,0632	0,0019	0,0002	4E-05	7E-06
Variância Acumulada	0,9346	0,9978	0,9997	1,0000	1,0000	1,0000

Figura 3.1: Ponderação dos componentes principais em cada taxa



O primeiro componente principal, responsável por 93,46% da variação das taxas, é claramente um componente de nível, como pode ser observado na figura 3.2:

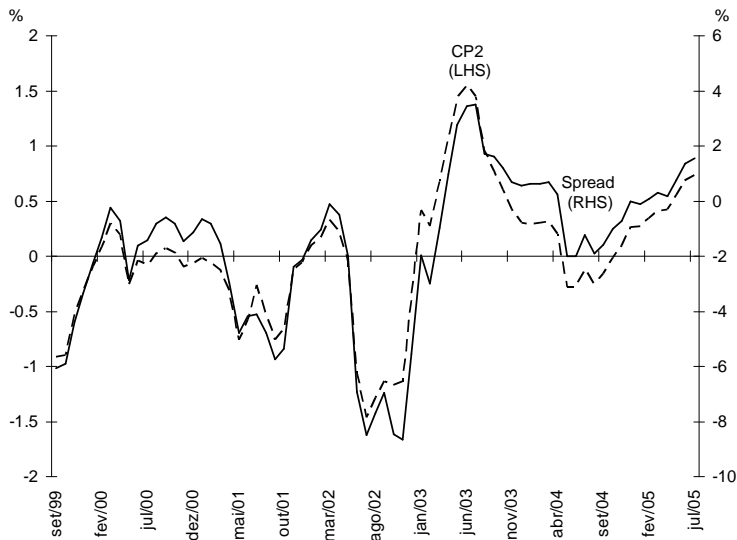
Figura 3.2: Primeiro componente principal e nível da curva de juros



Já o segundo componente principal, responsável por 6,32% da variação das taxas, pode ser comparado à inclinação da curva de juros, medida pela

diferença entre a taxa de 1 mês e a taxa de 12 meses, como pode ser observado na figura 3.3:

Figura 3.3: Segundo componente principal e inclinação da curva de juros



Dessa forma, utilizamos um modelo gaussiano com 2 fatores latentes com volatilidades constantes ou, na classificação de Dai e Singleton (2000), um modelo $A_0(2)$.

3.2.2 Estimação do Modelo

Para estimar o modelo, utilizamos máxima verossimilhança baseada no filtro de Kalman. Para modelos gaussianos, a utilização desta estratégia é ótima com relação à classe dos estimadores lineares, sendo os estimadores consistentes e eficientes, como pode ser visto em Bollerssev e Wooldridge (1992)⁴. Além disso, conseguimos evitar a hipótese usual de que apenas algumas taxas, selecionadas de forma arbitrária, são medidas com erro, como na maioria dos estudos realizados até aqui, que são baseados na estratégia introduzida por Chen e Scott (1993)⁵. Assim, podemos supor que todas as taxas são medidas com erro e deixar que a própria estimação nos forneça esses erros de medida. Dados

⁴Duffee e Stanton (2001) também defendem a utilização do filtro de Kalman como superior às demais abordagens.

⁵A estratégia de Chen e Scott (1993) consiste em obter os fatores latentes invertendo as equações das taxas. Assim, se temos N fatores latentes, temos que ter N taxas medidas sem erro.

os parâmetros do modelo e outras normalizações necessárias, estimadores não-viesados das variáveis latentes L_t e S_t , que representam respectivamente o nível e a inclinação da estrutura a termo da taxa de juros, podem ser obtido através do filtro de Kalman.

Como explicitado na seção anterior, o vetor de estados $F_t = (L_t, S_t)$ segue um processo VAR(1) gaussiano. A equação de estados é dada portanto por:

$$F_t = \Phi F_{t-1} + \Sigma \varepsilon_t \quad (3-21)$$

onde $\varepsilon_t \sim IIDN(0, I_2)$, $\Sigma_{2 \times 2}$ é diagonal e $\Phi_{2 \times 2}$ é triangular inferior.

As taxas utilizadas na estimação são os *swaps* de 1 mês, 2 meses, 3 meses, 4 meses, 6 meses e 12 meses. Dessa forma, as variáveis observáveis podem ser agrupadas no vetor $Z_t \equiv [i_t^{(1)} \ i_t^{(2)} \ i_t^{(3)} \ i_t^{(4)} \ i_t^{(6)} \ i_t^{(12)}]'$.

Por serem extraídos de forma endógena, os fatores latentes são invariantes a algumas transformações afim, tornando necessárias algumas normalizações para identificá-los. Impõem-se assim: (i) média nula para os fatores, ou seja, $\Phi_0 = 0$; (ii) pesos unitários sobre os fatores na equação da taxa de curta-prazo; e (iii) valor de δ_0 fixo na média amostral da taxa de 1 mês. Assim temos como equação da taxa de curto-prazo:

$$i_t^{(1)} = \delta_0 + \delta_1' F_t = 19,0114 + L_t + S_t$$

A equação de medição pode portanto ser escrita como:

$$Z_t = G' + H' F_t + \nu_t \quad (3-22)$$

onde

$$G = [a_1 \ a_2 \ a_3 \ a_4 \ a_6 \ a_{12}]$$

$$H' = \begin{bmatrix} b_1^{(1)} & b_2^{(1)} \\ b_1^{(2)} & b_2^{(2)} \\ b_1^{(3)} & b_2^{(3)} \\ b_1^{(4)} & b_2^{(4)} \\ b_1^{(6)} & b_2^{(6)} \\ b_1^{(12)} & b_2^{(12)} \end{bmatrix}$$

onde a_n e b_n são dados pelas equações recursivas derivadas na seção anterior. Assume-se que os erros de medida ν_t são i.i.d., com distribuição

multi-variada normal de médio zero e matriz de covariância diagonal R , ou seja, $\nu_t \sim IIDN(0, R)$.

Seja $\Upsilon_t = (Z_s : s \leq t)$ o conjunto informacional em t e:

$$\widehat{F}_{t+1|t} \equiv E(F_{t+1}|\Upsilon_t) \quad (3-23)$$

$$P_{t+1|t} \equiv E[(\widehat{F}_{t+1} - F_{t+1|t})^2|\Upsilon_t] \quad (3-24)$$

a previsão ótima do vetor de estados e o erro quadrático médio associado. O algoritmo do filtro de Kalman nos permite calcular previsões e os *EQMs* (erros quadráticos médios) associados de forma recursiva:

$$\begin{aligned} \widehat{F}_{t+1|t} &= \Phi \widehat{F}_{t|t-1} \\ &\quad + \Phi P_{t|t-1} H (H' P_{t|t-1} H + R)^{-1} (Z_t - G' - H' \widehat{F}_{t|t-1}) \end{aligned} \quad (3-25)$$

$$P_{t+1|t} = \Phi [P_{t|t-1} - P_{t|t-1} H (H' P_{t|t-1} H + R)^{-1} H' P_{t|t-1}] \Phi' + \Sigma \Sigma' \quad (3-26)$$

iniciando com média e matriz de covariância não-condicional $\widehat{F}_{1|0} = E(F_1)$ e $P_{1|0} = cov(F_1)$. Sob nossa especificação do VAR, a matriz de covariância não-condicional é dada por $vec(P_{1|0}) = [I - \Phi \otimes \Phi]^{-1} \times vec(\Sigma \Sigma')$.

A função de verossimilhança pode então ser construída sabendo-se que a distribuição condicional do vetor observável Z_{t+1} dado o conjunto informacional Υ_t é multivariada normal:

$$Z_{t+1}|\Upsilon_t \sim N(G' + H' \widehat{F}_{t+1|t}, H' P_{t+1|t} H + R), t \geq 0$$

Todos os parâmetros $\Theta = (\Phi, \Sigma, G, H, R)$ que determinam o comportamento do filtro de Kalman são determinados pelos parâmetros primitivos, através de transformações determinísticas e das restrições de ausência de arbitragem.

Definindo-se $\varepsilon_t(\Theta) \equiv Z_t - G' - H' \widehat{F}_{t+1|t}$ e $\Omega_t(\Theta) \equiv H' P_{t+1|t} H + R$, podemos escrever a função de verossimilhança condicional como:

$$f_{Z_t|\Upsilon_{t-1}}(Z_t|\Upsilon_{t-1}) = (2\pi)^{-1/2} |\Omega_t(\Theta)|^{-1/2} \times \exp \left\{ -\frac{1}{2} \varepsilon_t(\Theta)' \times \Omega_t(\Theta)^{-1} \times \varepsilon_t(\Theta) \right\} \quad (3-27)$$

para $t = 1, 2, \dots, T$. Assim, podemos escrever a função de log-verossimilhança amostral como:

$$\begin{aligned} -2 \times \log L(Z_T) &= -2 \times \sum_{t=1}^T \log f(Z_t | \Upsilon_{t-1}) = \\ &= \sum_{t=1}^T \log |\Sigma_t(\Theta)| \\ &\quad + \sum_{t=1}^T \varepsilon_t(\Theta)' \times \Omega_t(\Theta)^{-1} \times \varepsilon_t(\Theta) \end{aligned} \quad (3-28)$$

A otimização é realizada utilizando o toolbox para MATLAB E^4 , que permite a estimação de diferentes modelos em forma de espaço de estados⁶. O algoritmo utilizado para otimização da função de verossimilhança foi o algoritmo quasi-newton com atualização BFGS. Os erros-padrão são calculados através da inversa de uma aproximação numérica da matriz Hessiana da função de verossimilhança.

