

## Construção do modelo integrado e o exercício de previsão

O exercício de previsão realizado no presente trabalho será apresentado pelos passos de implementação, referentes aos modelos descritos no capítulo de metodologia.

Devemos atentar que o nosso modelo foi construído a partir do modelo proposto por Diebold e Li, para taxas de juros calculadas a partir dos preços dos *corporate bonds*, também com frequência mensal, e estendido segundo premissas discutidas por Diebold, Li e Yue para diferentes mercados, através de curvas de rendimentos heterogêneas. Buscando satisfazer o princípio KISS (*Keep it sophisticatedly simple*) seguido e defendido por esses autores, optamos pela estimação do modelo por mínimos quadrados ordinários lineares, mantendo o parâmetro  $\lambda$  fixo, mesmo diante da possibilidade da estimação conjunta dos parâmetros via mínimos quadrados ordinários não lineares. Destaca-se que o uso do  $\lambda$  fixo é comum na literatura, uma vez que eleva a eficiência dos estimadores, melhorando as previsões, segundo Diebold e Li (2006).

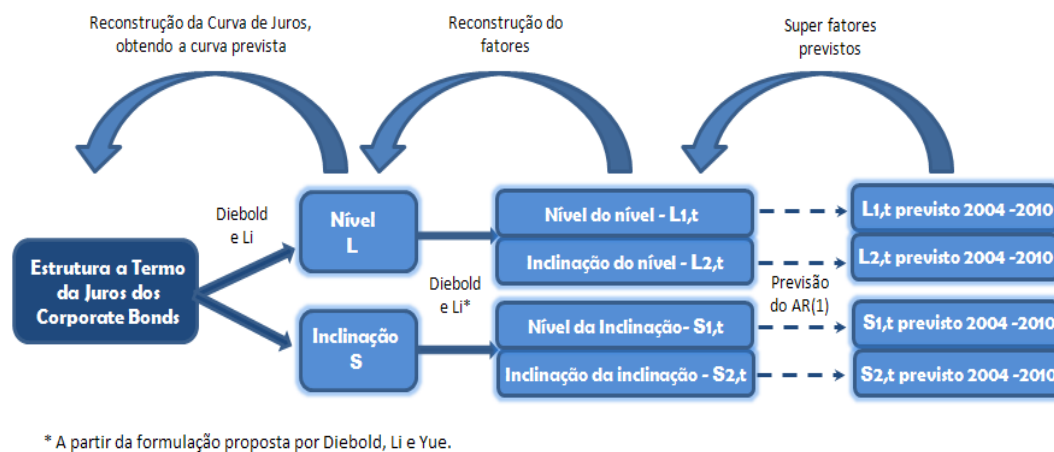
A principal diferença do nosso estudo em relação aos artigos base é a nossa amostra ser composta por títulos corporativos industriais, transacionados no mercado secundário americano e, que por sua vez, apresentam risco de inadimplência. Trabalhamos com títulos de 14 níveis de *ratings*<sup>10</sup> e 15 maturidades<sup>11</sup>. No entanto, destaca-se que o nosso modelo foi construído sobre o mesmo referencial de identificação de forças/fatores que são capazes de governar a dinâmica da ETTJ. Outras inovações presentes neste trabalho são: a estimação de dois super fatores que conduzem a trajetória dos fatores nível e inclinação, enquanto DLY estimam apenas um, e observar a performance preditiva do modelo hierárquico, enquanto esses autores estão centrados na construção e eficiência do modelo para explicar o movimento das curvas de juros dos diferentes países.

O esquema abaixo mostra as etapas realizadas e que serão descritas em cada tópico desse capítulo.

---

<sup>10</sup> Os ratings são AAA, AA, A+, A, A-, BBB+, BBB, BBB-, BB+, BB, BB-, B+, B e B-.

<sup>11</sup> As maturidades são maturidades 3M, 6M, 1A, 2A, 3A, 4A, 5A, 7A, 8A, 9A, 10A, 15A, 20A, 25A e 30A.



**Figura 4.1** Esquema da estrutura do exercício a ser realizado nesse trabalho.

#### 4.1

### Estimação dos fatores comuns a partir da estrutura a termo da taxa de juros para cada *rating*

O primeiro passo consiste na estimação dos fatores comuns presentes na função proposta do Modelo de Nelson e Siegel para descrever a estrutura a termo da taxa de juros. Neste trabalho, assim como proposto por Diebold, Li e Yue (2008), consideramos apenas os fatores  $\beta_{1,t}$  e  $\beta_{2,t}$  representando a curva de juros para as operações com maturidade igual a  $\tau$  períodos, descrita pela seguinte função:

$$y_t^R(\tau) = \beta_{1,t}^R + \beta_{2,t}^R \left( \frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \varepsilon_t^R \quad [4.1]$$

De tal forma que a função  $\left( \frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right)$  é denominada  $F(\tau)$  e  $\varepsilon_t^R$  refere-se ao componente idiossincrático. Essa equação modelará a estrutura a termo para cada nível de *rating* R, assim como realizado por DLY para diferentes países.

A decisão da descrição por apenas dois fatores que explicam a dinâmica da curva de juros é baseada na decomposição por componentes principais da ETTJ, para cada nível de risco. Através dessa técnica conseguimos mostrar que grande parte do comportamento da variável em análise, isto é a sua variabilidade, pode ser explicada por poucos componentes que guiam a sua dinâmica, eliminando os

demais sem perda de generalidade, o que reduz os cálculos necessários para previsões. Conforme demonstrado na Tabela 4.1, nas colunas do acumulado proporcional, em média, 99% do movimento dos juros de títulos corporativos norte-americanos é explicado por dois fatores apenas: nível e inclinação.

**Tabela 4.1** Decomposição das ETTJ por análise de componentes principais para cada *rating*.

ACP - Ratings						
	Variância Proporcional			Acumulado Proporcional		
	Nível	Inclinação	Curvatura	Nível	Inclinação	Curvatura
AAA	90.8%	8.1%	0.5%	90.8%	98.9%	99.4%
AA	90.6%	8.4%	0.5%	90.6%	99.1%	99.6%
A+	91.1%	8.1%	0.5%	91.1%	99.1%	99.6%
A	90.9%	8.3%	0.5%	90.9%	99.1%	99.6%
A-	90.0%	9.1%	0.5%	90.0%	99.2%	99.7%
BBB+	89.4%	9.7%	0.6%	89.4%	99.1%	99.6%
BBB	89.3%	9.7%	0.7%	89.3%	98.9%	99.6%
BBB-	88.7%	10.2%	0.7%	88.7%	98.9%	99.6%
BB+	88.8%	10.1%	0.7%	88.8%	98.9%	99.7%
BB	91.8%	7.2%	0.7%	91.8%	98.9%	99.6%
BB-	92.6%	6.4%	0.6%	92.6%	99.0%	99.7%
B+	94.0%	5.1%	0.6%	94.0%	99.1%	99.7%
B	94.3%	4.5%	0.8%	94.3%	98.9%	99.7%
B-	95.0%	3.6%	1.2%	95.0%	98.6%	99.8%

Nosso argumento é reforçado uma vez que, essa hipótese foi assumida por Diebold, Li e Yue como descrito no tópico 2.3.

Para a estimação desses fatores nível e inclinação, indicados pela Análise de Componentes Principais como os mais relevantes ( $\beta_{1,t}$  e  $\beta_{2,t}$ ), foi utilizado o método de Mínimos Quadrados Ordinários, como sugerido por Diebold e Li (2006), da seguinte forma:

$$Y_t^R(\tau) = AX_t^R + E_t^R \therefore \begin{bmatrix} y_t(3) \\ y_t(6) \\ \vdots \\ y_t(30) \end{bmatrix} = \begin{bmatrix} 1 & F(3) \\ 1 & F(6) \\ \vdots & \vdots \\ 1 & F(20) \end{bmatrix} \begin{bmatrix} \beta_{1,t} \\ \beta_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_t(3) \\ \varepsilon_t(6) \\ \vdots \\ \varepsilon_t(30) \end{bmatrix} \quad [4.2]$$

$$\hat{X} = (A'A)^{-1} A'Y$$

Conforme descrito na metodologia, associamos esses *loadings* aos fatores nível e inclinação, identificados nos trabalhos de Litterman e Scheinkman (1991) e Knez, Litterman e Scheinkman (1994). Desta forma, em nosso modelo nos referiremos a  $\beta_{1,t}$  como nível e  $\beta_{2,t}$  como inclinação. No processo de estimação,

para a construção da função  $F(\tau)$ , usamos  $\lambda = 0.0609$ , invariante com o tempo, também conforme utilizado por DL e a variável  $\tau$  assumindo os valores das maturidades 3M, 6M, 1A, 2A, 3A, 4A, 5A, 7A, 8A, 9A, 10A, 15A, 20A, 25A e 30A em meses<sup>12</sup>. Como a base de dados é composta por estruturas a termo de 14 níveis de *rating* diferentes, estimamos os fatores nível e inclinação associados a cada classificação. Destaca-se que esses fatores são variantes no tempo.

Como citado no tópico 2.2, relacionaremos as recomendações dos autores com as premissas utilizadas em nosso modelo:

- Modelo com três fatores comuns sendo utilizado para ajustar a curva de juros, optamos por apenas dois fatores, como descrito em [4.1] devido, a boa capacidade de explicação conforme indicado anteriormente.

- O parâmetro  $\lambda_t$  é considerado constante e calibrado de tal forma que a função  $G(\tau)$  do *loading*  $\beta_{3,t}$  atinja seu máximo na maturidade média. Em nosso caso, utilizamos o  $\lambda$  calibrado por DL, para ajustarmos a curva de juros a apenas dois fatores, assim como Diebold, Li e Yue (2008), sem perda de generalidade.

- Os fatores  $\beta_{1,t}$ ,  $\beta_{2,t}$  e  $\beta_{3,t}$  sejam estimados por mínimos quadrados ordinários (opção factível porque  $\lambda_t$  é igual a um número definido a priori, o que faz com que a função  $F(\tau)$  seja perfeitamente conhecida). Também utilizamos esse método de estimação, no entanto nos detivemos à estimação de  $\beta_{1,t}$ ,  $\beta_{2,t}$  apenas, como descrito anteriormente.

- As séries temporais correspondentes aos fatores sejam modeladas por processos estocásticos univariados. Supusemos que a lei de movimento de  $\beta_{i,t}$  seja a de um processo AR(1) convencional, ou seja:

$$\beta_{i,t+1} = c_i + \gamma_i \beta_{i,t} + \varepsilon_{i,t+1}, \quad i = 1, 2, \dots, I$$

A forma proposta por DL e descrita pela equação 2.3, pode ser vista como uma maneira mais livre de gerar as previsões. A suposição dos *loadings* seguirem um processo AR(1) será utilizada mais adiante, já que no presente trabalho assumimos a presença de super fatores que coordenam o movimento dos fatores globais, nível e inclinação.

<sup>12</sup> Os valores assumidos pela variável T são 3, 6, 12, 24, 36, 48, 60, 84, 96, 108, 120, 180, 240, 300 e 360.

Para facilitar o entendimento, nos referiremos aos fatores nível e inclinação para a descrição do exercício e das equações não mais como  $\beta_{1,t}$  e  $\beta_{2,t}$  e sim, como  $L_t$  e  $S_t$ .

Podemos observar o comportamento dos fatores  $L_t$  e  $S_t$  estimados conforme descrição anterior, nos Gráficos 4.1 e 4.2. Notamos nas séries de nível estimadas que em períodos de maior estabilidade econômica, há um fechamento entre as curvas de diferentes níveis de risco, enquanto em períodos caracterizados por crises, ocorrem afastamentos. Destaca-se, também, que a amplitude da variável nível é maior para as curvas mais arriscadas, apresentando maior volatilidade, isto é, uma medida governamental que afete uma variável de nível, isto é, aquela que afeta todas as maturidades da curva simultaneamente, possuem um impacto maior nas curvas com classificações piores.

No que se refere às séries de inclinação estimadas, essas se apresentam mais voláteis, fato esse por característica inerente ao fator, que representa os choques que afetam diferentemente a ponta longa (maturidades mais longas) e a ponta curta (maturidades mais curtas). Observamos que diante de períodos de crise como a Bolha da Internet em 2001 e a crise econômica de 2008/2009, a diferença entre as maturidades mais longas e mais curtas, que determinam a inclinação, ficaram maiores em módulo, resultando em fatores cada vez mais negativos, enquanto que em períodos de maior estabilidade econômica, possuem um comportamento mais regular diferenciando-se pelo nível de risco. Podemos perceber, em nossa amostra, que diante de choques, a resposta do fator inclinação, em períodos de crise, se revela em movimentos conjuntos de todos os *ratings*, e em período mais estáveis a diferença deve se concentrar principalmente no diferencial de risco a ser assumido pelos agentes.