3.2.3

Principais Resultados

O modelo foi estimado de Setembro de 1999 a Dezembro de 2004, utilizando-se o período de Janeiro de 2005 a Julho de 2005 para avaliar sua performance de previsão fora da amostra. A tabela 3.2 apresenta os resultados da estimação. Os fatores L_t e S_t são bem persistentes. Existe também uma pequena correlação incondicional entre os fatores. Como pode ser observado na figura 3.4, podemos interpretar novamente os fatores latentes como um fator de nível e outro fator de inclinação da curva de juros. Todos os parâmetros encontrados são altamente significativos⁷ e os desvios-padrão dos erros de medida para as taxas de 1,2, 3, 4, 6 e 12 meses são, respectivamente, 14, 0, 7, 10, 17 e 95 *basis points*, ligeiramente inferiores aos obtidos em Almeida (2004). A variação temporal no prêmio de risco depende primordialmente do fator de inclinação, que apresenta os maiores coeficientes (em módulo) da matriz λ_1 .

⁶Agradecemos a Jaime Terceiro, Jose Manuel Casals, Miguel Jerez, Gregorio R. Serrano e Sonia Sotoca por fornecerem o toolbox para a estimação do modelo e especialmente a Miguel Jerez pela ajuda com questões técnicas relacionadas ao pacote.

⁷Em geral, não se consegue na literatura prêmios de risco muito significativos. Uma potencial razão para a alta significância dos prêmios de risco no caso brasileiro pode ser o fato das taxas de juros não serem estacionárias no período estudado (apesar de na teoria termos que taxas de juros são variáveis estacionárias).

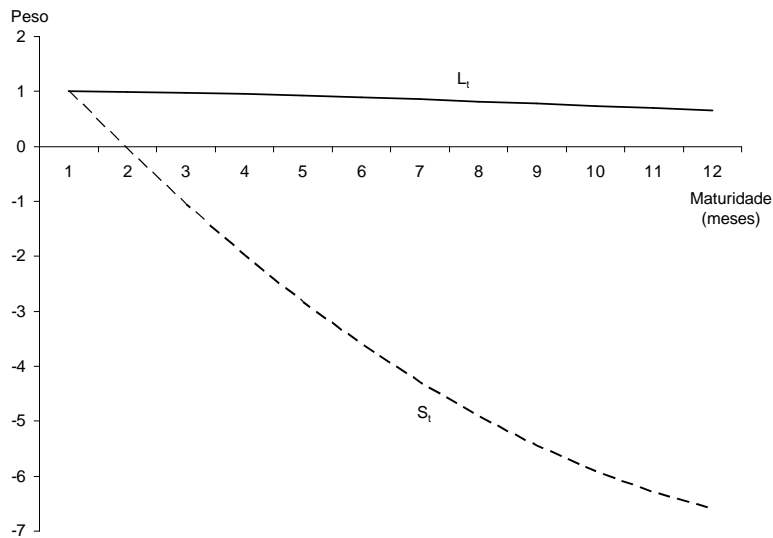
Tabela 3.2: Parâmetros estimados - Modelo com variáveis latentes

Dinâmica dos Fatores (Φ)			
	L_{t-1}	S_{t-1}	
L_t	0,8926 (0,0138)	-	
S_t	0,0175 (0,0014)	0,9067 (0,0097)	
Parâmetros da Taxa de Curto-prazo (δ)			
δ_0	δ_L	δ_S	
19,0114*	1,0000*	1,0000*	
Preços de Risco (λ)			
	λ_0	L_t	S_t
$\lambda_{L,t}$	-1,9768 (0,1563)	-0,0405 (0,0036)	1,0694 (0,0741)
$\lambda_{S,t}$	13,4697 (0,6341)	0,0684 (0,0059)	-0,2033 (0,0181)
Desvios-Padrão (Σ)			
σ_L	1,9086 (0,1603)		
σ_S	0,1278 (0,0047)		
Erros de Medida (R)			
σ_1^2	0,0198 (0,0015)		
σ_2^2	0,0000 (0,0000)		
σ_3^2	0,0045 (0,0001)		
σ_4^2	0,0115 (0,0008)		
σ_6^2	0,0347 (0,0027)		
σ_{12}^2	0,8832 (0,0785)		

Nota: Desvios-padrão das estimativas entre parênteses

Os Parâmetros com asterisco são mantidos fixos como explicitado na subseção anterior

Figura 3.4: Ponderação dos fatores latentes em cada taxa



A adequação do modelo dentro da amostra é muito boa para todas as taxas, como pode ser observado nas figuras 3.5 e 3.6, com um R^2 ajustado superior a 93% para todas as taxas. A precisão é um pouco menor para as taxas mais longas, essencialmente nos períodos de maior volatilidade. No entanto, como já foi observado em outros estudos⁸, o modelo apresenta pior performance para previsões fora da amostra.

⁸Duffee (2002), por exemplo, avalia o modelo afim padrão para os Estados Unidos e verifica que sua performance de previsão é pior que a de um simples passeio aleatório.

Figura 3.5: *Swap* de 1 mês - real vs. estimado

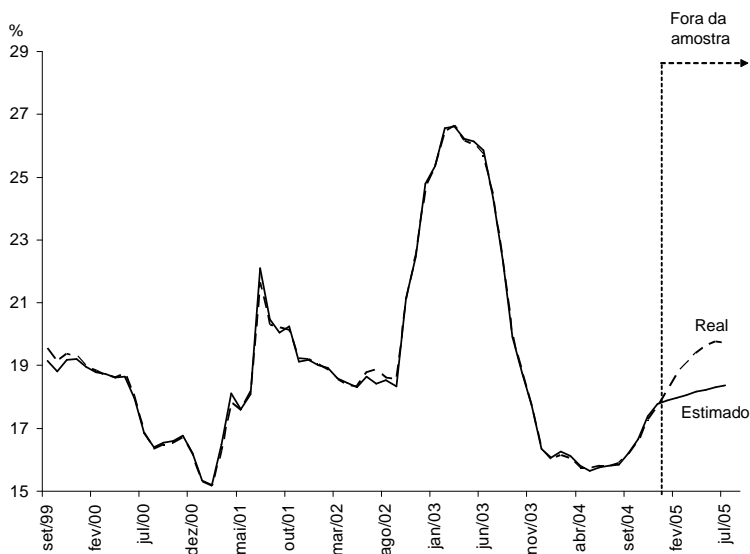
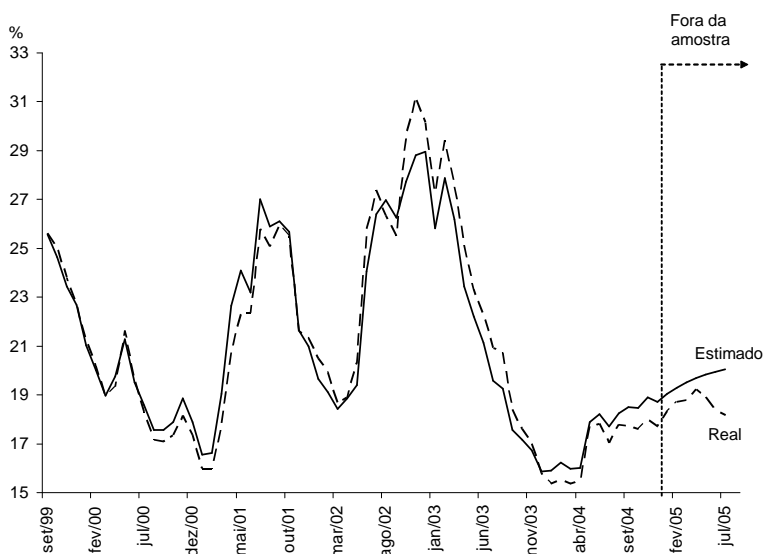
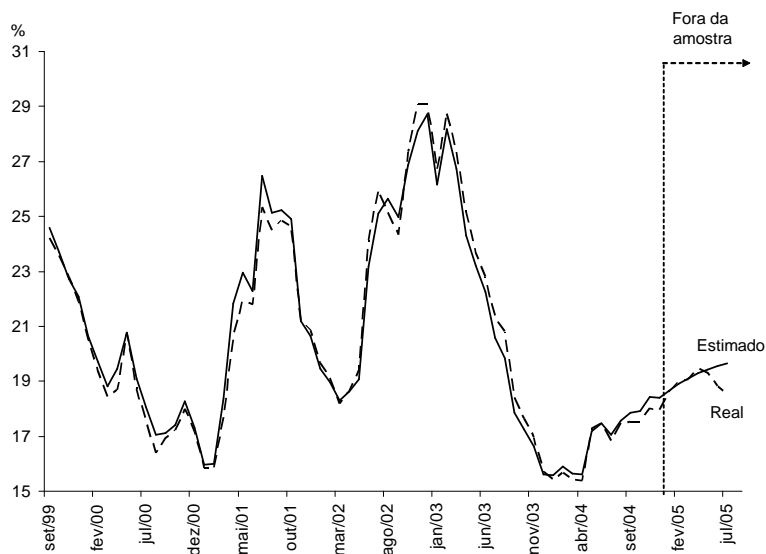


Figura 3.6: *Swap* de 12 meses - real vs. estimado



Como a utilização de um modelo de curva de juros visa também fornecer boas aproximações para as demais taxas, a figura 3.7 mostra a adequação do modelo para a taxa de 9 meses, que não foi incluída na estimação. Como pode ser observado, o modelo também apresenta uma ótima adequação dentro da amostra para esta taxa, confirmando a boa performance do modelo para o caso brasileiro.

Figura 3.7: Swap de 9 meses - real vs. estimado



Sendo nosso modelo gaussiano, é equivalente a um VAR tradicional e podemos assim calcular funções impulso-resposta (FRIs) e decomposição de variâncias, que nos fornece a proporção da variância das taxas que pode ser atribuída a cada um dos fatores. A tabela 3.3 mostra a decomposição de variâncias para horizontes de previsão de 1, 12 e 60 meses:

Tabela 3.3: Decomposição de variâncias

	horizonte (meses)	L	S
$i_t^{(1)}$	1	99,56	0,45
	12	99,57	0,44
	60	99,57	0,43
$i_t^{(2)}$	1	100,00	0,00
	12	100,00	0,00
	60	100,00	0,00
$i_t^{(3)}$	1	99,47	0,53
	12	99,34	0,66
	60	99,30	0,70
$i_t^{(4)}$	1	98,01	1,99
	12	97,28	2,72
	60	97,07	2,93
$i_t^{(6)}$	1	92,66	7,34
	12	88,27	11,73
	60	87,90	12,10
$i_t^{(12)}$	1	64,28	35,72
	12	51,41	48,59
	60	65,91	34,09

O fator de nível domina a decomposição de variâncias para as taxas até 6 meses. Já para a taxa de 12 meses, temos uma maior proporção da variação decorrente do fator inclinação, especialmente em horizontes intermediários.

Podemos concluir que o modelo consegue uma ótima aproximação para a estrutura a termo da taxa de juros brasileira, com exceção dos períodos de grande volatilidade. No entanto, apesar de sua ótima performance dentro da amostra, o modelo não apresenta a mesma precisão para a previsão das taxas fora da amostra. Uma potencial solução para essas questões seria a utilização de uma especificação estocástica para a volatilidade (veja, por exemplo, Balduzzi e outros (1996)). No entanto, modelos com volatilidade estocástica devem respeitar restrições adicionais de admissibilidade, já que os fatores determinantes da volatilidade entram sob a forma de raízes quadradas e assim tem que ser positivos, (como mostrado em Dai e Singleton (2000)). Outro fator limitante é a grande dificuldade de estimação desses modelo, ainda mais se considerarmos a pequena quantidade de dados disponíveis no caso brasileiro. Além disso, Duffee (2002) mostra que estes modelos apresentam resultados de previsão fora da amostra piores que o modelo gaussiano para a curva americana, não sendo claro portanto qual seria o resultado para o caso brasileiro.

Uma alternativa a esta solução é a utilização de variáveis macroeconômicas como fatores adicionais no modelo. Na seção 2, pudemos observar que os períodos de grande volatilidade das taxas coincide com grande volatilidade de variáveis macroeconômicas. Além disso, devido à presença de conteúdo informacional sobre nível de atividade econômica futura na estrutura a termo da taxa de juros, é intuitivo imaginarmos que a adição de variáveis macroeconômicas poderia melhorar seu poder de previsão. Por outro lado, se acreditamos que uma regra de Taylor é uma boa aproximação para a forma de atuação do banco central, esta intuição fica ainda mais clara para a taxa de curto-prazo. Como pode ser visto em Ang, Dong e Piazzesi (2005), uma especificação afim para a estrutura a termo da taxa de juros é condizente com qualquer especificação da regra de Taylor (padrão, *backward* e *forward-looking*). Dessa forma, incluiremos agora algumas variáveis para sumarizar o ambiente macroeconômico e verificaremos a performance deste modelo completo com relação ao modelo somente com variáveis latentes, tanto dentro quanto fora da amostra.

3.3

Modelo Completo

3.3.1

Seleção das Variáveis Macroeconômicas e Dinâmica das Variáveis de Estado

Desde o trabalho seminal de Taylor (1993), uma série de estudos tem utilizado alguma variante da regra de Taylor para representar a atuação do banco central na determinação da taxa de curto-prazo. Num regime de metas de inflação em pequena economia aberta, a função de reação do Banco Central pode ser representada como:

$$i_t = r^* + \pi_t^* + \alpha_\pi(E_t\pi_{t+j} - \pi_{t+j}^*) + \alpha_y y_t + \alpha_{\Delta e} \Delta e_t + u_t \quad (3-29)$$

onde i_t é a taxa de curto-prazo, r^* é a taxa de juros real (potencialmente variante no tempo), $E_t\pi_t$ é a expectativa de inflação, π_t^* é a meta de inflação, y_t é o hiato do produto, Δe_t é a variação da taxa de câmbio nominal e u_t representa choques de política monetária. Neste caso, a autoridade monetária determina a taxa de curto-prazo em seu valor de longo-prazo adicionado de ajustes cíclicos, relacionados a desvios da expectativa de inflação com relação à meta, variações do hiato do produto e movimentos da taxa de câmbio⁹. Além disso, como pode ser visto em Ang, Dong e Piazzesi (2005), regras de Taylor *forward-looking* como estas são compatíveis com o modelo afim já que, após colocar o sistema sob a forma de espaço de estados, as expectativas de inflação podem ser obtidas como função das variáveis correntes.

Por outro lado, num regime de livre flutuação cambial, a paridade descoberta da taxa de juros caracteriza a relação entre a taxa de juros doméstica e a taxa de juros internacional:

$$i_t = i_t^* + E_t(\Delta e_{t+1}) + x_t$$

onde i_t^* é a taxa de juros internacional e x_t é o prêmio de risco ou risco-país. Assim temos como variáveis potenciais para o modelo π_t^* , π_t , y_t , Δe_t , i_t^* e x_t . No entanto, torna-se impossível incluir mais de três variáveis endógenas no modelo sem prejudicar a qualidade da estimação devido ao aumento da

⁹Minella e outros (2003) defendem a utilização de uma função de reação do Banco Central Brasileiro muito semelhante para o período após a adoção do regime de metas de inflação.

dimensão do modelo e da pequena amostra disponível. Além disso, devido ao processo de construção de credibilidade da autoridade monetária no Brasil recente, a meta de inflação percebida pelos agentes pode ser diferente do centro da meta determinado pelo banco central¹⁰. Finalmente, o risco-país não é diretamente observável e também está associado aos fundamentos macroeconômicos. Sendo assim, decidimos incluir como variáveis macroeconômicas aquelas responsáveis pelo ajuste cíclico da taxa de curto-prazo:

$$X_t^o = \left(y_t \quad \pi_t \quad \Delta e_t \right)'$$

Para captar as variações do nível de longo-prazo da taxa de juros, os choques de política monetária e a variação dos fatores externos, temos que incluir também 2 fatores latentes, $X_t^u = \left(L_t \quad S_t \right)'$ ¹¹. Assim, o modelo completo consiste de 3 fatores macro e 2 fatores latentes.

A seleção do número de defasagens é uma questão delicada. Ao mesmo tempo que a inclusão de um maior número de defasagens é condizente com uma caracterização mais realista do mecanismo de transmissão de política monetária (como pode ser visto no capítulo 2), ela pode prejudicar a precisão da estimação dos coeficientes, especialmente em ocasiões em que o tamanho da amostra é pequeno, como é o caso brasileiro. Ao estimarmos um VAR com taxas no mesmo estilo de Evans e Marshall (1998 e 2001), a inclusão de duas defasagens parece acomodar melhor este *trade-off*¹².

Dessa forma, especificamos o vetor de estados $X_t = (X_t^o, X_t^u) = \left(y_t \quad \pi_t \quad \Delta e_t \quad L_t \quad S_t \right)'$ como um processo gaussiano VAR(2):

$$X_t = \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \Theta u_t \quad (3-30)$$

com $u_t \sim IIDN(0, I_5)$.

Definindo $F_t = \left(X_t \quad X_{t-1} \right)$, podemos reescrever a dinâmica na forma compacta de um VAR gaussiano de primeira ordem:

$$F_t = \Phi_0 + \Phi F_{t-1} + \Sigma \varepsilon_t \quad (3-31)$$

¹⁰Diversos estudos mostraram que, na crise eleitoral de 2002 por exemplo, os agentes já trabalhavam com uma meta de inflação implícita diferente daquela determinada pelo banco central, fato este confirmado pela alteração posterior da meta.

¹¹Tentou-se uma especificação com somente 1 variável latente, mas os resultados obtidos foram pouco satisfatórios.

¹²A inclusão de um maior número de defasagens torna a estimação bastante imprecisa, o que é comprovado pela grande instabilidade das respostas impulsionais.

onde

$$\begin{aligned}\Sigma &= \begin{bmatrix} \Theta & \mathbf{0}_{5 \times 5} \\ \mathbf{0}_{5 \times 5} & \mathbf{0}_{5 \times 5} \end{bmatrix} \\ \Phi_0 &= \left(\phi'_0 \quad \mathbf{0}_{1 \times 5} \right)' \\ \Phi &= \begin{bmatrix} \phi_1 & \phi_2 \\ \mathbf{I}_{5 \times 5} & \mathbf{0}_{5 \times 5} \end{bmatrix}\end{aligned}$$

3.3.2

Estimação do Modelo

Utiliza-se novamente máxima verossimilhança baseada no filtro de Kalman. Adicionalmente às vantagens relacionados na estimação do modelo só com variáveis latentes, esta abordagem nos permite maior flexibilidade na especificação do modelo com relação à abordagem tradicional de Chen e Scott (1993), que nos obrigaria a fazer hipóteses extremamente restritivas para que conseguíssemos estimar o modelo. Ang e Piazzesi (2003), por exemplo, assumem que a dinâmica macroeconômica é independente das variáveis latentes, enquanto Rudebusch e Wu (2005) definem valores arbitrários do prêmio de risco como nulos. Por outro lado, a estimação de modelos de maior dimensão e altamente não-lineares como o utilizado aqui através do filtro de Kalman torna-se especialmente sensível às condições iniciais, sendo necessário cuidado redobrado para garantir que não estamos utilizando um mínimo local da função de verossimilhança¹³.