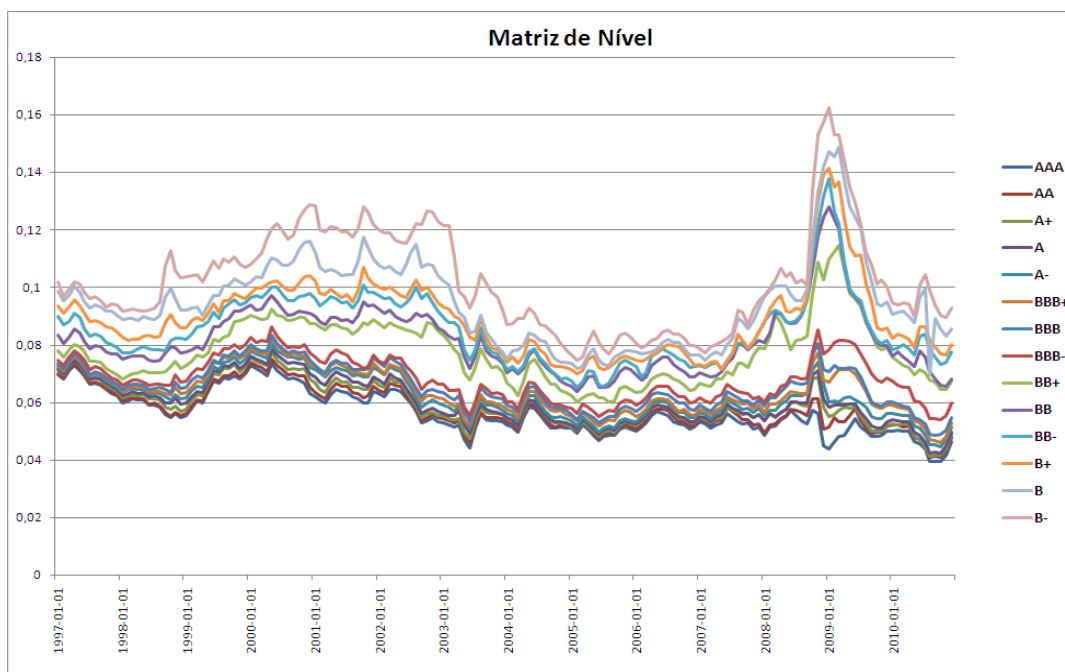


Gráfico 4.1 Fator nível estimado para cada rating.

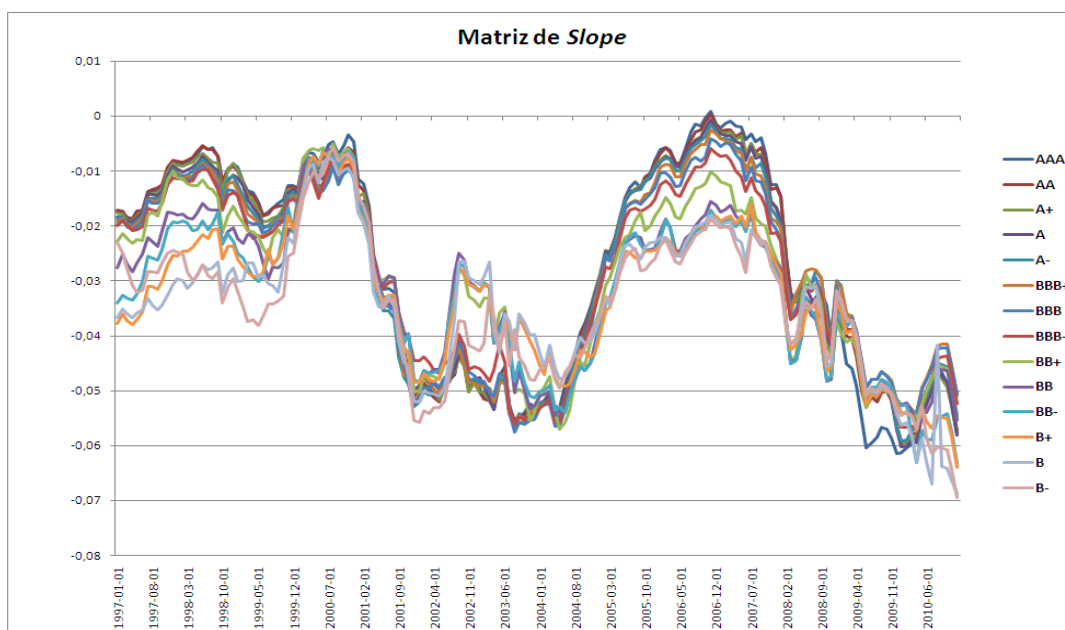


Gráfico 4.2 Fator inclinação estimado para cada rating.

Através da análise das estatísticas descritivas dos fatores nível e inclinação, podemos perceber que uma maior variabilidade no caso do nível, está associada aos piores ratings, podendo ser justificada pela maior fuga dos títulos de pior classificação que diante de choques que afetam todas as maturidades. Enquanto para o fator inclinação, se concentra nas melhores classificações, uma vez que essa componente representa, realmente, a variabilidade dos títulos menos arriscados. A curva média é crescente, no sentido dos títulos mais arriscados, para

o nível, e para a inclinação é decrescente no mesmo sentido, se tornando mais negativa. Quanto às medidas de assimetria, o fator inclinação apresenta-se assimétrico a esquerda para todos os níveis de risco se diferenciando pouco entre eles, o coeficiente está, em torno, de -0,2, com exceção dos títulos classificados como B+,B e B-. Para o nível, as séries são assimétricas a direita, com os valores variando entre si, sendo que a série relativa ao *rating* B+ possui maior coeficiente, igual a 1,3. Em relação a curtose, para o fator denotado por  $L_t$ , essa medida é menor do que 3 para os títulos melhor classificados, e para os títulos a partir do BB+, decrescendo na classificação, esse coeficiente é maior que 3. Para o fator  $S_t$ , essa medida está em torno de 1,5 para quase todos os ratings e a partir da classificação BB+ vai aumentando, atingindo o máximo igual a 3 para o rating B quando volta a cair.

Para a análise da normalidade das séries realizamos o Teste Jarque-Bera, onde percebemos que para todas as classificações de risco, com exceção dos títulos B+,B e B-, rejeitamos a hipótese de normalidade no intervalo de 95% de confiança para o fator inclinação. Enquanto que para o nível a normalidade é rejeitada para todos os *ratings* no intervalo de confiança de 90%, como pode ser observado nas Tabelas 4.2 e 4.3.

**Tabela 4.2** Estatísticas descritivas do fator nível.

Estatísticas Descritivas - Fator Nível														
	AAA	AA	A+	A	A-	BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-
Média	0.056	0.058	0.059	0.060	0.062	0.064	0.065	0.068	0.077	0.082	0.085	0.089	0.095	0.102
Mediana	0.055	0.056	0.058	0.059	0.060	0.063	0.065	0.067	0.075	0.080	0.083	0.088	0.093	0.100
Máximo	0.074	0.075	0.077	0.079	0.080	0.082	0.084	0.086	0.115	0.128	0.138	0.142	0.149	0.162
Mínimo	0.039	0.041	0.041	0.042	0.044	0.046	0.049	0.054	0.060	0.064	0.065	0.070	0.070	0.075
Desvio Padrão	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.011	0.012	0.012	0.014	0.016	0.018
Assimetria	0.345	0.358	0.276	0.272	0.265	0.181	0.142	0.263	0.795	1.258	1.155	1.293	0.946	0.811
Curtose	2.437	2.456	2.397	2.319	2.215	2.201	2.191	2.069	3.609	5.614	5.447	5.627	4.327	3.580
Teste Jarque-Bera	5.546	5.653	4.682	5.312	6.287	5.392	5.151	8.010	20.294	92.092	79.228	95.099	37.397	20.751
p-valor	0.062	0.059	0.096	0.070	0.043	0.067	0.076	0.018	0.000	0.000	0.000	0.000	0.000	0.000

**Tabela 4.3** Estatísticas descritivas do fator inclinação.

Estatísticas Descritivas - Fator Inclinação														
	AAA	AA	A+	A	A-	BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-
Média	-0.028	-0.028	-0.028	-0.028	-0.028	-0.028	-0.029	-0.029	-0.030	-0.031	-0.033	-0.033	-0.034	-0.035
Mediana	-0.021	-0.020	-0.021	-0.021	-0.022	-0.023	-0.023	-0.024	-0.027	-0.028	-0.030	-0.031	-0.032	-0.033
Máximo	0.001	0.000	-0.001	-0.001	-0.001	-0.003	-0.004	-0.006	-0.005	-0.006	-0.007	-0.007	-0.008	-0.006
Mínimo	-0.062	-0.060	-0.060	-0.060	-0.059	-0.057	-0.058	-0.057	-0.057	-0.057	-0.063	-0.064	-0.069	-0.070
Desvio Padrão	0.020	0.019	0.019	0.018	0.018	0.017	0.017	0.016	0.015	0.014	0.014	0.013	0.013	0.013
Assimetria	-0.227	-0.201	-0.197	-0.213	-0.233	-0.219	-0.223	-0.259	-0.217	-0.173	-0.199	-0.097	-0.398	-0.126
Curtose	1.470	1.416	1.427	1.467	1.471	1.480	1.481	1.556	1.694	1.874	2.121	2.293	3.001	2.666
Teste Jarque-Bera	17.832	18.698	18.409	17.718	17.897	17.509	17.547	16.478	13.253	9.703	6.511	3.764	4.429	1.227
p-valor	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.008	0.039	0.152	0.109	0.541

Pela análise da matriz de autocorrelação, tanto do nível quanto da inclinação, nas Tabelas 4.4 e 4.5, notamos que os *ratings* vizinhos são mais correlacionados. Destaca-se que para o fator  $S_t$ , a correlação é bastante persistente entre os diferentes níveis de risco, o que não é válido para o  $L_t$ .

**Tabela 4.4** Matriz de autocorrelação fator nível.

LEVEL	AAA	AA	A_PLUS	A	A_MINUS	BBB_PLUS	BBB	BBB_MINUS	BB_PLUS	BB	BB_MINUS	B_PLUS	B	B_MINUS
AAA	1,000	0,989	0,968	0,954	0,938	0,877	0,857	0,739	0,343	0,316	0,332	0,279	0,277	0,260
AA	0,989	1,000	0,990	0,983	0,971	0,930	0,915	0,816	0,451	0,432	0,443	0,393	0,385	0,365
A_PLUS	0,968	0,990	1,000	0,996	0,988	0,956	0,947	0,852	0,523	0,513	0,522	0,473	0,458	0,439
A	0,954	0,983	0,996	1,000	0,995	0,969	0,961	0,875	0,561	0,552	0,565	0,518	0,503	0,491
A_MINUS	0,938	0,971	0,988	0,995	1,000	0,977	0,972	0,895	0,606	0,594	0,607	0,560	0,547	0,540
BBB_PLUS	0,877	0,930	0,956	0,969	0,977	1,000	0,995	0,962	0,722	0,702	0,701	0,676	0,667	0,646
BBB	0,857	0,915	0,947	0,961	0,972	0,995	1,000	0,964	0,748	0,735	0,733	0,704	0,688	0,672
BBB_MINUS	0,739	0,816	0,852	0,875	0,895	0,962	0,964	1,000	0,838	0,810	0,806	0,798	0,790	0,760
BB_PLUS	0,343	0,451	0,523	0,561	0,606	0,722	0,748	0,838	1,000	0,974	0,961	0,961	0,952	0,936
BB	0,316	0,432	0,513	0,552	0,594	0,702	0,735	0,810	0,974	1,000	0,984	0,977	0,950	0,930
BB_MINUS	0,332	0,443	0,522	0,565	0,607	0,701	0,733	0,806	0,961	0,984	1,000	0,980	0,957	0,948
B_PLUS	0,279	0,393	0,473	0,518	0,560	0,676	0,704	0,798	0,961	0,977	0,980	1,000	0,981	0,960
B	0,277	0,385	0,458	0,503	0,547	0,667	0,688	0,790	0,952	0,950	0,957	0,981	1,000	0,963
B_MINUS	0,260	0,365	0,439	0,491	0,540	0,646	0,672	0,760	0,936	0,930	0,948	0,960	0,963	1,000

**Tabela 4.5** Matriz de autocorrelação fator inclinação.

SLOPE	AAA	AA	A_PLUS	A	A_MINUS	BBB_PLUS	BBB	BBB_MINUS	BB_PLUS	BB	BB_MINUS	B_PLUS	B	B_MINUS
AAA	1.000	0.996	0.994	0.992	0.989	0.984	0.984	0.988	0.942	0.908	0.890	0.873	0.822	0.869
AA	0.996	1.000	0.999	0.998	0.995	0.992	0.992	0.990	0.941	0.906	0.890	0.866	0.816	0.869
A_PLUS	0.994	0.999	1.000	0.999	0.995	0.992	0.993	0.991	0.943	0.908	0.892	0.866	0.815	0.868
A	0.992	0.998	0.999	1.000	0.997	0.993	0.993	0.990	0.941	0.906	0.892	0.865	0.819	0.868
A_MINUS	0.989	0.995	0.995	0.997	1.000	0.997	0.996	0.989	0.940	0.902	0.885	0.857	0.813	0.864
BBB_PLUS	0.984	0.992	0.992	0.993	0.997	1.000	0.999	0.990	0.941	0.902	0.879	0.848	0.796	0.855
BBB	0.984	0.992	0.993	0.993	0.996	0.999	1.000	0.992	0.947	0.911	0.888	0.854	0.800	0.860
BBB_MINUS	0.988	0.990	0.991	0.990	0.989	0.990	0.992	1.000	0.961	0.931	0.911	0.875	0.818	0.876
BB_PLUS	0.942	0.941	0.943	0.941	0.940	0.941	0.947	0.961	1.000	0.986	0.967	0.932	0.880	0.909
BB	0.908	0.906	0.908	0.906	0.902	0.902	0.911	0.931	0.986	1.000	0.985	0.956	0.907	0.918
BB_MINUS	0.890	0.890	0.892	0.892	0.885	0.879	0.888	0.911	0.967	0.985	1.000	0.981	0.941	0.937
B_PLUS	0.873	0.866	0.866	0.865	0.857	0.848	0.854	0.875	0.932	0.956	0.981	1.000	0.969	0.941
B	0.822	0.816	0.815	0.819	0.813	0.796	0.800	0.818	0.880	0.907	0.941	0.969	1.000	0.943
B_MINUS	0.869	0.869	0.868	0.868	0.864	0.855	0.860	0.876	0.909	0.918	0.937	0.941	0.943	1.000

No tópico seguinte, nos concentraremos na descrição do procedimento realizado para a estimação dos super fatores que conduzem a trajetória dos fatores de nível e inclinação estimados nessa etapa.