As taxas utilizadas para construir o filtro de Kalman, da mesma forma que no modelo da seção anterior, são as taxas de 1 mês, 2 meses, 3 meses, 4 meses, 6 meses e 12 meses. Dessa forma, as variáveis observáveis podem ser agrupadas no vetor

¹³Tecnicamente, não temos como garantir que atingimos um mínimo global. No entanto, utilizamos uma série de estratégias diferentes para reduzir o risco de termos utilizado os resultados de um mínimo local. Entre estas, a cada valor encontrado, tentou-se a reestimação do modelo supondo as variáveis latentes como observáveis e obtendo a estimação das matrizes de preços de risco minimizando o erro quadrático médio das taxas estimadas. Dessa forma, acreditamos termos obtido resultados confiáveis.

$$Z_t \equiv \begin{bmatrix} i_t^{(1)} \\ i_t^{(2)} \\ i_t^{(3)} \\ i_t^{(4)} \\ i_t^{(6)} \\ i_t^{(12)} \\ y_t \\ \pi_t \\ \Delta e_t \\ y_{t-1} \\ \pi_{t-1} \\ \Delta e_{t-1} \end{bmatrix}$$

A equação de medição pode ser escrita novamente como:

$$Z_t = G' + H'F_t + \nu_t \tag{3-32}$$

onde agora temos que:

$$G = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_6 & a_{12} & 0_{1 \times 6} \end{bmatrix}$$

$$H' = \begin{bmatrix} b_1^{(1)} & b_2^{(1)} & b_3^{(1)} & b_4^{(1)} & b_5^{(1)} & b_6^{(1)} & b_7^{(1)} & b_8^{(1)} & b_9^{(1)} & b_{10}^{(1)} \\ b_1^{(2)} & b_2^{(2)} & b_3^{(2)} & b_4^{(2)} & b_5^{(2)} & b_6^{(2)} & b_7^{(2)} & b_8^{(2)} & b_9^{(2)} & b_{10}^{(2)} \\ b_1^{(3)} & b_2^{(3)} & b_3^{(3)} & b_4^{(3)} & b_5^{(3)} & b_6^{(3)} & b_7^{(3)} & b_8^{(3)} & b_9^{(3)} & b_{10}^{(3)} \\ b_1^{(4)} & b_2^{(4)} & b_3^{(4)} & b_4^{(4)} & b_5^{(4)} & b_6^{(4)} & b_7^{(4)} & b_8^{(4)} & b_9^{(4)} & b_{10}^{(4)} \\ b_1^{(6)} & b_2^{(6)} & b_3^{(6)} & b_4^{(6)} & b_5^{(6)} & b_6^{(6)} & b_7^{(6)} & b_8^{(6)} & b_9^{(6)} & b_{10}^{(6)} \\ b_1^{(12)} & b_2^{(12)} & b_3^{(12)} & b_4^{(12)} & b_5^{(12)} & b_6^{(12)} & b_7^{(12)} & b_8^{(12)} & b_9^{(12)} & b_{10}^{(12)} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Assume-se que os erros de medida ν_t são i.i.d., com distribuição multivariada normal de médio zero e matriz de covariância R . Adicionalmente, assume-se que as variáveis de estado observáveis não contém erros de medida,

e assim os últimos 6 elementos de ν_t são nulos e R é identicamente nula com exceção da sub-matriz 6×6 esquerda-superior, que representa a matriz de covariância dos erros de medida nas taxas observáveis e é uma matriz diagonal. Utilizando-se o mesmo procedimento de filtragem da seção anterior, podemos obter os parâmetros relevantes para descrever a dinâmica do sistema, a taxa de curto-prazo e os preços de risco de mercado, além de uma série temporal para os fatores latentes.

Para tornar o modelo mais tratável e motivados por regressões irrestritas utilizando os fatores latentes obtidos no modelo afim, restringimos A_n e B_n e os preços de risco a dependerem apenas de X_t . Tecnicamente, esta restrição é obtida impondo que a taxa de curto-prazo tenha pesos diferentes de zero somente nas variáveis contemporâneas e que a dinâmica de F_t seja um VAR(1) sob a medida neutra-risco. Além disso, podemos, sem perda de generalidade, utilizar variáveis de estado com média nula e assim temos que $\Phi_0 = \mathbf{0}$.¹⁴

Temos também que realizar algumas normalizações para obter um estimador não-viesado das variáveis latentes L_t e S_t através do filtro de Kalman. Para isso, fixamos novamente δ_0 como a média incondicional da taxa de 1 mês e utilizamos pesos unitários nos fatores latentes. Temos agora portanto como equação da taxa de curto-prazo:

$$i_t^{(1)} = \delta_0 + \delta_1' F_t = 19,0114 + \delta_y * y_t + \delta_\pi * \pi_t + \delta_{\Delta e} * \Delta e + L_t + S_t \quad (3-33)$$

Além disso, como a autoridade monetária brasileira no período segue um regime de metas de inflação com uma regra de política monetária *forward-looking*, temos que realizar algumas normalizações adicionais. Como pode ser visto em Ang, Dong e Piazzesi (2005), para este tipo regra de política monetária, alguns parâmetros associados ao horizonte de previsão não estão identificados. Para identificá-los, temos que impor a hipótese de identificação adicional de ausência de correlação condicional entre os fatores latentes e macroeconômicos. Finalmente, seguindo a estratégia da seção anterior, impusemos ausência de correlação entre as variáveis latentes.

¹⁴Para isso, utilizamos o hiato do produto, inflação e variação do câmbio nominal sem a sua média amostral.

3.3.3

Principais Resultados

As tabelas 3.4 e 3.5 mostram os resultados da estimação. Todos os parâmetros encontrados são altamente significativos¹⁵ e os desvios-padrão dos erros de medida para as taxas de 1,2, 3, 4, 6 e 12 meses são, respectivamente, 18, 0, 10, 10, 0 e 50 *basis points*.

Tabela 3.4: Parâmetros estimados - Dinâmica dos fatores - Modelo completo

Dinâmica dos Fatores (ϕ_1)					
	y_{t-1}	π_{t-1}	Δe_{t-1}	L_{t-1}	S_{t-1}
y_t	0,7044 (0,0215)	-0,6765 (0,0206)	0,0040 (0,0001)	-0,0292 (0,0009)	1,5059 (0,0384)
π_t	0,0729 (0,0023)	0,4874 (0,0152)	0,0315 (0,0010)	0,0740 (0,0024)	1,1116 (0,0311)
Δe_t	-0,6771 (0,0203)	0,2068 (0,0066)	0,3452 (0,0110)	-0,8599 (0,0250)	-6,4580 (0,0601)
L_t	-0,0062 (0,0002)	-0,0188 (0,0006)	0,0364 (0,0012)	1,0103 (0,0140)	1,6570 (0,0405)
S_t	0,0006 (0,0000)	-0,0170 (0,0005)	-0,0236 (0,0007)	-0,0282 (0,0009)	0,8519 (0,0253)
Dinâmica dos Fatores (ϕ_2)					
	y_{t-2}	π_{t-2}	Δe_{t-2}	L_{t-2}	S_{t-2}
y_t	-0,0945 (0,0030)	-0,0037 (0,0001)	0,0521 (0,0017)	-0,2366 (0,0076)	-1,2406 (0,0337)
π_t	-0,0534 (0,0017)	-0,0314 (0,0010)	0,0433 (0,0014)	-0,0394 (0,0013)	-1,1321 (0,0315)
Δe_t	0,5930 (0,0183)	-0,7098 (0,0215)	-0,0753 (0,0024)	0,5666 (0,0175)	4,8207 (0,0582)
L_t	0,0450 (0,0014)	0,4345 (0,0137)	0,0697 (0,0022)	-0,1005 (0,0032)	-2,3086 (0,0480)
S_t	0,0105 (0,0003)	-0,0101 (0,0003)	-0,0006 (0,0000)	0,0382 (0,0012)	-0,1244 (0,0040)
Decomposição de Cholesky da Matriz de Covariância (Θ)					
	y_t	π_t	Δe_t	L_t	S_t
y_t	1,2786 (0,0409)	0*	0*	0*	0*
π_t	0,0440 (0,0014)	0,3392 (0,0094)	0*	0*	0*
Δe_t	0,0223 (0,0007)	-0,1376 (0,0044)	1,7605 (0,0537)	0*	0*
L_t	0*	0*	0*	0,7567 (0,0232)	0*
S_t	0*	0*	0*	0*	0,0765 (0,0014)

¹⁵Como no caso somente com variáveis latentes, a alta significância das variáveis pode ser decorrente da não-estacionariedade das taxas de juros no período estudado.

Tabela 3.5: Parâmetros estimados - Preços de risco - Modelo completo

Parâmetros da Taxa de Curto-prazo (δ)					
δ_0	δ_y	δ_π	$\delta_{\Delta e}$	δ_L	δ_S
19,0114*	-0,0129 (0,0000)	-0,1145 (0,0102)	0,0355 (0,0000)	1,0000*	1,0000*

Matriz de Risco (λ)					
	y_t	π_t	Δe_t	L_t	S_t
$\lambda_{y,t}$	0,4044 (0,0129)	2,7682 (0,0513)	-0,2353 (0,0075)	1,6523 (0,0404)	6,0230 (0,0597)
$\lambda_{\pi,t}$	0,5297 (0,0165)	0,2510 (0,0080)	0,0463 (0,0015)	0,9248 (0,0268)	4,0563 (0,0566)
$\lambda_{\Delta e,t}$	0,7975 (0,0226)	-1,2714 (0,0340)	-1,1561 (0,0318)	0,0673 (0,0022)	8,7186 (0,0612)
$\lambda_{L,t}$	-0,1413 (0,0044)	0,1407 (0,0045)	0,0595 (0,0020)	-0,0406 (0,0013)	4,2363 (0,0565)
$\lambda_{S,t}$	0,6496 (0,0193)	-0,7251 (0,0218)	-0,8201 (0,0243)	-0,1783 (0,0057)	7,1741 (0,0605)

Preços de Risco (λ_0)		Erros de Medida (R)	
$\lambda_{y,t}$	7,2835 (0,0606)	σ_1^2	0,0316 (0,0005)
$\lambda_{\pi,t}$	-12,3730 (0,0620)	σ_2^2	0,0000 (0,0000)
$\lambda_{\Delta e,t}$	-5,8443 (0,0595)	σ_3^2	0,0097 (0,0001)
$\lambda_{L,t}$	-0,8252 (0,0245)	σ_4^2	0,0103 (0,0001)
$\lambda_{S,t}$	-1,7859 (0,0424)	σ_6^2	0,0000 (0,0000)
		σ_{12}^2	0,2453 (0,0064)

As figura 3.8 e 3.9 mostram os pesos b_n correspondentes a cada fator e cada maturidade, que representam a resposta inicial a choques nestas variáveis de cada taxa. Novamente, como já era esperado, os fatores latentes representam claramente um fator de nível e outro de inclinação, enquanto o impacto das variáveis macroeconômicas é crescente em módulo com a maturidade.

Figura 3.8: Ponderação dos fatores latentes nas taxas

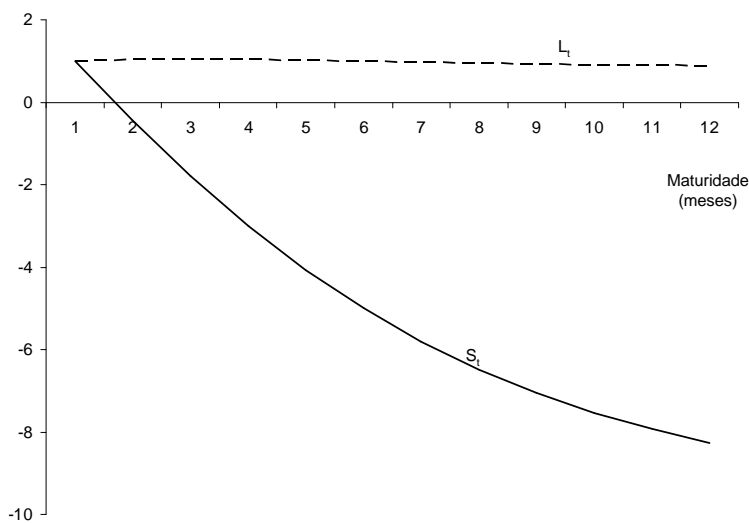
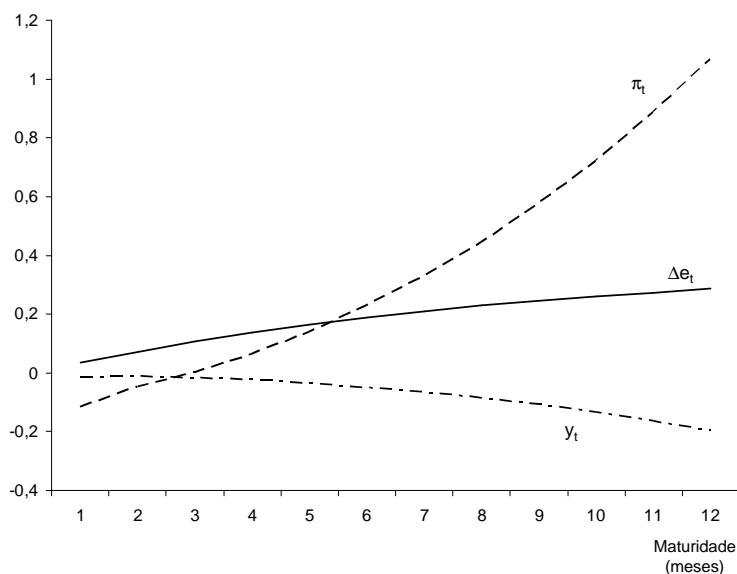


Figura 3.9: Ponderação dos fatores macroeconômicos nas taxas



Podemos interpretar essas respostas iniciais como a presença de uma certa inércia na taxa de inflação, que faz com que um choque inflacionário hoje aumente as taxas de longo-prazo devido ao aumento das expectativas de inflação futura. Além disso, representam as defasagens do *pass-through* das variações cambiais para a inflação levando por isso também a um aumento maior nas taxas de maior maturidade. Finalmente, a relação crescentemente negativa com o hiato do produto está relacionada com os períodos de crise, onde temos ao mesmo tempo uma queda no hiato do produto e um aumento na inclinação da curva de juros.

Repetindo os resultados de estudos anteriores, os fatores latentes no modelo só com variáveis latentes e no modelo completo são semelhantes, com correlação de 0,99 e 0,92 respectivamente para o nível e inclinação, como pode ser visto nas figuras 3.10 e 3.11. Além disso, o fator de inclinação é menos volátil no modelo completo, já que parte da resposta às crises agora está sendo captada pelas variáveis macroeconômicas.

Figura 3.10: Fator de nível nos modelos

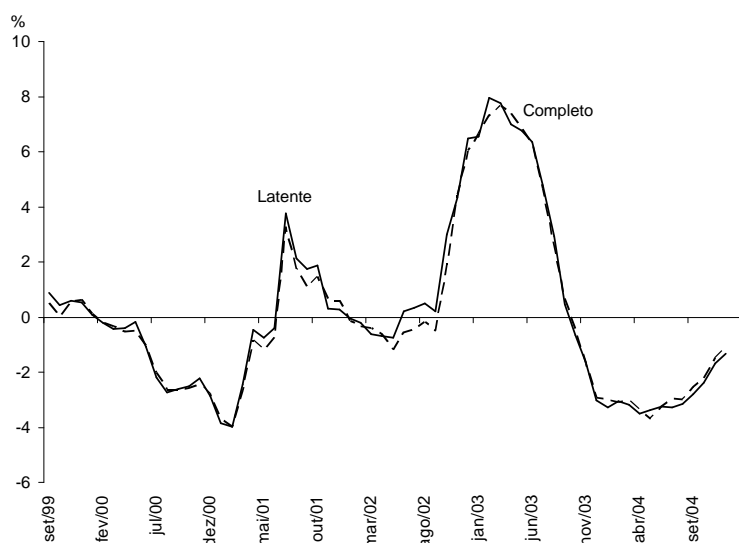
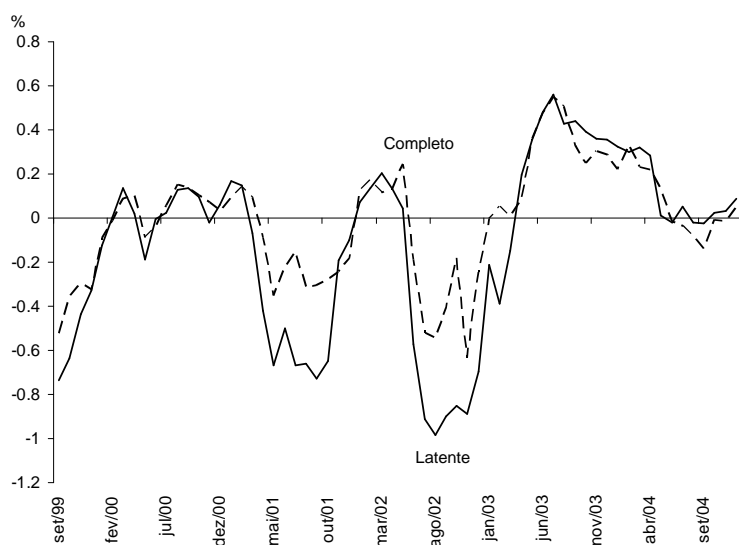


Figura 3.11: Fator de inclinação nos modelos



A adequação do modelo dentro da amostra é novamente muito boa para todas as taxas, como pode ser observado nas figuras 3.12 e 3.13, melhorando a adequação das taxas mais longas com relação ao modelo somente com as variáveis latentes e com um R^2 ajustado superior a 97% para todas as taxas.