## 4.2

### Estimação dos super fatores referentes aos fatores globais

O segundo passo consistiu na estimação dos super fatores, referentes aos fatores que governam o movimento da curva de juros. Essa etapa é baseada no trabalho de Diebold, Li e Yue, em que, assim como os autores, definiremos leis de



movimento específicas<sup>13</sup> para os fatores que conduzem a dinâmica da estrutura a termo. No entanto, os últimos, como descrito na seção 2.3, estimam um componente comum (um super fator) e um componente idiossincrático como responsáveis pelas trajetórias dos fatores nível e inclinação ( $L_t$  e  $S_t$  respectivamente), enquanto que em nosso modelo estimamos dois super fatores, além do componente idiossincrático que não nos detivemos na estimação como um processo definido, mas descreve também cada fator. A estimação será explicada neste tópico e consiste em uma inovação introduzida pelo nosso trabalho, até onde estamos cientes. A esses denominaremos:  $L_1, L_2, S_1$  e  $S_2$  onde os super fatores de  $L_t$  se referem ao nível, e os de  $S_t$  se referem à inclinação. Os *loadings* com índice 1 são responsáveis pelos movimentos em nível, isto é, fatores que afetam simultaneamente todas as maturidades da ETTJ, a grosso modo, chamados de “nível do nível” e “nível da inclinação”. Enquanto os *loadings* com índice 2 indicam os movimentos que acontecem de forma diferenciada dependendo da maturidade. São normalmente mais voláteis e afetam majoritariamente a ponta curta, são a grosso modo denominados: “inclinação do nível” e “inclinação da inclinação”.

Para a estimação também utilizamos a função proposta no modelo de Nelson e Siegel. No entanto, com uma diferença, enquanto para a descrição da curva de juros temos como variável dependente a maturidade, denotada pela variável  $\tau$ , para os fatores nível e inclinação temos como variável dependente o *rating*, denotado pela variável  $R$ . Desta forma, o modelo proposto por Nelson e Siegel passa a ser descrito por:

$$L_t(R) = L_{1,t} + L_{2,t} \left( \frac{1 - e^{-\lambda_t R}}{\lambda_t R} \right) + \varepsilon_t^l \quad [4.3]$$

$$S_t(R) = S_{1,t} + S_{2,t} \left( \frac{1 - e^{-\lambda_t R}}{\lambda_t R} \right) + \varepsilon_t^s \quad [4.4]$$

Onde a função  $F(R)$  é denotada por  $\left( \frac{1 - e^{-\lambda_t R}}{\lambda_t R} \right)$  e a variável  $R$  assume uma escala de números em ordem crescente de 1 a 14, associada aos *ratings*, onde 1 se refere ao AAA e conseqüentemente, 14 se refere a B- (a variável  $\tau$  assumia os valores

<sup>13</sup> As equações que descrevem as leis de movimento dos fatores definidas por Diebold, Li e Yue são [2.6] e [2.7].

referentes às maturidades). Nesta etapa, consideramos como a base de dados, duas matrizes construídas a partir do nível estimado e da inclinação estimada, de modo que os *loadings*  $L_1$  e  $L_2$  são obtidos por mínimos quadrados da matriz de nível e os *loadings*  $S_1$  e  $S_2$  obtidos por mínimos quadrados da matriz de inclinação.

A decisão da descrição por apenas dois fatores que explicam a dinâmica de cada fator é baseada na decomposição por componentes principais de cada matriz, e sustentada pelo argumento de DLY, citado anteriormente na seção 2.3. Conforme mostrado nas Tabelas 4.6 e 4.7, tanto para o nível quanto para a inclinação, apenas dois fatores respondem em média por 98% da dinâmica das curvas.

**Tabela 4.6** Decomposição por análise de componentes principais do fator nível.

ACP - Nível			
	$L_1$	$L_2$	$L_3$
Autovalor	10.726	2.974	0.091
Variância Proporcional	76.6%	21.2%	0.7%
Acumulado Proporcional	76.6%	97.9%	98.5%

**Tabela 4.7** Decomposição por análise de componentes principais do fator inclinação.

ACP - Inclinação			
	$S_1$	$S_2$	$S_3$
Autovalor	13.043	0.688	0.132
Variância Proporcional	93.16%	4.91%	0.94%
Acumulado Proporcional	93.16%	98.08%	99.02%

O método de Mínimos Quadrados Ordinários foi utilizado da seguinte forma:

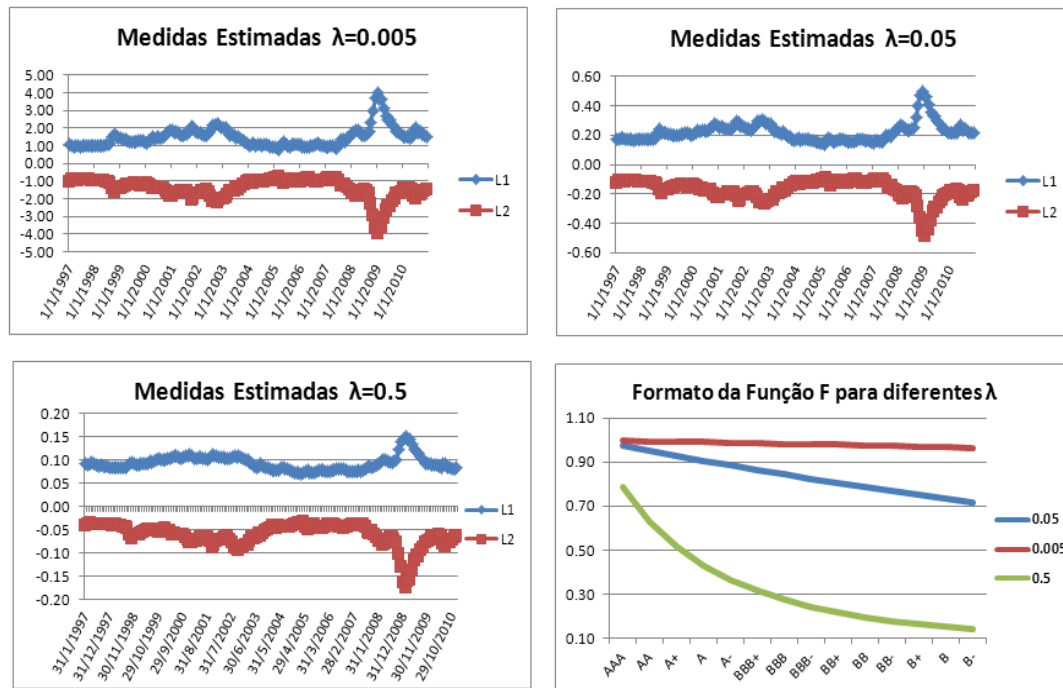
$$L_t(R) = AM_t + E_t^l \therefore \begin{bmatrix} L_t^{AAA} \\ L_t^{AA} \\ \vdots \\ L_t^{B-} \end{bmatrix} = \begin{bmatrix} 1 & F(AAA) \\ 1 & F(A) \\ \vdots & \vdots \\ 1 & F(B-) \end{bmatrix} \begin{bmatrix} L_{1,t} \\ L_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^l(AAA) \\ \varepsilon_t^l(AA) \\ \vdots \\ \varepsilon_t^l(B-) \end{bmatrix} \quad [4.5]$$

$$\hat{M} = (A'A)^{-1} A'L$$

$$S_t(R) = AM_t + E_t^s \therefore \begin{bmatrix} S_t^{AAA} \\ S_t^{AA} \\ \vdots \\ S_t^{B-} \end{bmatrix} = \begin{bmatrix} 1 & F(AAA) \\ 1 & F(A) \\ \vdots & \vdots \\ 1 & F(B-) \end{bmatrix} \begin{bmatrix} S_{1,t} \\ S_{2,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_t^s(AAA) \\ \varepsilon_t^s(AA) \\ \vdots \\ \varepsilon_t^s(B-) \end{bmatrix} \quad [4.6]$$

$$\hat{M} = (A'A)^{-1} A'S$$

Para a construção da função  $F(R)$ , o valor assumido pela variável  $R$  é a escala de números em ordem crescente de 1 a 14, como descrito anteriormente. Essa ordena os *ratings* em termos da possibilidade de *default*, de modo que 1 é o AAA com a menor probabilidade de *default*, enquanto 14 é o título classificado como B-, mais arriscado entre nossos dados. Essa escolha foi arbitrária, de modo que outras estruturas poderiam ser definidas para a variável. Não nos detivemos na elaboração mais complexa por acreditarmos que não seria um fator de forte impacto no objetivo central deste trabalho, a projeção das ETTJ. A escolha do  $\lambda$ , trata-se de um processo complexo devido à presença de multicolinearidade, que foi detectada nas séries dos super fatores, ao tentar selecionar esse parâmetro buscando a minimização do erro entre o fator estimado e o realizado. O comportamento de  $L_{1,t}$ ,  $L_{2,t}$  e o formato da função  $F(R)$  podem ser observados na Figura 4.2. Os  $\lambda$ 's menores que geram erros reduzidos também geram super fatores do nível estimado altamente correlacionados - série em que foram testados os diferentes  $\lambda$  - em torno de 99%, como pode ser verificado na Tabela 4.8. Desta forma, o parâmetro  $\lambda$  foi selecionado buscando maior aderência às medidas de mercado e menor autocorrelação entre os super fatores, sendo assim, diferentes na estimação dos *loadings* referentes ao nível e à inclinação. As medidas de mercado são: a média dos níveis que foi comparada ao  $L_{1,t}$ , a média das inclinações que foi comparada ao  $S_{1,t}$  e a medida construída a partir da diferença entre a taxa mais arriscada e a menos arriscada, tanto para o nível, quanto para a inclinação, que foram comparadas respectivamente com  $L_{2,t}$  e  $S_{2,t}$ . Essas foram construídas a partir das matrizes geradas pela estimação dos fatores  $L_t$  e  $S_t$ : a matriz de nível e a matriz de inclinação, cada uma com 14 colunas, e em cada coluna o fator referente a cada *rating*, as mesmas matrizes a partir dos quais os super fatores foram estimados.



**Figura 4.2.** Comportamento dos super fatores do nível estimados para diferentes  $\lambda$ 's e o formato da função F.

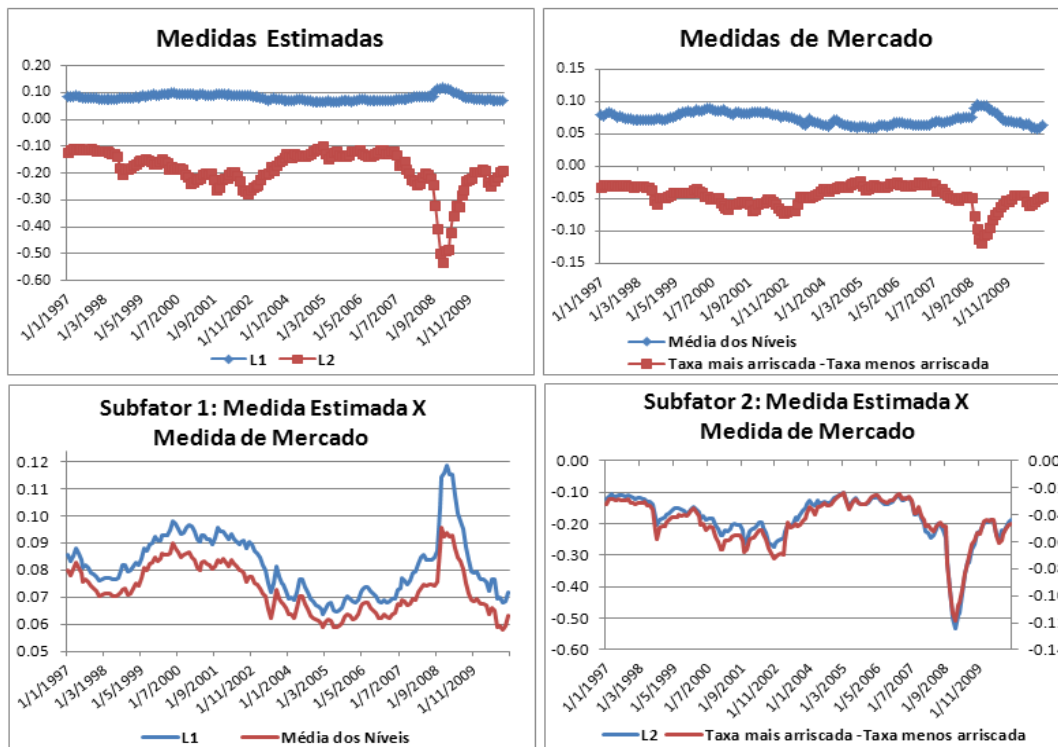
**Tabela 4.8** Autocorrelação das medidas de nível estimadas para diferentes  $\lambda$ 's.