Figura 3.12: *Swap* de 1 mês - Comparação dos modelos

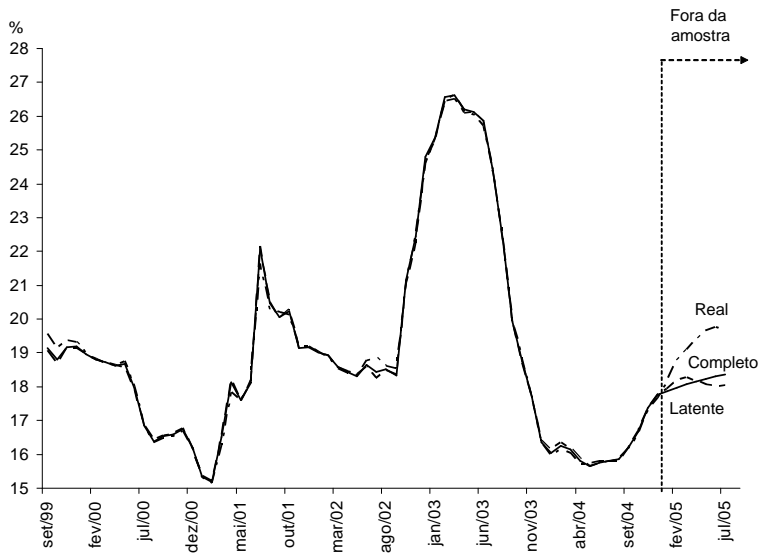
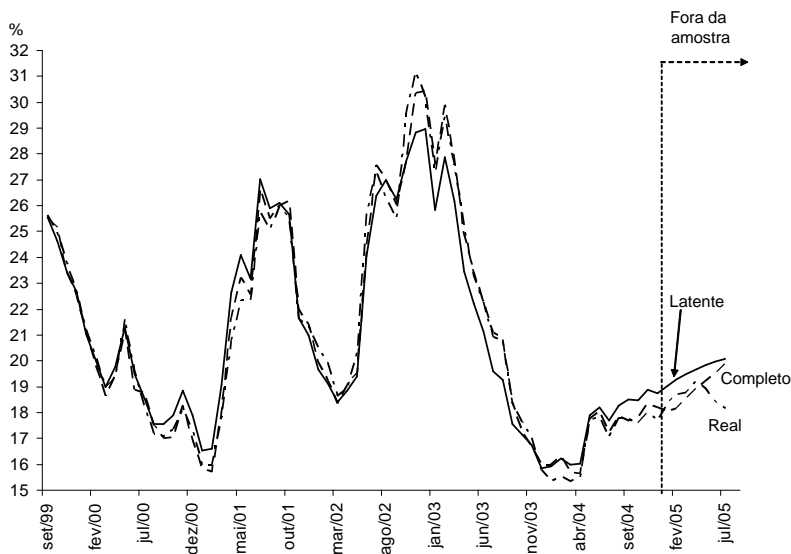
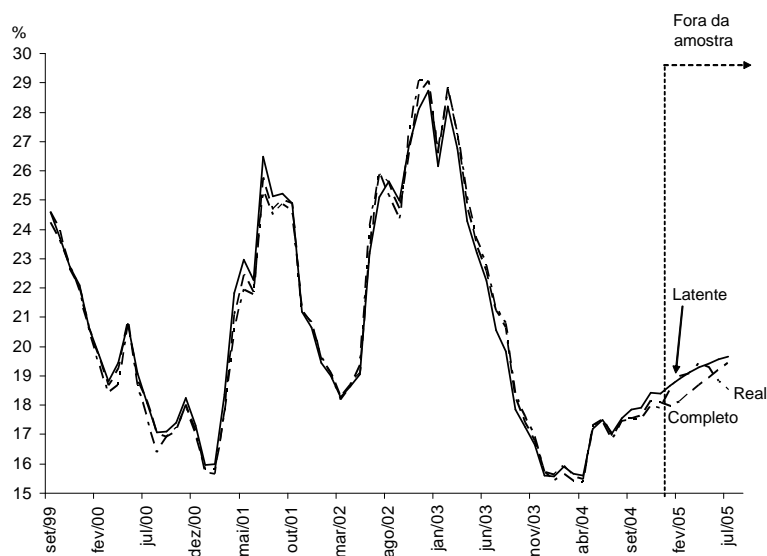


Figura 3.13: *Swap* de 12 meses - Comparação dos modelos



Consequimos novamente obter uma boa aproximação para taxas não incluídas na estimação, como pode ser observado na figura 3.14, que mostra os valores estimados para a taxa de 9 meses.

Figura 3.14: *Swap* de 9 meses - Comparação dos modelos

Comparamos também a performance de previsão fora da amostra dos dois modelos e de um *random-walk* para diferentes períodos à frente, utilizando a métrica RMSE¹⁶. A tabela 3.6 mostra os resultados desta comparação:

¹⁶RMSE é a raiz quadrada da média do quadrado dos erros para cada horizonte de previsão.

Tabela 3.6: Poder de previsão (RMSE)

	horizonte (meses)	<i>RW</i>	<i>Latente</i>	<i>Completo</i>
$i_t^{(1)}$	1	0,52	0,31	0,23
	3	1,08	0,76	0,58
	6	1,57	1,12	1,16
$i_t^{(2)}$	1	0,58	0,44	0,46
	3	1,08	0,80	0,74
	6	1,52	1,07	1,17
$i_t^{(3)}$	1	0,59	0,46	0,58
	3	1,06	0,77	0,83
	6	1,46	0,98	1,13
$i_t^{(4)}$	1	0,62	0,44	0,67
	3	1,07	0,69	0,89
	6	1,41	0,83	1,06
$i_t^{(6)}$	1	0,65	0,26	0,71
	3	1,06	0,45	0,90
	6	1,31	0,48	0,89
$i_t^{(12)}$	1	0,62	0,72	0,26
	3	0,93	0,68	0,36
	6	1,07	0,90	0,56

Ambos os modelos possuem um poder de previsão melhor que um *random-walk*¹⁷ para quase todas as taxas e horizontes. O modelo completo apresenta melhor performance para as taxas de 1 e 12 meses, enquanto o modelo só com variáveis latentes vence os demais para as taxas intermediárias. Assim, não podemos afirmar de forma clara qual modelo é melhor para previsões da curva como um todo fora da amostra.

O modelo completo também é capaz de representar diferentes formatos da curva de juros, conseguindo uma boa aproximação mesmo em casos de grande inclinação, como pode ser observado nas figuras 3.15 a 3.17, que mostram as curvas implícitas pelo modelo e as reais em meses selecionados.

¹⁷Em geral, é extremamente difícil que um modelo em finanças tenha um poder de previsão melhor que o de um *random-walk*, o que valoriza muito os resultados aqui obtidos.

Figura 3.15: Curva de juros - Julho de 2002

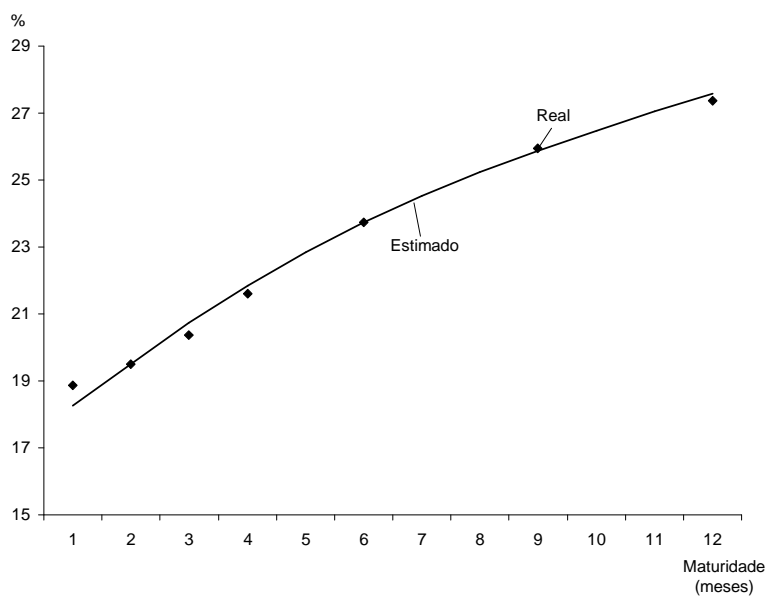


Figura 3.16: Curva de juros - Março de 2000

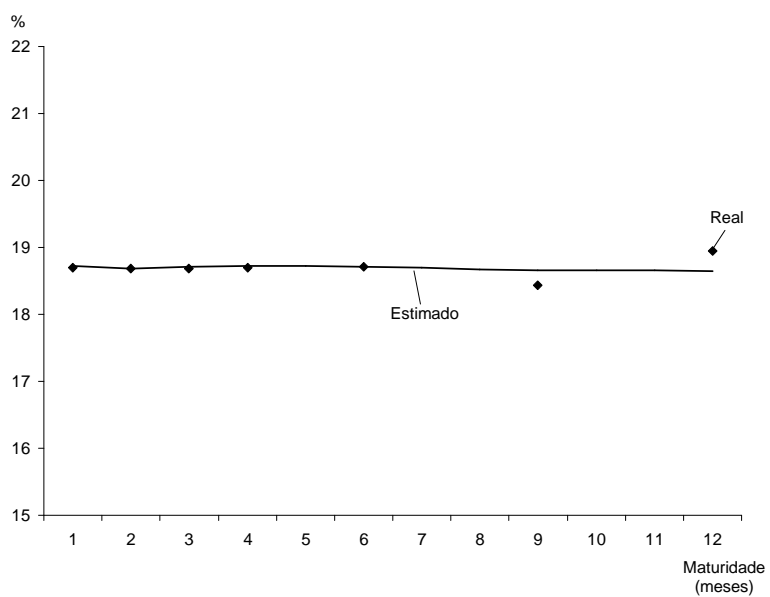
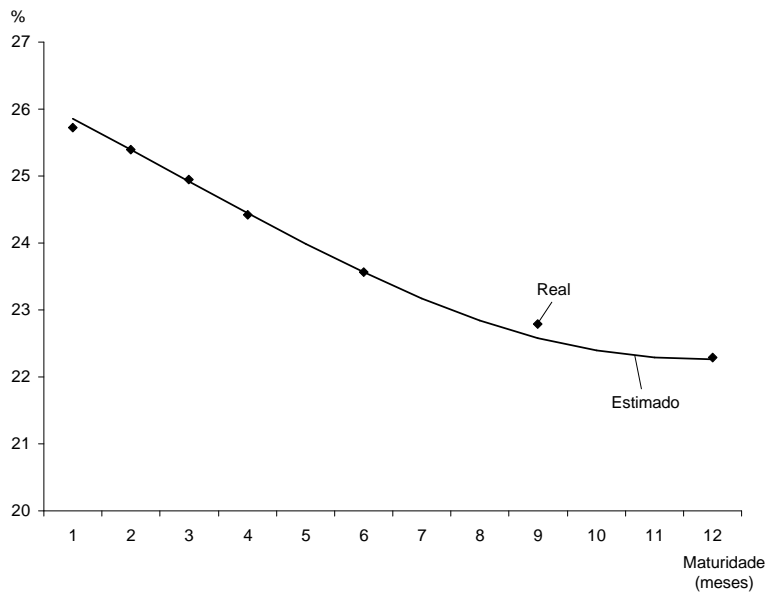