Autocorrelação entre subfatores	
$\lambda = 0.005$	-99.99%
$\lambda = 0.05$	-99.30%
$\lambda = 0.5$	-86.97%

Para o cálculo dos super fatores, referente ao nível, foi selecionado o  $\lambda = 5$ , que possui uma aderência com as medidas de mercado de 97% individualmente, e autocorrelação negativa de 75%, como pode ser observado na Tabela 4.9. As medidas estimadas e de mercado podem ser observadas nos gráficos que compõem a Figura 4.3.

**Tabela 4.9** Autocorrelação das medidas estimadas e medidas de mercado para os super fatores do nível.

CORRELAÇÃO			
Medidas Estimadas (L1,L2)	-75%	(L1, Média dos Níveis)	97%
Medidas de Mercado ( Média dos Níveis, Taxa mais arriscada -Taxa menos arriscada)	-62%	( L2, Taxa mais arriscada -Taxa menos arriscada)	97%

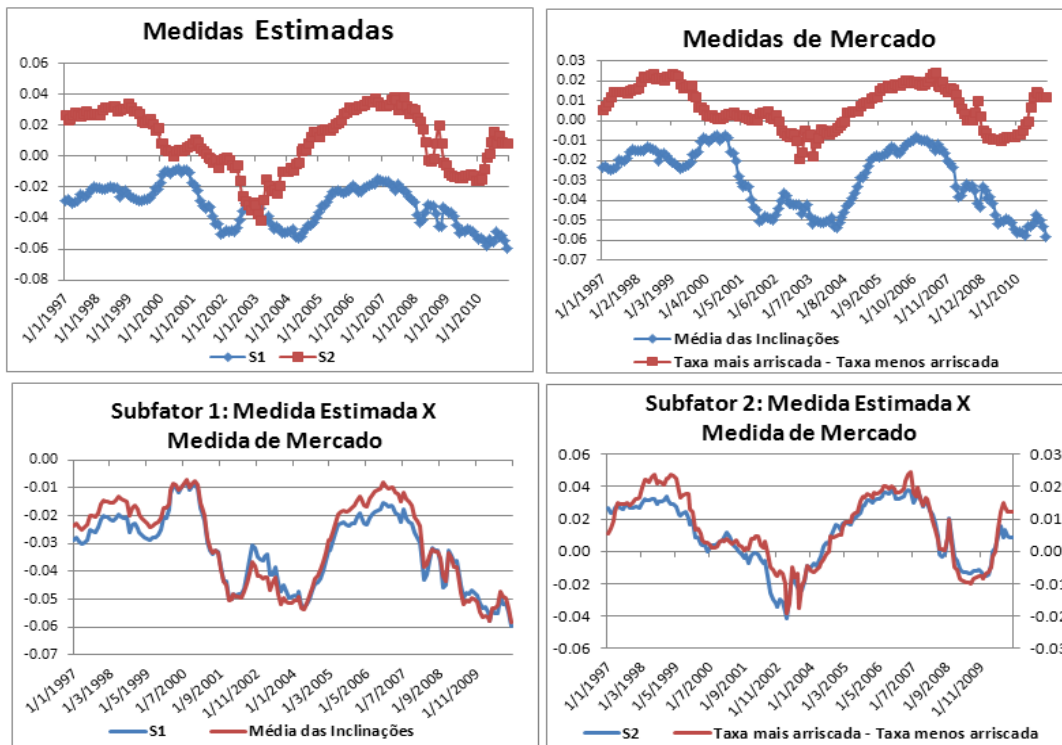


**Figura 4.3** Medidas estimadas e medidas de mercado para os super fatores do nível.

No caso da estimação dos super fatores relativos à inclinação foi selecionado o  $\lambda = 1$ , que possui uma aderência com as medidas de mercado de 98% para  $S_{1,t}$  e 94% para  $S_{2,t}$ , enquanto autocorrelação entre  $S_{1,t}$  e  $S_{2,t}$  é de 56%, como pode ser observado na Tabela 4.10. As medidas estimadas e de mercado podem ser observadas nos gráficos que compõem a Figura 4.4.

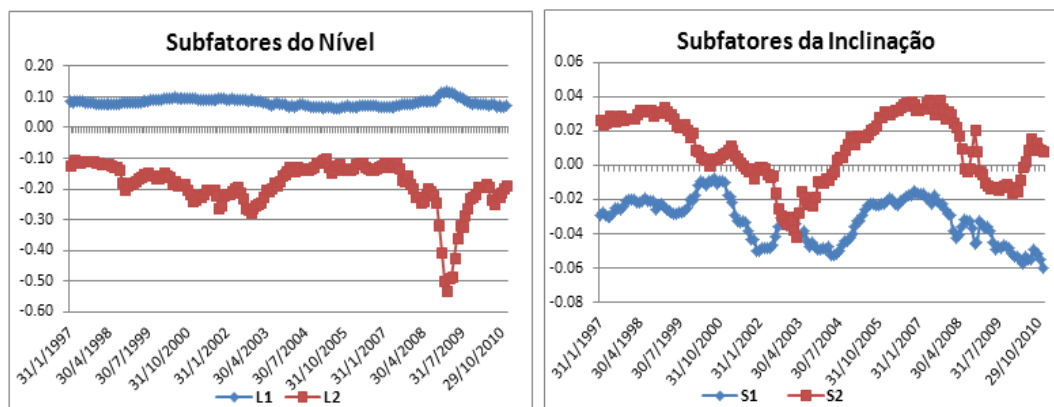
**Tabela 4.10** Autocorrelação das medidas estimadas e medidas de mercado para os super fatores da inclinação.

CORRELAÇÃO			
Medidas Estimadas (S1,S2)	56%	(S1, Média das Inclinações)	98%
Medidas de Mercado ( Média das Inclinações, Taxa mais arriscada -Taxa menos arriscada)	70%	( S2, Taxa mais arriscada -Taxa menos arriscada)	94%



**Figura 4.4** Medidas estimadas e medidas de mercado para os super fatores da inclinação.

O comportamento dos super fatores estimados pode ser observado na Figura 4.5.



**Figura 4.5** Comportamento dos Super Fatores Estimados

Pela análise das estatísticas descritivas apresentadas na Tabela 4.11, podemos perceber maior variabilidade em  $L_{2,t}$ , enquanto as demais séries possuem a mesma variação. Para as medidas de assimetria, a série  $L_{2,t}$  é a que apresenta menor simetria, com coeficiente igual a -2,08, enquanto os *loadings*  $S_{1,t}$  e  $S_{2,t}$  são -0,18 e -0,47, respectivamente, levemente assimétricas a esquerda. A

série de  $L_{1,t}$  apresenta leve assimetria à direita, com coeficiente igual a 0,76. No que refere-se à curtose, essa medida é menor do que 3 para  $S_{1,t}$  e  $S_{2,t}$ , enquanto que para  $L_{1,t}$  e  $L_{2,t}$  é maior que 3, principalmente para  $L_{2,t}$ , onde é igual a 8,83. Realizamos o Teste Jarque-Bera para a análise da normalidade das séries, onde percebemos que para todos os super fatores rejeitamos a hipótese de normalidade no nível de 5%. Destacamos que o comportamento mais discrepante da variável  $L_{2,t}$ , pode ser explicado pela própria definição da variável: inclinação do nível. O nível é o fator que afeta conjuntamente todas as maturidades da estrutura a termo, e pode ser decomposto em duas forças: sendo uma representa os movimentos mais regulares (em nível) e outra os mais voláteis (na inclinação). A última deve apresentar um comportamento mais variante, uma vez que o objetivo é que capture todas as variações de  $L_t$ . O mesmo não acontece com  $S_{1,t}$  e  $S_{2,t}$ , porque são medidas de inclinação, um fator que tem menor representatividade<sup>14</sup> na dinâmica dos juros.

**Tabela 4.11** Estatísticas descritivas dos super fatores.

Estatísticas Descritivas				
	$L_1$	$L_2$	$S_1$	$S_2$
Média	0.08	-0.19	-0.03	0.01
Mediana	0.08	-0.17	-0.03	0.01
Máximo	0.12	-0.10	-0.01	0.04
Mínimo	0.06	-0.53	-0.06	-0.04
Desvio Padrão	0.01	0.08	0.01	0.02
Assimetria	0.76	-2.08	-0.18	-0.47
Curtose	3.53	8.83	1.93	2.28
Teste Jarque-Bera	18.36	359.68	8.89	9.91
p-valor	0.000	0.000	0.012	0.007

### 4.3

#### Previsão dos super fatores

Esta etapa consiste na modelagem dos super fatores como processos auto-regressivos univariados e, sob esse enfoque, a realização da previsão desses processos um, seis e doze passos à frente ( $h=1,6$  e  $12$ ), conforme sugerido por

<sup>14</sup> O fator inclinação responde, em média, por 8% da trajetória da curva de juros como pode ser observado na Tabela 4.1.

Diebold e Li (2006). A diferença do nosso estudo em relação à base teórica referencial é que estaremos projetando os super fatores que conduzem a trajetória dos fatores globais, que por sua vez guiam o movimento da taxa de juros. Repetiremos algumas equações descritas anteriormente para que seja mais clara a visualização. O artigo de Diebold, Li (2006), modela a curva de juros pela seguinte equação:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} + \beta_{3,t} \left( \frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \right) \quad [2.2]$$

Onde  $\beta_{1,t}$ ,  $\beta_{2,t}$  e  $\beta_{3,t}$  são os fatores comuns que descrevem a trajetória da curva de juros. Esses fatores são modelados como processos auto-regressivos univariados, para o qual se realiza a projeção. Em nosso modelo descrevemos a curva de juros como:

$$y_t^R(\tau) = \beta_{1,t}^R + \beta_{2,t}^R \left( \frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \right) + \varepsilon_t^R \quad [4.1]$$

Observa-se que não realizaremos a projeções de  $\beta_{1,t}$  e  $\beta_{2,t}$  como os fatores nível e inclinação, mas sim, a projeção dos super fatores  $L_{1,t}$ ,  $L_{2,t}$ ,  $S_{1,t}$  e  $S_{2,t}$ , citados no tópico 4.2<sup>15</sup>, que governam o movimento desses ( $L_t$  e  $S_t$ ).

Para a projeção de  $L_{1,t}$ ,  $L_{2,t}$ ,  $S_{1,t}$  e  $S_{2,t}$  modelamos esses como meros processos AR(1), que é uma simplificação do modelo auto-regressivo univariado descrito em [2.3], estando de acordo com o desejo dos autores de “... *intentionally impose substantial a priori structure, motivated by simplicity, parsimony, and theory*”. A estrutura desse passo será descrita pelas equações:

$$L_{1,t} = a_1 + b_1 L_{1,t-1} + \varepsilon_t^{l_1} \quad [4.7]$$

$$L_{2,t} = a_2 + b_2 L_{2,t-1} + \varepsilon_t^{l_2} \quad [4.8]$$

$$S_{1,t} = c_1 + d_1 S_{1,t-1} + \varepsilon_t^{s_1} \quad [4.9]$$

$$S_{2,t} = c_2 + d_2 S_{2,t-1} + \varepsilon_t^{s_2} \quad [4.10]$$

A partir dessas equações estimaremos os coeficientes  $a, b, c$  e  $d$  por mínimos quadrados ordinários para cada super fator, utilizando a subamostra referente ao

<sup>15</sup> As leis de movimento dos fatores estão descritas pelas equações [4.3] e [4.4].



período de janeiro de 1997 a dezembro de 2003. Assim, de posse dos coeficientes referentes a toda a subamostra, em dezembro de 2003, realizamos a previsão um, seis e doze passos à frente. As datas de previsão são janeiro, junho e dezembro de 2004 respectivamente, destacando que a projeção é calculada individualmente para  $L_{1,t}$ ,  $L_{2,t}$ ,  $S_{1,t}$  e  $S_{2,t}$ . A construção da série temporal desses fatores projetados é dinâmica, realizando um *looping*, isto é, a cada nova informação inserida são reestimados os coeficientes do modelo auto-regressivo e calculados os fatores projetados para  $h=1,6$  e  $12$ , de forma que a amostra é constantemente atualizada. A estimação dos coeficientes e a projeção incluem a última informação disponível.