Figura 3.17: Curva de juros - Junho de 2003



Novamente, podemos interpretar nosso modelo como um VAR gaussiano e assim podemos computar funções impulso-resposta e decomposição de variâncias. O cálculo das funções impulso-resposta e da decomposição de variâncias é baseado na decomposição de Cholesky da matriz de covariância das inovações das variáveis de estado (Σ) na ordem $(y_t \ \pi_t \ \Delta e_t \ L_t \ S_t)$. Os resultados das respostas impulsivas podem ser observados nas figuras 3.18 a 3.20, onde apresentamos o efeito dos choques macroeconômicos nas taxas.

Figura 3.18: Impacto de choques de inflação nas taxas

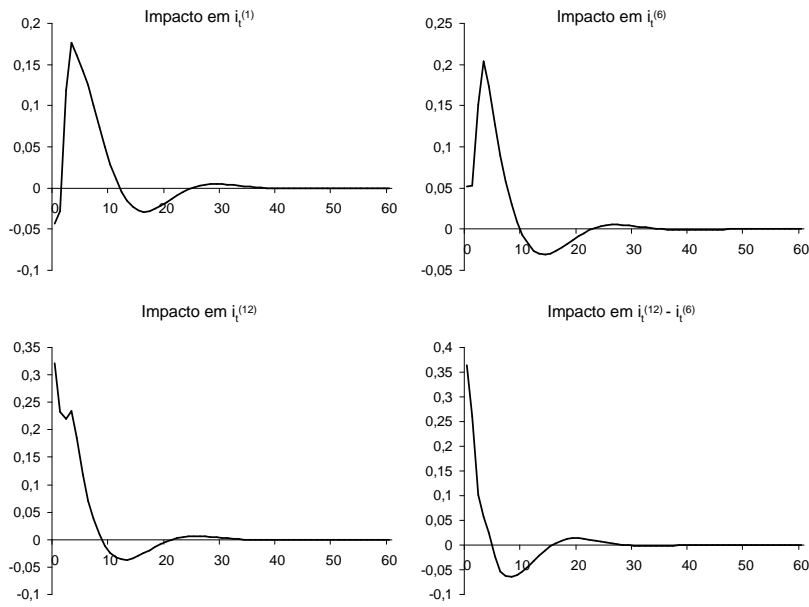


Figura 3.19: Impactos de choques cambiais nas taxas

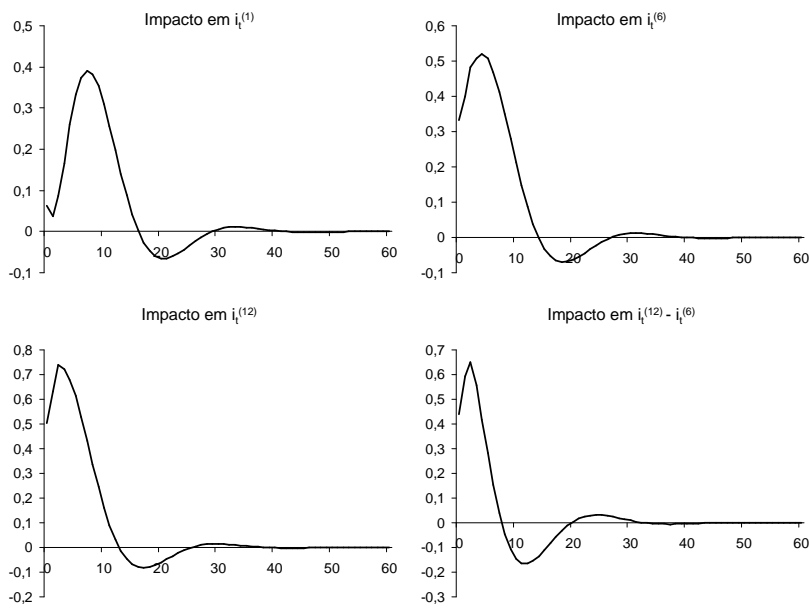
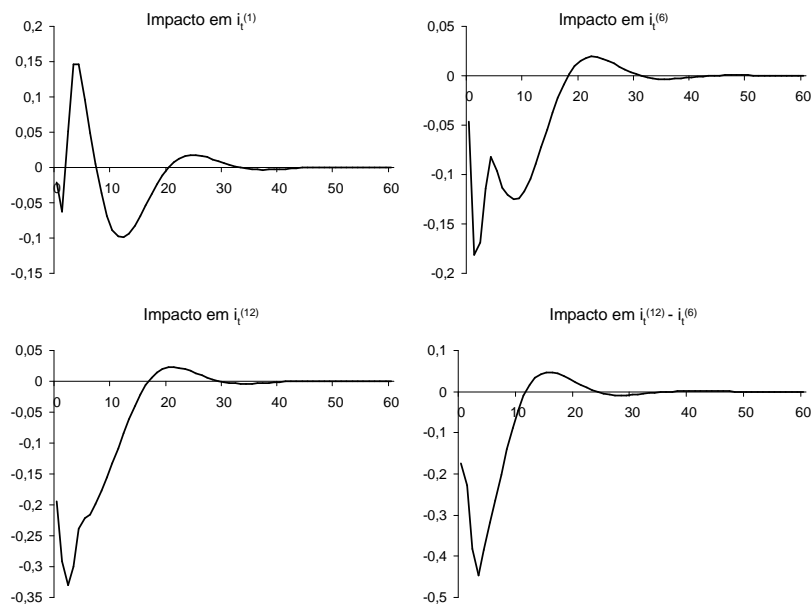


Figura 3.20: Impacto de choques de produto nas taxas



Os resultados obtidos com relação a choques cambiais e na inflação estão dentro do esperado, com uma desvalorização cambial ou uma surpresa inflacionária levando a um aumento nas taxas de todas maturidades. Já o efeito dos choques no hiato do produto levam a uma pequena redução inicial para um posterior aumento da taxa de 1 mês, enquanto a redução na taxa de 12 meses é mais persistente. Este efeito na taxa de 12 meses é resultado da resposta das variáveis a crise externas, onde temos ao mesmo tempo depreciações cambiais e reduções no hiato do produto. Assim, um aumento no hiato do produto é relacionado com apreciações cambiais e, conseqüentemente, com reduções nas taxas longas decorrentes de menores expectativas inflacionárias no futuro. Finalmente, o *spread* apresenta resultados dentro do esperado, com uma redução após choques de inflação (após um breve aumento inicial) e produto e aumento após choques cambiais, decorrentes de uma maior expectativa de inflação futura. Tanto o choque de um desvio-padrão na inflação quanto no produto tem efeitos absolutos bem menores que os choques cambiais, o que é condizente com a percepção da alta vulnerabilidade externa da economia brasileira no período estudado. Além de serem coerentes com o mecanismo tradicional de transmissão da política monetária, estes resultados podem evidenciar respostas concomitantes das variáveis macroeconômicas e da curva de juros às crises externas sofridas pelo país no período.¹⁸

¹⁸Isto ocorre pois, após uma crise externa, temos uma desvalorização cambial (que leva a um aumento da inflação) e um aumento das taxas de juros como resposta à fuga de capitais, sem que tenhamos uma causalidade direta entre o aumento da inflação esperada e o aumento das taxas.