Para a previsão de um modelo auto-regressivo de ordem 1<sup>16</sup>, dado na forma geral por:

$$y_{t+1} = a_0 + a_1 y_t + \varepsilon_{t+1} \quad [4.11]$$

A equação de previsão para  $j$  passos a frente é dada por:

$$E_t y_{t+j} = a_0(1 + a_1 + a_1^2 + \dots + a_1^{j-1}) + a_1^j y_t \quad [4.12]$$

Em nosso modelo, as equações para a previsão um, seis e doze passos à frente, são dadas no formato geral, como:

$$h=1 \quad E_t y_{t+1} = a_0 + a_1 y_t \quad [4.13]$$

$$h=6 \quad E_t y_{t+6} = a_0(1 + a_1 + a_1^2 + a_1^3 + a_1^4 + a_1^5) + a_1^6 y_t \quad [4.14]$$

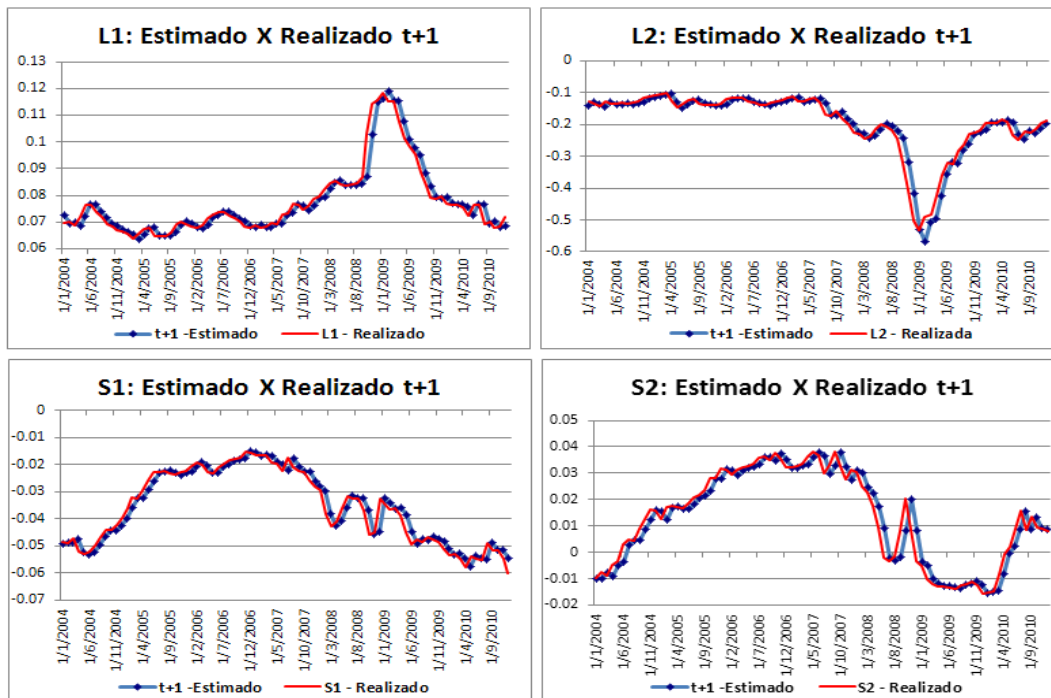
$$h=12 \quad E_t y_{t+12} = a_0(1 + a_1 + a_1^2 + \dots + a_1^{11}) + a_1^{12} y_t \quad [4.15]$$

Nota-se que para a previsão um passo à frente, a última observação computada é novembro de 2010, para  $h=6$  a observação é junho de 2010 e conseqüentemente para  $h=12$  é dezembro de 2009.

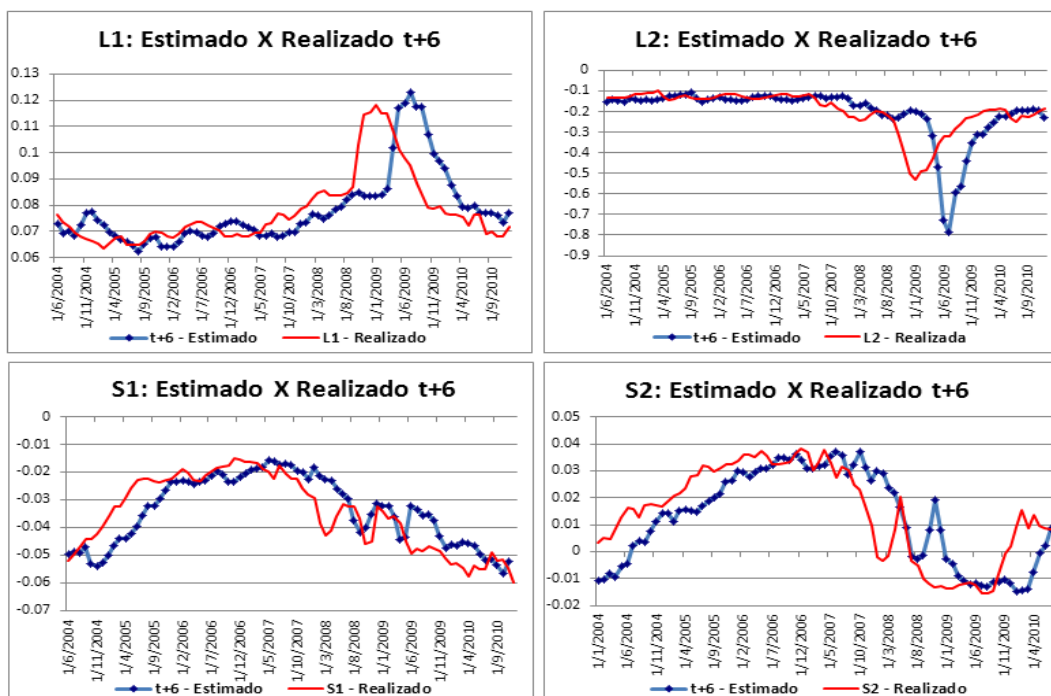
Os resultados das curvas estimadas podem ser observados nos Gráficos 4.3, 4.4 e 4.5, onde podemos perceber que nosso modelo tem um bom poder preditivo para  $h=1$ . No entanto, para  $h=6$  e  $12$  percebemos defasagem na projeção, com o distanciamento maior acontecendo quanto maior o horizonte de previsão,

<sup>16</sup> Ver equações [4.11] e [4.12] e suas derivações, no tópico 9: The Forecast Function, do capítulo 2: Stationary Time-Series Models em ENDERS, W.. Applied Econometric Time Series.

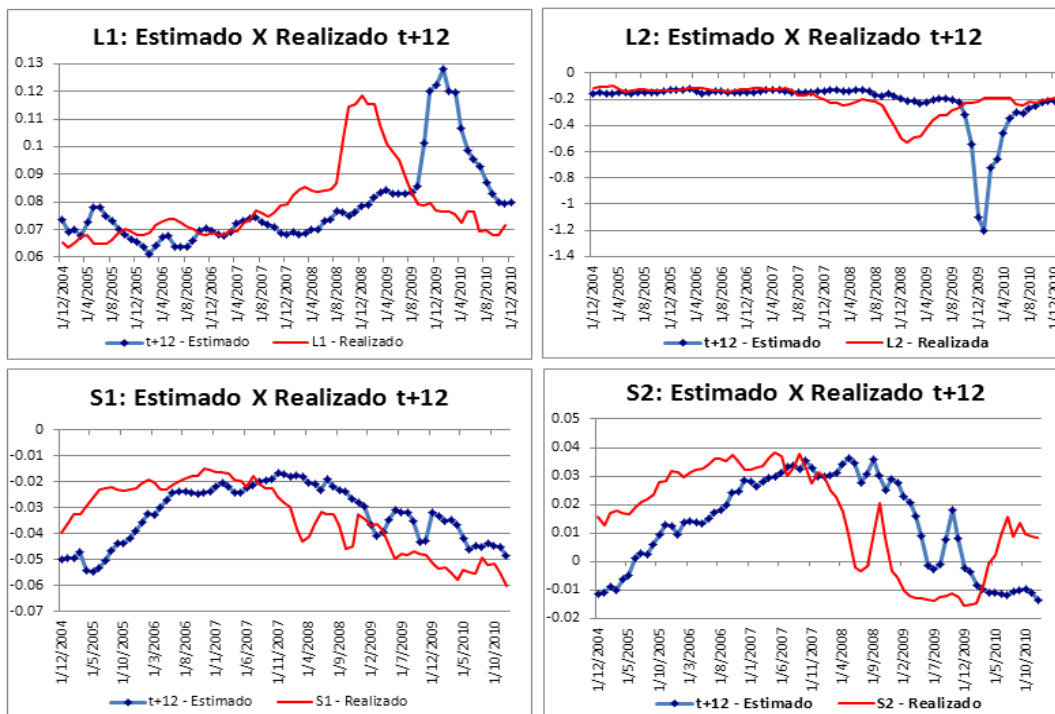
como esperado.



**Figura 4.6** Previsão t+1: Super fatores estimados pelo modelo integrado x realizada no período de jan/2004 a dez/2010.



**Figura 4.7** Previsão t+6: Super fatores estimados pelo modelo integrado x realizada no período de jan/2004 a dez/2010.



**Figura 4.8** Previsão t+12: Super fatores estimados pelo modelo integrado x realizada no período de jan/2004 a dez/2010.

As conclusões acima também podem ser verificadas através das estatísticas dos erros de previsão, nas Tabelas 4.12 a 4.15, onde percebemos que na estimação de todos os super fatores a variabilidade, medida pelo desvio padrão, é maior quanto maior o horizonte de previsão ( $h$ ), como esperado. As séries temporais dos erros associadas às projeções um, seis e doze passos a frente são compostas por 84,79 e 73 observações respectivamente. Na seção 4.2, referente à estimação de  $L_{1,t}$ ,  $L_{2,t}$ ,  $S_{1,t}$  e  $S_{2,t}$ , percebemos que o fator  $L_{2,t}$  apresenta maior variabilidade, curtose e assimetria elevadas. Tal fato é refletido nos erros, devido à maior dificuldade em prever esse comportamento muito volátil. Destaca-se que os erros apresentam uma estrutura de correlação crescente com  $h$ , sendo que os erros mais correlacionados são referentes às medidas de inclinação.

**Tabela 4.12** L1: Estatística dos erros de previsão.

<b>L1</b>						
	média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
<b>t+1</b>	0.0000	0.0033	0.0000	0.4744	-0.0929	-0.0398
<b>t+6</b>	0.0002	0.0126	0.0001	0.9178	-0.0617	-0.1172
<b>t+12</b>	0.0000	0.0184	0.0003	0.9560	0.3578	-0.3918

Tabela 4.13 L2: Estatística dos erros de previsão.

L2						
	média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
t+1	-0.0012	0.0254	0.0006	0.5811	-0.2766	0.0095
t+6	-0.0064	0.1199	0.0136	0.8867	-0.2786	-0.0913
t+12	-0.0191	0.2059	0.0372	0.8822	0.0527	-0.3902

Tabela 4.14 S1: Estatística dos erros de previsão.

S1						
	média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
t+1	0.0002	0.0031	0.0000	0.2374	0.1409	-0.1433
t+6	0.0004	0.0082	0.0001	0.8804	0.4136	0.1202
t+12	0.0011	0.0140	0.0002	0.9449	0.7166	0.4696

Tabela 4.15 S2: Estatística dos erros de previsão.

S2						
	média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
t+1	-0.0003	0.0043	0.0000	0.2867	0.0590	0.0691
t+6	-0.0021	0.0125	0.0002	0.8941	0.2942	0.2151
t+12	-0.0020	0.0191	0.0004	0.9554	0.6742	0.2838

Abaixo pode-se observar um esquema que ilustra as etapas do exercício de forma a resumir o que foi realizado até a presente seção.

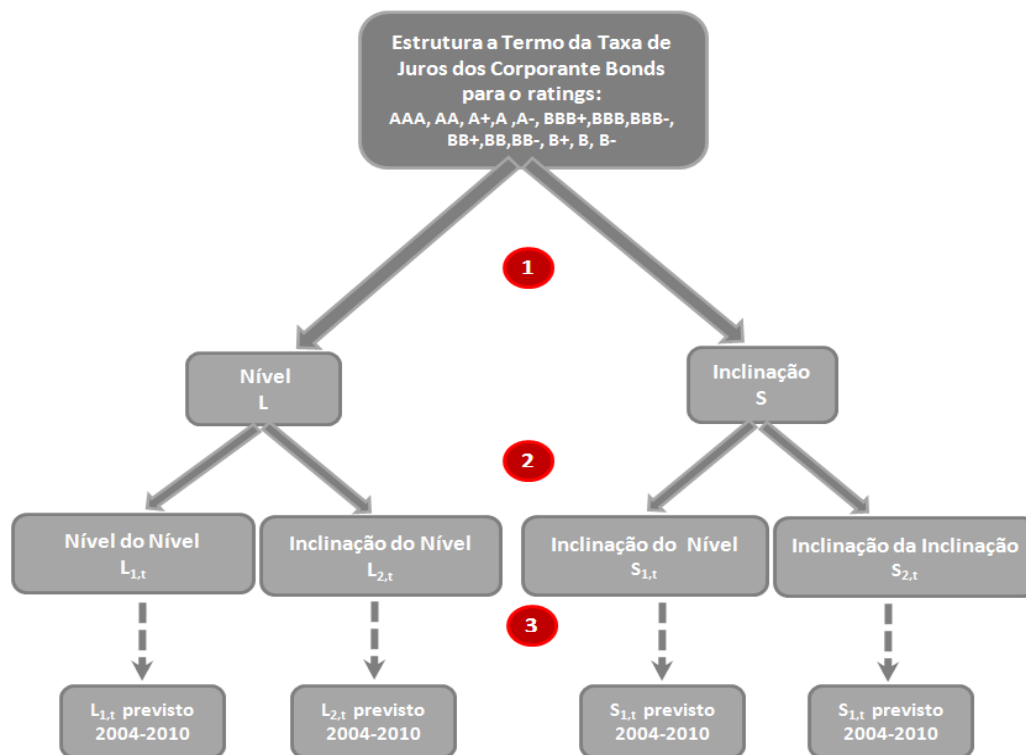


Figura 4.9 Esquema da estrutura do exercício realizado neste trabalho até esta seção.

Etapas:

- 1 Modelagem da curva de juros segundo a formulação de Diebold e Li para a extração dos fatores nível e inclinação. Os fatores são obtidos para cada *rating*. Após essa etapa, montamos uma matriz do nível e uma matriz da inclinação, onde as colunas são as classificações de risco e nas linhas são as datas.
- 2 Extração conjunta dos super fatores, através das matrizes descritas anteriormente, modelando os fatores nível e a inclinação de acordo com a formulação de Diebold, Li e Yue (2008).
- 3 Previsão dos super fatores que são modelados como processos auto-regressivos de ordem 1. A projeção é realizada um, seis e doze passos a frente, a partir de janeiro de 2004 e a estimação dos coeficiente do AR(1) é realizada através de mínimos quadrados ordinários.

Na próxima etapa nos concentraremos na reconstrução dos fatores a partir dos valores projetados neste tópico.

#### 4.4

#### Reconstrução dos fatores a partir da previsão dos super fatores

De posse dos super fatores previstos que conduzem a dinâmica de  $L_t$  e  $S_t$ , conseguimos reconstruir as leis de movimento, descritas na seção 4.2 pelas equações [4.3] a [4.4] e, assim, refizemos a estimação dos fatores no sentido da reconstrução da curva de juros prevista, que será descrita no tópico seguinte. Relembrando a lei da dinâmica dos fatores:

$$L_t(R) = L_{1,t} + L_{2,t} \left( \frac{1 - e^{-\lambda_t R}}{\lambda_t R} \right) + \varepsilon_t^l \quad [4.3]$$

$$S_t(R) = S_{1,t} + S_{2,t} \left( \frac{1 - e^{-\lambda_t R}}{\lambda_t R} \right) + \varepsilon_t^s \quad [4.4]$$

Para a reestimação dos fatores de nível e inclinação referente a cada *rating*, reescrevemos a equação acima pelas matrizes descritas em [4.16] e [4.17], onde

através da multiplicação matricial, obtivemos a série temporal desses fatores previstos para janeiro de 2004 a dezembro de 2010. Ao fim da estimação, obtivemos três séries para cada fator, referentes a cada horizonte de previsão ( $h=1,6$  e  $12$ ).

$$\begin{bmatrix} L_t^{AAA} \\ L_t^{AA} \\ \vdots \\ L_t^{B-} \end{bmatrix} = \begin{bmatrix} 1 & F(AAA) \\ 1 & F(A) \\ \vdots & \vdots \\ 1 & F(B-) \end{bmatrix} \begin{bmatrix} L_{1,t} \\ L_{2,t} \end{bmatrix} \quad [4.16]$$

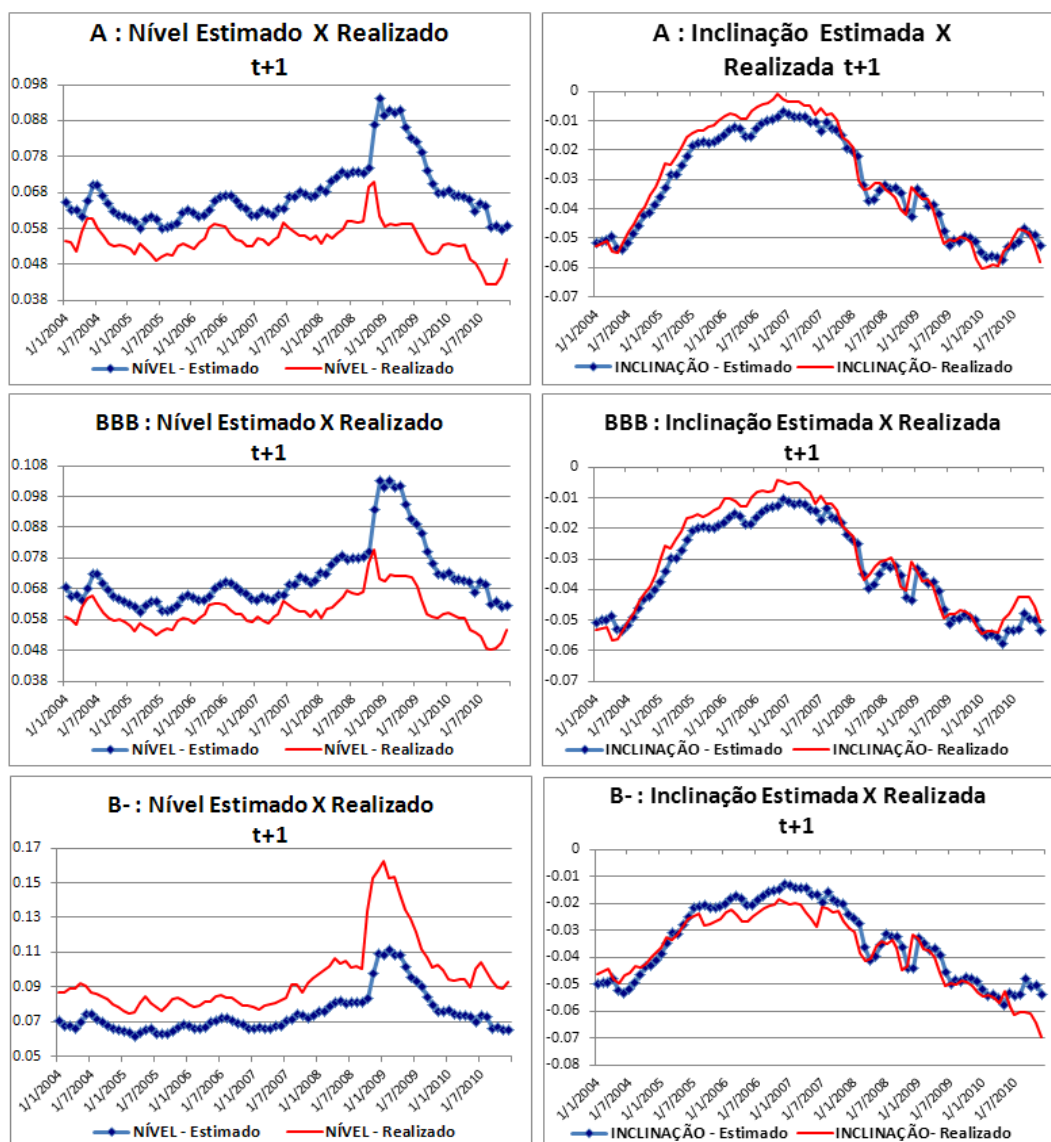
$$\begin{bmatrix} S_t^{AAA} \\ S_t^{AA} \\ \vdots \\ S_t^{B-} \end{bmatrix} = \begin{bmatrix} 1 & F(AAA) \\ 1 & F(A) \\ \vdots & \vdots \\ 1 & F(B-) \end{bmatrix} \begin{bmatrix} S_{1,t} \\ S_{2,t} \end{bmatrix} \quad [4.17]$$

Destaca-se que para o cálculo da função  $F(R)$ , assim como na estimação dos super fatores, utilizamos a variável  $R$  assumindo valores na escala de números em ordem crescente de 1 a 14, associada aos *ratings*, onde 1 se refere ao AAA e conseqüentemente, 14 refere-se ao B-. O parâmetro  $\lambda$ , foi mantido o mesmo quanto da estimação de  $L_{1,t}, L_{2,t}, S_{1,t}$  e  $S_{2,t}$ . Pela própria estrutura do modelo, mantivemos essas premissas já que consiste na reconstrução dos fatores. Sendo assim, para o fator nível foi utilizado  $\lambda = 5$  e para inclinação,  $\lambda = 1$ .

Destaca-se que a realização de cada passo nesse trabalho, se faz assumindo a etapa anterior como realizada. Por exemplo, para a decomposição dos fatores  $L_t$  e  $S_t$  em  $L_{1,t}, L_{2,t}, S_{1,t}$  e  $S_{2,t}$ , assumimos que os fatores nível ( $L_t$ ) e inclinação ( $S_t$ ) fossem os realizados, obtendo assim a série dos super fatores. Esse fato nos permite a comparação dos fatores previstos com os então realizados.

Para a apresentação dos resultados dessa etapa, selecionamos os níveis de risco: A, BBB e B-, de maneira arbitrária, uma vez que a apresentação gráfica de todos os *ratings* se tornaria exaustiva e, não apresenta diferenças consideráveis com relação aos selecionados. Desta forma, podemos perceber através da comparação dos fatores previstos, a partir da projeção dos super fatores, com os fatores realizados que graficamente a projeção um passo à frente para o fator inclinação é mais precisa do que para o nível para as classificações de risco

analisadas. O distanciamento para ambos os fatores se torna mais relevante conforme aumenta o horizonte de previsão. Optamos pela disponibilização dos gráficos referentes à previsão seis e doze passos à frente no Apêndice (Gráficos 7.21 e 7.22), capítulo 7 deste trabalho, para possibilitar maior dinamismo na leitura.



**Figura 4.10** Previsão t+1: Fatores estimados pelo modelo integrado x realizada no período de jan/2004 a dez/2010 para os ratings A, BBB E B-.

Pela análise das estatísticas dos erros apresentadas nas Tabelas 4.16 a 4.18, percebemos que o desvio padrão aumenta com o horizonte de previsão para todos os níveis de risco, como o esperado, sendo levemente maior para o *rating* mais arriscado, B-. Nota-se também, que o valor médio dos erros do nível é maior que

do fator inclinação. O erro médio quadrático da previsão do título mais arriscado é maior do que os demais, possivelmente devido à variabilidade inerente a títulos mais arriscados, enquanto que para os títulos A e BBB o EQM é o mesmo. Destaca-se ainda, que os erros apresentam uma estrutura de correlação significativa, tanto de ordem um, seis, quanto de ordem doze, indicando séries bastante persistentes. Sendo que os erros na previsão do fator inclinação são mais correlacionados.

**Tabela 4.16 A:** Estatística dos erros de previsão dos fatores.

<b>A</b>						
	Média	Desv. Padrão	EMQ	correl(1)	correl(6)	correl(12)
<b>t+1</b>						
Nível	0.013	0.007	0.0002	0.878	0.484	0.318
Inclinação	-0.002	0.004	0.0000	0.654	0.404	0.236
<b>t+6</b>						
Nível	0.013	0.009	0.0002	0.916	0.446	0.325
Inclinação	-0.003	0.011	0.0001	0.940	0.619	0.439
<b>t+12</b>						
Nível	0.013	0.010	0.0003	0.870	0.614	0.168
Inclinação	-0.002	0.019	0.0004	0.975	0.809	0.644

**Tabela 4.17 BBB:** Estatística dos erros de previsão dos fatores.

<b>BBB</b>						
	Média	Desv. Padrão	EMQ	correl(1)	correl(6)	correl(12)
<b>t+1</b>						
Nível	0.011	0.007	0.0002	0.861	0.425	0.326
Inclinação	-0.003	0.004	0.0000	0.607	0.202	-0.087
<b>t+6</b>						
Nível	0.011	0.009	0.0002	0.928	0.362	0.220
Inclinação	-0.004	0.010	0.0001	0.921	0.476	0.219
<b>t+12</b>						
Nível	0.011	0.012	0.0003	0.937	0.583	-0.047
Inclinação	-0.003	0.016	0.0003	0.968	0.773	0.567

**Tabela 4.18 B-:** Estatística dos erros de previsão dos fatores.

<b>B-</b>						
	Média	Desv. Padrão	EMQ	correl(1)	correl(6)	correl(12)
<b>t+1</b>						
Nível	-0.021	0.010	0.0005	0.870	0.451	0.220
Inclinação	0.003	0.005	0.0000	0.693	0.344	-0.039
<b>t+6</b>						
Nível	-0.021	0.017	0.0007	0.915	0.052	-0.110
Inclinação	0.003	0.008	0.0001	0.889	0.459	0.223
<b>t+12</b>						
Nível	-0.022	0.022	0.0009	0.948	0.333	-0.311
Inclinação	0.005	0.014	0.0002	0.951	0.752	0.520



No próximo capítulo, será descrito processo de reestimação da curva de juros.

## 4.5

### Reconstrução da curva de juros e análise da previsão

Esse tópico trata do enfoque preditivo dado ao modelo hierárquico construído baseado nas premissas de Diebold e Li (2006) e Diebold, Li e Yue (2008). Após a estimação dos super fatores, onde ampliamos a modelagem de DLY, realizamos a previsão dos super fatores por um modelo univariado e reconstruímos os fatores, que por sua vez, guiam o movimento da curva de juros. Assim, partimos então, finalmente, para a obtenção da estrutura a termo da taxa de juros prevista para cada *rating*.

A partir da obtenção dos fatores nível e inclinação, para cada nível de risco na secção 4.4, pela própria estrutura do modelo passamos a deter as séries que compõem a modelagem da curva de juros, que é dada pela equação [4.1], reescrita abaixo.