Podemos também decompor a variância dos erros de previsão decorrentes dos diferentes tipos de choque. A tabela 3.7 mostra a decomposição de variâncias para horizontes de previsão de 1, 12 e 60 meses:

Tabela 3.7: Decomposição de variâncias

	horizonte (meses)	y	π	Δe	L	S
$i_t^{(1)}$	1	0,39	0,24	0,47	96,01	2,88
	12	2,21	3,01	22,93	61,46	10,39
	60	2,72	2,99	23,15	60,15	10,98
$i_t^{(2)}$	1	0,60	0,07	2,63	96,30	0,40
	12	1,56	2,94	26,25	58,66	10,58
	60	2,03	2,94	26,25	57,76	11,02
$i_t^{(3)}$	1	0,97	0,02	5,95	91,51	1,55
	12	1,52	2,83	29,60	53,89	12,16
	60	1,90	2,84	29,46	53,37	12,42
$i_t^{(4)}$	1	1,35	0,02	9,56	84,19	4,89
	12	1,94	2,68	32,57	48,26	14,56
	60	2,23	2,70	32,37	48,03	14,67
$i_t^{(6)}$	1	2,08	0,32	15,98	68,37	13,25
	12	3,48	2,46	36,89	37,40	19,77
	60	3,64	2,49	36,68	37,48	19,71
$i_t^{(12)}$	1	4,78	6,08	25,21	38,83	25,10
	12	7,78	4,09	41,44	20,09	26,60
	60	7,78	4,09	41,34	20,33	26,47

Uma série de conclusões podem ser obtidas desta decomposição. Em geral, a variância decorrente dos fatores macroeconômicos é maior para taxas mais longas, ao contrário do que ocorre para as taxas americanas¹⁹. Além disso, mesmo para as taxas de 12 meses, onde obtemos o maior poder explicativo para as variáveis macroeconômicas, os resultados são menos significativos²⁰, evidenciando a necessidade da utilização de uma maior quantidade de variáveis para caracterizar a dinâmica das taxas numa economia emergente. O fator macroeconômico que explica a maior parcela das variações é a variação do câmbio nominal, que pode estar representando tanto variações na expectativa de inflação futura (decorrente do *pass-through* de desvalorizações cambiais para a inflação) quanto variações no prêmio de risco externo (o chamado risco-país). Para o horizonte de previsão de 1 mês, as variáveis macroeconômicas explicam em média em torno de 13% da variância, enquanto para horizontes mais longos explicam em média em torno de 38%. Finalmente, a proporção de variância

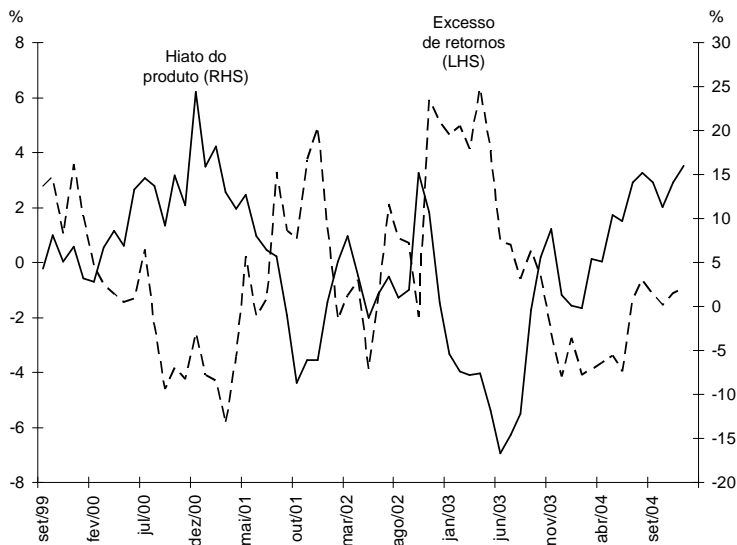
¹⁹Como pode ser visto em Ang e Piazzesi (2003).

²⁰A variância explicada pelos fatores macroeconômicos em Ang e Piazzesi (2003) é de 67%, 79% e 78% para os horizontes de 1, 12 e 60 meses respectivamente.

explicada pelo fator de nível decresce com a maturidade, enquanto a explicada pelo fator de inclinação aumenta com a maturidade.

Podemos também observar a evolução do excesso de retornos esperado (ou prêmio a termo) implícito pelo modelo para diferentes maturidades. A figura 3.21 mostra uma comparação do excesso de retornos esperado para a taxa de 12 meses com o hiato do produto.

Figura 3.21: Excesso de retornos esperado do *swap* de 12 meses e hiato do produto

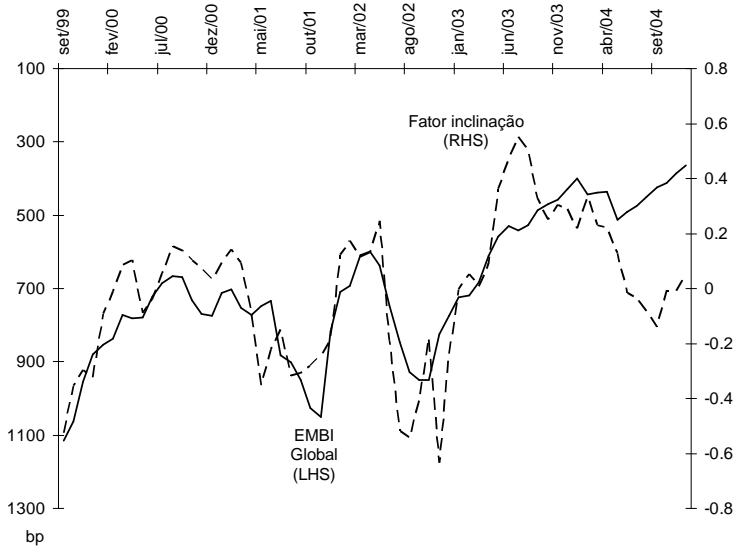


Podemos verificar que, em geral, o prêmio a termo é anti-cíclico. Este resultado já era esperado, já que os agentes requerem uma compensação maior para poupar em períodos de recessão (que é o caso quando o hiato do produto é menor), pois têm o desejo de suavizar o consumo durante toda a vida.

Finalmente, dada a grande parcela da variação das taxas associada aos fatores latentes, gostaríamos de verificar quais variáveis adicionais poderiam potencialmente explicar esses fatores. Como pode ser observado em Megale (2003), o prêmio de risco presente na paridade descoberta da taxa de juros, além de depender dos fundamentos macroeconômicos (que aqui podem ser representados pelo hiato do produto, inflação corrente e variação cambial), depende também da aversão ao risco dos investidores internacionais. Assim, o fator de inclinação deveria estar relacionado com esta variável. De fato, como pode ser observado na figura 3.22, o fator de inclinação tem uma dinâmica muito similar ao índice EMBI Global do JP Morgan, que pode ser utilizado

como proxy para a aversão ao risco global²¹, especialmente até maio de 2004, com uma correlação de 0,70 no período todo (0,83 até maio de 2004). Neste período, somente durante a crise eleitoral de 2002 e em meados de 2003 há um grande deslocamento do fator de inclinação com relação ao EMBI Global.

Figura 3.22: Fator de inclinação e EMBI Global

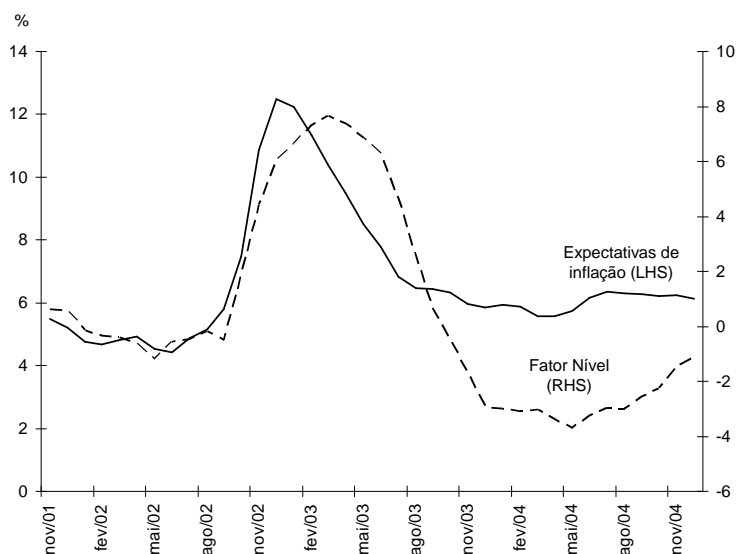


Por outro lado, o fator de nível está intimamente relacionado com a média das taxas. Assim, podemos supor que esteja capturando variações na parcela da taxa de juros relacionada ao seu nível de longo-prazo, ou seja, à taxa de juros real e à meta implícita de inflação percebida pelos agentes (ou às expectativas de inflação 12 meses à frente). Para o período estudado, é razoável supor que a dinâmica do nível de longo-prazo da taxa de juros seja dominada pela variação das expectativas de inflação, já que a taxa de juros real podem ser considerada praticamente constante em curtos períodos de tempo. De fato, como podemos observar na figura 3.23, a dinâmica do fator de nível está muito relacionada às expectativas de inflação²², apresentando correlação de 0,79. A queda mais pronunciada do fator de nível no segundo semestre de 2003 é decorrente do afrouxamento da política monetária e poderíamos explicar esse resultado pela presença de um choque de política monetária (neste caso negativo).

²¹Na realidade, como pode ser observado no gráfico, o EMBI Global tem uma relação com -1^*S , o que não é um problema já que os fatores latentes podem ser rotacionados arbitrariamente.

²²As expectativas de inflação utilizadas são as coletadas pelo Banco Central com diversas instituições (financeiras e não-financeiras), que só começaram a ser divulgadas em novembro de 2001.

Figura 3.23: Fator de nível e expectativas de inflação 12 meses adiante



Podemos concluir portanto que, mesmo com todas as questões ligadas ao mecanismo de transmissão da política monetária no Brasil, a dinâmica da curva de juros é bastante influenciada pelo ambiente macroeconômico, explicando uma parcela significativa da variação das taxas. Vale ressaltar a importância da inclusão de variáveis relacionadas ao setor externo numa pequena economia aberta como o Brasil, intuição esta confirmada pela grande proporção da variação das taxas decorrente de variações no câmbio nominal (até 41%). Além disso, a inclusão de variáveis macroeconômicas melhora a adequação do modelo aos dados, especialmente nos períodos de maior volatilidade, indicando ser esta uma alternativa robusta a modelos com volatilidade estocástica (que são de difícil estimação) para a modelagem de economias sujeitas a grandes instabilidades e altamente vulneráveis a choques.