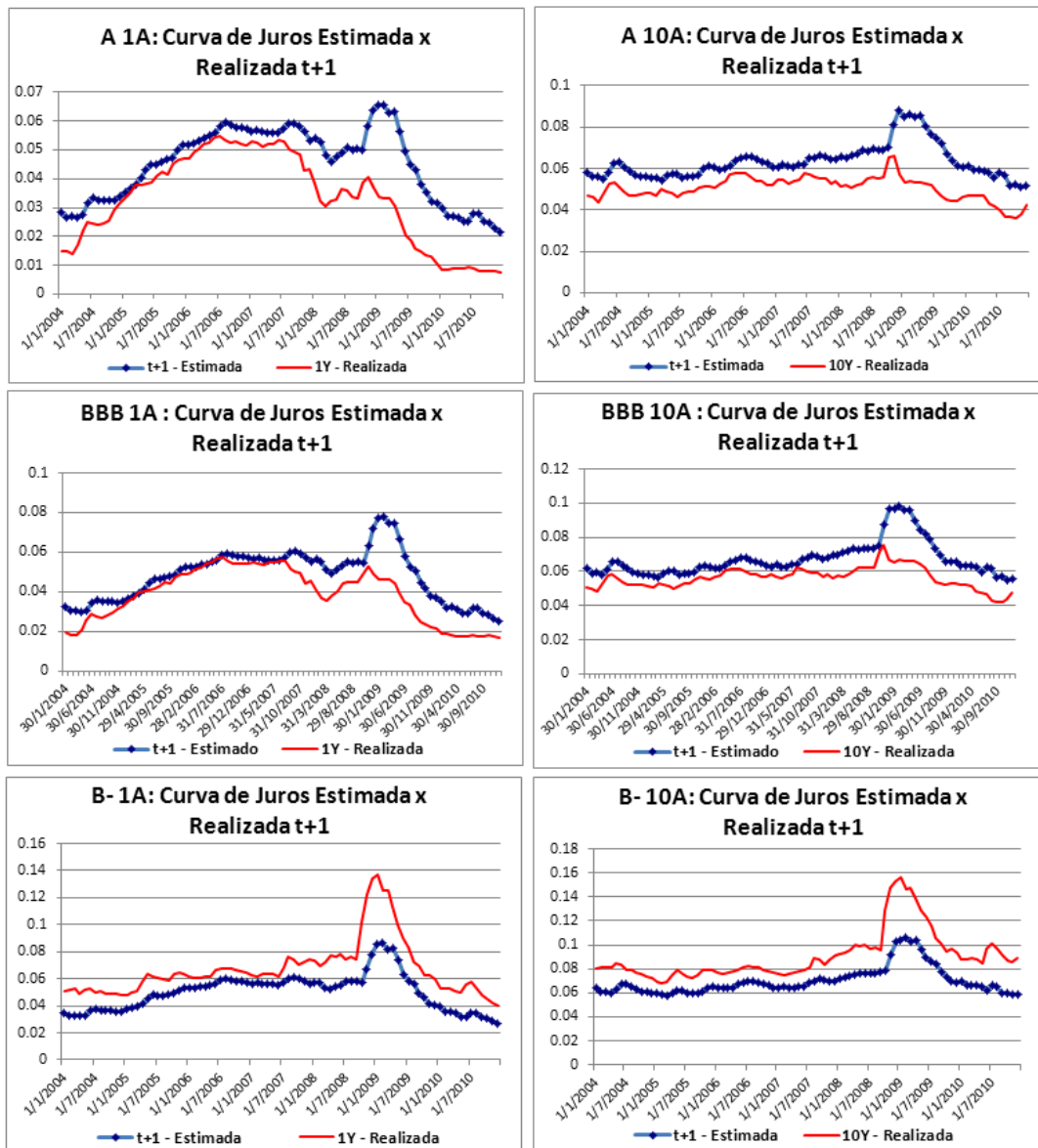
$$y_t^R(\tau) = \beta_{1,t}^R + \beta_{2,t}^R \left( \frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} \right) + \varepsilon_t^R \quad [4.1]$$

A equação está completamente definida uma vez que, consideramos, assim como em 4.1, o parâmetro  $\lambda = 0.0609$ , invariante no tempo e a variável  $\tau$  assumindo os valores das maturidades 3M, 6M, 1A, 2A, 3A, 4A, 5A, 7A, 8A, 9A, 10A, 15A, 20A, 25A e 30A em meses. A curva de juros foi obtida pela multiplicação matricial, para cada nível de risco, representada na estrutura abaixo:

$$\begin{bmatrix} y_t(3) \\ y_t(6) \\ \vdots \\ y_t(30) \end{bmatrix} = \begin{bmatrix} 1 & F(3) \\ 1 & F(6) \\ \vdots & \vdots \\ 1 & F(20) \end{bmatrix} \begin{bmatrix} L_t \\ S_t \end{bmatrix} \quad [4.18]$$

Destaca-se que para cada classificação de risco obtivemos a estrutura a termo prevista um, seis e doze passos à frente no período de janeiro de 2004 a dezembro de 2010. Da mesma maneira que no tópico anterior, apresentaremos os resultados da estimação para os *ratings* A, BBB e B-, sem perda de generalidade, já que o foco é a comparação do modelo com o passeio aleatório, a qual apresentaremos

para outras classificações de risco na próxima secção. Os gráficos relativos a previsões seis e doze passos à frente, encontram-se no Apêndice (Gráficos 7.23 e 7.24).



**Figura 4.11** Previsão t+1: Curva de juros estimada pelo modelo integrado x realizada no período de jan/2004 a dez/2010 para os ratings A, BBB E B-.

Analisaremos as estatísticas dos erros de previsão no próximo tópico, quando realizaremos a comparação com o passeio aleatório, modelo selecionado como *benchmark*.

## 4.6

### Comparação da previsão do modelo integrado com o passeio aleatório

Essa etapa consiste na comparação da estrutura a termo da taxa de juros projetada pelo modelo integrado com o modelo mais tradicional e simples da literatura de previsão: o passeio aleatório. A projeção pelo passeio aleatório é o último valor realizado, seja para a previsão um, seis, doze ou  $h$  passos à frente, que pode ser descrita pela equação:

$$\hat{y}_{t+h|t}(\tau) = y_t(\tau) \quad [4.19]$$

onde  $\tau$  representa a maturidade da estrutura a termo da taxa de juros.

Realizamos a comparação com o competidor para todos os *ratings*, no entanto, nos restringimos em apresentar graficamente o erro quadrático e as estatísticas dos erros de previsão para cinco níveis de risco selecionados: AAA, A, BBB, BB, e B- e para duas maturidades: 1 e 10 anos. Destacando que além de trabalharmos com 14 níveis de *rating*, a estrutura a termo de cada título é formada por 15 maturidades e a apresentação completa seria exaustiva e sem grandes ganhos analíticos. Na subamostra relativa à classificação dos títulos, conseguimos captar a variabilidade de risco assumido pelos agentes, já que as selecionadas são níveis médios da nossa amostra, representando de forma significativa a sua variabilidade, sem perda de generalidade. Quanto à maturidade, preferimos as taxas de 1 e 10 anos por acreditar na boa representatividade dessas, em termos de negociação.

Essa comparação também foi realizada por Diebold e Li (2006), no entanto, esses autores não se restringiram ao passeio aleatório apenas, uma vez que obtiveram ganhos em relação a esse. Analisaram a eficiência da previsão do seu modelo com outros competidores, como descrito na seção 2.2. Seguindo a métrica desses autores com relação à eficiência da previsão, calculamos as estatísticas dos erros também utilizadas por eles: média, desvio padrão, erro médio quadrático e as correlações de ordem um, seis e doze, referente ao período previsto de janeiro de 2004 a dezembro de 2010. Essas estatísticas e a comparação com o passeio aleatório podem ser verificadas nas Tabelas 4.19 a 4.21. Pela análise dessas, percebemos que o erro médio de previsão e o desvio padrão, em

nosso modelo, são maiores do que o erro do competidor para ambas às maturidades, e em todos os níveis de risco apresentados, com destaque, para o BBB. O mesmo ocorre para o erro quadrático médio que será analisado graficamente, mais adiante.

**Tabela 4.19** A: Estatística dos erros de previsão da curva de juros.

Estatísticas dos Erros - A						
	Média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
<b>t+1</b>						
<b>1A</b>	0.012	0.009	0.00022	0.957	0.748	0.569
<b>P.A. 1A</b>	0.000	0.002	0.00000	0.467	0.265	0.068
<b>10A</b>	0.013	0.007	0.00021	0.895	0.471	0.289
<b>P.A. 10A</b>	0.000	0.002	0.00001	0.289	-0.036	-0.049
<b>t+6</b>						
<b>1A</b>	0.012	0.014	0.00034	0.979	0.708	0.559
<b>P.A. 1A</b>	0.001	0.009	0.00007	0.952	0.416	0.259
<b>10A</b>	0.013	0.009	0.00024	0.933	0.474	0.357
<b>P.A. 10A</b>	0.001	0.005	0.00003	0.815	-0.209	0.043
<b>t+12</b>						
<b>1A</b>	0.012	0.019	0.00050	0.977	0.863	0.743
<b>P.A. 1A</b>	0.002	0.015	0.00021	0.978	0.742	0.526
<b>10A</b>	0.012	0.011	0.00026	0.893	0.686	0.357
<b>P.A. 10A</b>	0.001	0.006	0.00004	0.872	0.365	-0.027

No que tange a análise da estrutura de correlação dos erros, percebemos que esses são bastante correlacionados, não só para o competidor, como era o esperado pela própria construção do mesmo, mas também em nosso modelo. A correlação nos erros do nosso modelo são mais elevadas quanto maior o horizonte de previsão, indicando séries mais persistentes. Para os títulos de classificação de risco intermediária, essa análise é complementada pelas Tabelas 7.6 e 7.7 encontradas no Apêndice, relativas aos *ratings* A e BB. Destaca-se que a presença de uma estrutura de autocorrelação nos erros, foi detectada por Diebold e Li (2006), no entanto, não foi capaz de comprometer a boa performance preditiva do modelo desses autores, no horizonte de seis e doze passos. Desta forma, a detecção de uma estrutura nos erros do nosso modelo não seria um indicativo de ineficiência na previsão.

Tabela 4.20 BBB: Estatística dos erros de previsão da curva de juros.

Estatísticas dos Erros - BBB						
	Média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
<b>t+1</b>						
1A	0.009	0.008	0.00015	0.939	0.569	0.437
P.A. 1A	0.000	0.002	0.00000	0.454	0.192	-0.054
10A	0.011	0.007	0.00016	0.880	0.439	0.308
P.A. 10A	0.000	0.002	0.00001	0.308	0.076	-0.059
<b>t+6</b>						
1A	0.009	0.014	0.00028	0.970	0.558	0.331
P.A. 1A	0.000	0.009	0.00007	0.950	0.338	0.010
10A	0.011	0.010	0.00022	0.940	0.403	0.255
P.A. 10A	0.001	0.006	0.00003	0.853	0.074	-0.121
<b>t+12</b>						
1A	0.009	0.019	0.00045	0.980	0.748	0.507
P.A. 1A	0.001	0.014	0.00021	0.975	0.636	0.250
10A	0.011	0.013	0.00028	0.946	0.622	0.095
P.A. 10A	0.001	0.008	0.00007	0.917	0.460	-0.260

Tabela 4.21 B-: Estatística dos erros de previsão da curva de juros.

Estatísticas dos Erros - B-						
	Média	Desv. Padrão	EQM	Correl(1)	Correl(6)	Correl(12)
<b>t+1</b>						
1A	-0.017	0.010	0.00039	0.890	0.413	0.142
P.A. 1A	0.000	0.006	0.00003	0.495	-0.163	-0.096
10A	-0.021	0.010	0.00053	0.885	0.486	0.273
P.A. 10A	0.000	0.006	0.00003	0.273	-0.097	-0.091
<b>t+6</b>						
1A	-0.017	0.017	0.00057	0.922	-0.065	-0.250
P.A. 1A	0.000	0.021	0.00045	0.912	-0.135	-0.177
10A	-0.021	0.017	0.00070	0.917	0.020	-0.155
P.A. 10A	-0.001	0.019	0.00037	0.904	-0.117	-0.240
<b>t+12</b>						
1A	-0.017	0.023	0.00079	0.949	0.238	-0.497
P.A. 1A	0.000	0.029	0.00083	0.953	0.327	-0.419
10A	-0.021	0.021	0.00091	0.946	0.272	-0.378
P.A. 10A	-0.002	0.027	0.00071	0.948	0.311	-0.439

O principal critério que utilizamos para a avaliação do desempenho preditivo foi o erro médio quadrático<sup>17</sup>, o mesmo adotado por DL. Para facilitar a análise, construímos gráfico do erro quadrático e, optamos arbitrariamente pela apresentação do *rating* BBB, nessa secção. Os demais gráficos referentes à previsão um, seis e doze passos à frente para os títulos classificados como AAA, A, BB e B- sugerem as mesmas conclusões e, podem ser encontrados no Apêndice (Gráficos 7.10 a 7.21). Os resultados indicam que o nosso modelo não apresenta ganhos de eficiência em relação ao passeio aleatório, principalmente, na previsão um passo a frente, DL verificaram performance ligeiramente melhor que esse *benchmark* na previsão neste horizonte. Com efeito, nas palavras dos autores

<sup>17</sup> O erro é definido como o quadrado da diferença entre o valor previsto para o *yield* em um dado período e o valor efetivamente observado.

se referindo a projeção para um período, “*In relative terms, RMSE comparison at various maturities reveals that our forecasts, although slightly better than the random walk and slope regression forecasts, are indeed only very slightly better*”. Quanto às projeções para seis e doze períodos, o nosso modelo não apresenta ganhos muito significativos, como pode ser observado graficamente, isto é, para esses horizontes existem pequenos períodos em que o nosso modelo é ligeiramente superior.

Destacamos que nosso período preditivo, janeiro de 2004 a dezembro de 2010, compreende uma fase de grande instabilidade na economia mundial, principalmente americana, cujos primeiros sinais da grande crise que viria se abater sobre os mercados financeiros em 2008/2009 começaram a ser sentidos em 2007 com a crise do *subprime*. Esse fato dificulta extremamente a nossa previsão, como pode ser observado nos gráficos, principalmente em 2008/2009. Esses erros se devem a uma característica intrínseca da construção do modelo, a nossa estrutura para o cálculo das previsões é realizada nos super fatores, através de um modelo dinâmico, em que os coeficientes da equação desses *loadings* e as projeções são estimados a cada tempo  $t$ , através da inserção de uma nova informação, como descrito no tópico 4.3. Essa peculiaridade fez com que o modelo fosse capturando toda a variação das taxas no período da crise, descrita nos super fatores e propagando-a em nossas previsões. Diante disso, apresentamos também o gráfico dos erros no período de 2004 a 2007, com o objetivo de retirar os efeitos da crise sobre a previsão e com isso, percebemos maior eficiência relativa do nosso modelo.

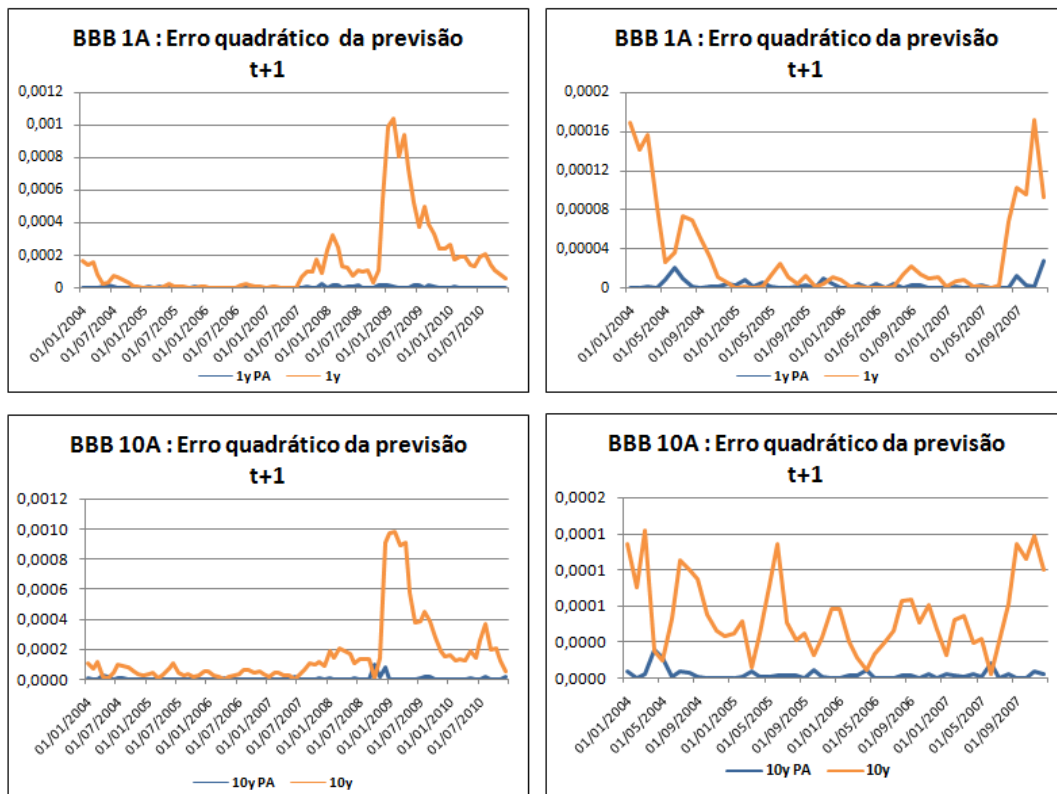


Figura 4.12 BBB: Erro quadrático da previsão t+1 para curvas de 1 e 10 anos nos períodos de jan/2004 a dez/2010 e jan/2004 a dez/2007.

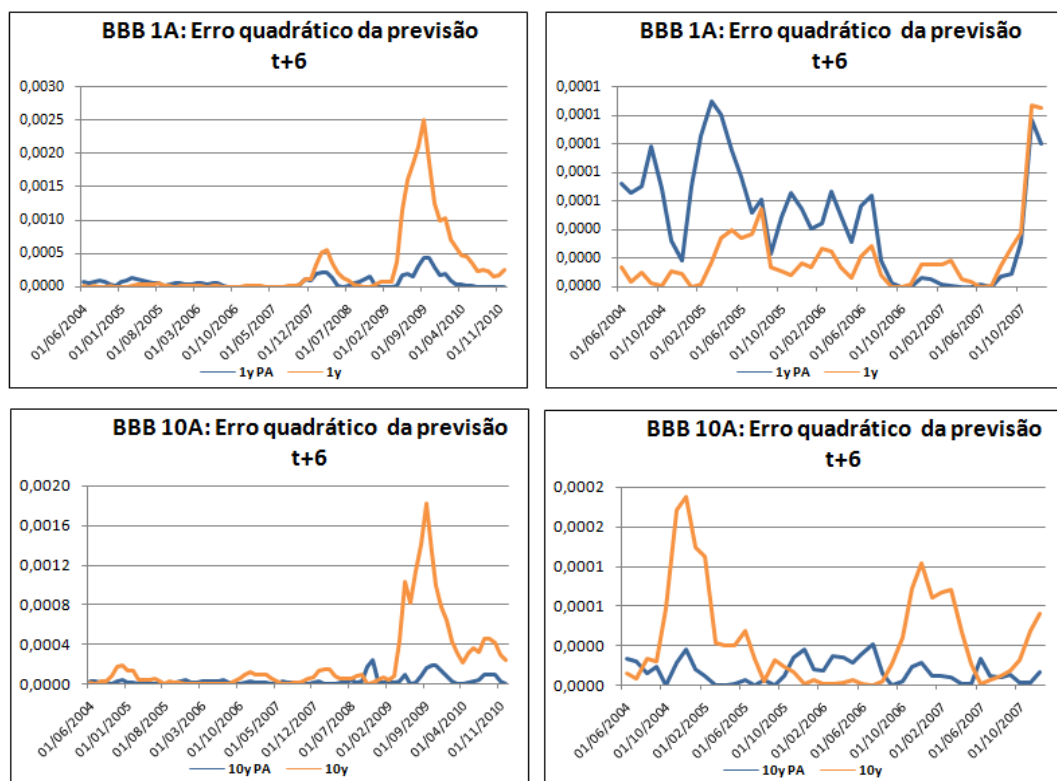
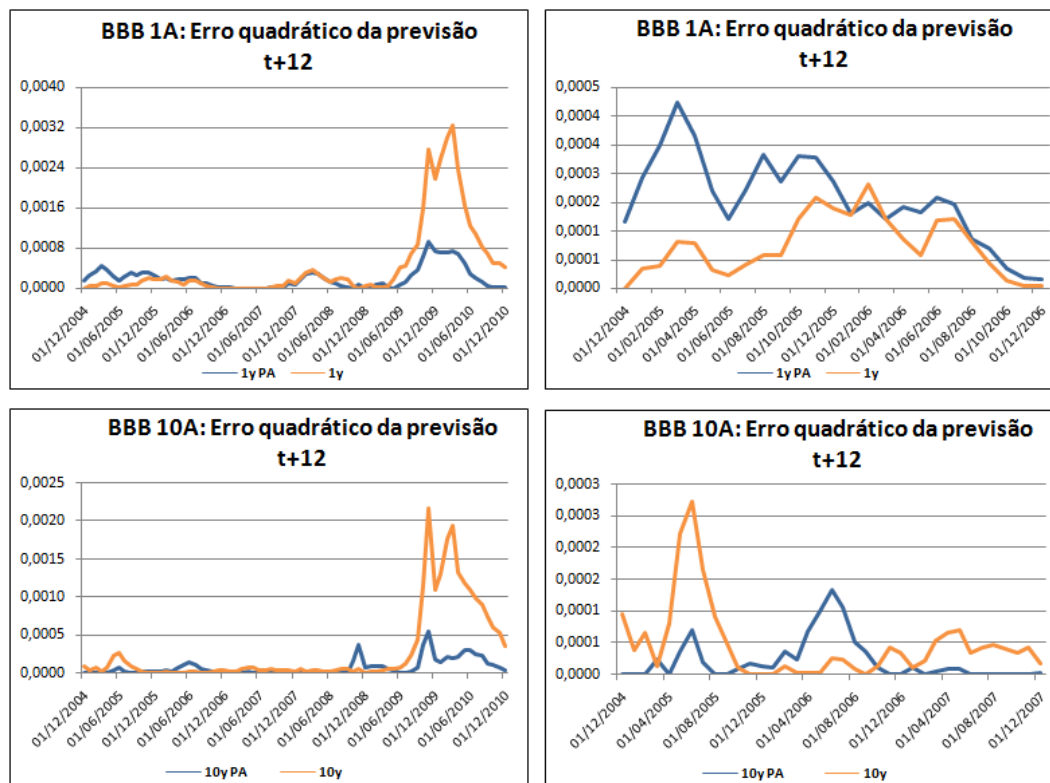


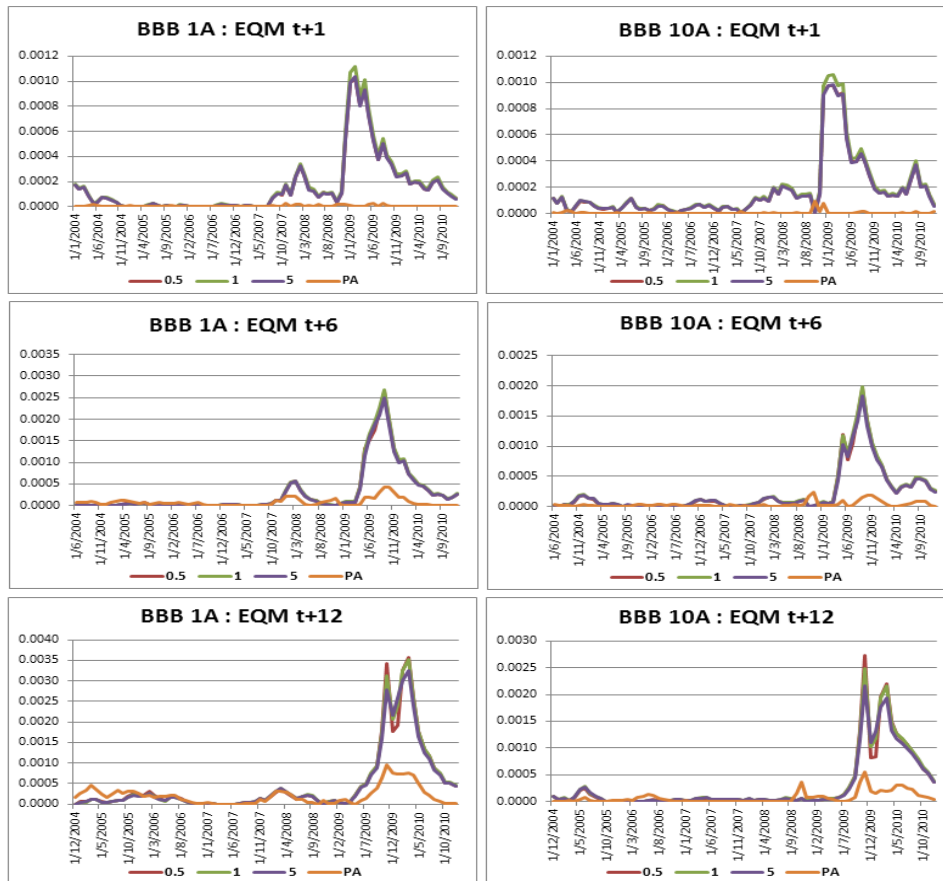
Figura 4.13 BBB: Erro quadrático da previsão t+6 para curvas de 1 e 10 anos nos períodos de jan/2004 a dez/2010 e jan/2004 a dez/2007.



**Figura 4.14** BBB: Erro quadrático da previsão t+12 para curvas de 1 e 10 anos nos períodos de jan/2004 a dez/2010 e jan/2004 a dez/2007.

Diante dos resultados apresentados acima, buscamos uma alternativa para melhorar o desempenho preditivo do nosso arcabouço de previsão. Essa se referiu a recalcular todo o modelo integrado tentando ajustar novos  $\lambda$ 's para a estimação dos super fatores referentes ao nível, onde acreditávamos, até então, que poderia estar a maior fonte de erros. Esperávamos que o erro estivesse na componente referente ao nível porque esse fator explica cerca de 90% da trajetória da curva de juros e, pelas análises gráficas e de correlação, como apresentado na seção 4.2, o fator inclinação está bem ajustado. Desta forma, reestimamos a curva de juros testando outros  $\lambda$ 's, fazendo-o assumir os valores: 0,5, 1 e 5 e, comparando-as com a previsão realizada pelo passeio aleatório. Pela análise do gráfico 4.10, percebemos que a diferenciação desse parâmetro não trouxe ganhos, uma vez que a estimativa foi pouco afetada, a curva referente ao  $\lambda = 1$  fica quase que sobreposta a do  $\lambda = 0,5$  e para esse nível de risco erram mais do que assumir o parâmetro igual a 5. Pela análise de todos os *ratings*, cujos gráficos comparativos estão no apêndice (Gráficos 7.22 a 7.25), optamos pela utilização do  $\lambda = 5$ , uma vez que os ganhos de previsão dos demais valores não são significativos.





**Figura 4.15** BBB: Erro quadrático da previsão das curvas de juros de 1 e 10 anos para diferentes  $\lambda$ 's no período de jan/2004 a dez/2010.

Acreditamos que a eficiência na previsão detectada por Diebold e Li (2006) e não verificada em nosso modelo, não se referem apenas a mudanças na estrutura, podem também ter relação como o fato de a amostra se composta de *treasuries*, séries mais estáveis, do que os títulos corporativos. De qualquer forma, os resultados apresentados sugerem que mudanças na estrutura do modelo devem ser realizadas, a fim de garantir melhor desempenho preditivo. As sugestões para trabalhos posteriores serão apresentadas no próximo capítulo.