5
A Metodologia Diebold e Li

O principal *insight* de Diebold e Li (2006) é utilizar a forma funcional proposta por Nelson e Siegel (1987) para ajustar a dinâmica da curva de juros norte-americana da maneira mais parcimoniosa possível e, com isso, extrair benefícios em termos de previsão fora da amostra.

Nelson e Siegel propõem um modelo onde a *forward rate curve* observada em um dado período é aproximada pela seguinte função:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t} e^{-\lambda_t \tau} + \beta_{3,t} \lambda_t e^{-\lambda_t \tau} \tag{11}$$

onde τ e $f_t(\tau)$ representam o prazo e a respectiva taxa *forward*. A curva de juros correspondente é aproximada pela função:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{3,t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right) \tag{12}$$

A taxa *forward* de prazo τ observada no período *t* é igual ao retorno que um investidor tem ao realizar a seguinte operação:

• adquirir em *t* um *zero coupon bond* com vencimento em $\tau+1$ períodos pagando um preço $P_{t,\tau+1}$.

• vender uma quantidade igual a $P_{t,\tau+1} / P_{t,\tau}$ de *zeros* de τ períodos pelo preço $P_{t,\tau}$ no mesmo momento.

Repare que essa operação equivale a desembolsar um montante igual a $\$\,{}^{P_{t,\tau+1}}\!/_{P_{t,\tau}}$ em $t=\tau$ e receber \$1 em $t=\tau+1$, sendo a rentabilidade dessa operação justamente a *forward rate* $f_t(\tau)$. Já $y_t(\tau)$ denota o retorno de um *zero coupon bond* com τ períodos até o vencimento adquirido no período *t*. Como $P_{t,\tau}=\frac{1}{(1+y_t(\tau))^{\tau}}$, $f_t(\tau)$ e $y_t(\tau)$ obedecem a seguinte relação[17]:

$$f_t(\tau)=\frac{(1+y_t(\tau+1))^{\tau+1}}{(1+y_t(\tau))^{\tau}}-1 \tag{13}$$

A modelagem de Nelson e Siegel faz com que existam três fatores comuns, denotados por $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$. Seus *loadings* são tais que:

• Uma perturbação positiva (negativa) no primeiro fator gera um deslocamento paralelo e para cima (baixo) da curva de juros.

• Uma perturbação positiva (negativa) no segundo fator faz com que os retornos dos títulos mais curtos aumentem (diminuam) mais do que os retornos dos títulos mais longos. Isso decorre do formato da função $F(\tau)=\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}$, que é monotonicamente decrescente para qualquer valor de λ_t. O parâmetro λ_t controla a velocidade de decaimento da função $F(\tau)$, com efeito, à medida que λ_t aumenta, os valores assumidos pela função $F(\tau)$ vão se aproximando mais rapidamente de zero.

• Uma perturbação positiva (negativa) no terceiro fator faz com que os retornos dos títulos de prazos intermediários aumentem (diminuam) mais do que

[17] Se trabalhássemos em um regime de capitalização contínua, o preço $P_{t,\tau}$ de um título de τ períodos emitido em *t* seria dado por $P_{t,\tau}=e^{-\tau y_t(\tau)}$. Já o relacionamento entre $f_t(\tau)$ e $y_t(\tau)$ seria definido por $f_t(\tau)=-\frac{dP_t(\tau)}{d\tau}\frac{1}{P_t(\tau)}$.

os retornos dos títulos de prazos curtos e longos. Isso decorre do formato da função $G(\tau)=\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}-e^{-\lambda_t \tau}$, que é estritamente côncava e atinge um máximo para um determinado τ^* intermediário. O valor de τ^* varia negativamente com λ_t.

O comportamento acima descrito permite que associemos os três fatores do modelo de Nelson e Siegel aos fatores nível, inclinação e curvatura identificados nos trabalhos discutidos no Capítulo 3. A Figura 4, que mostra os perfis correspondentes a cada *loading* (denotados por L1, L2 e L3), reforça essa interpretação. Comparando a Figura 1 com a Figura 4 percebe-se que, pelo menos a grosso modo, os formatos dos *loadings* propostos por Nelson e Siegel são aproximadamente iguais aos formatos dos *loadings* estimados por Varga e Valli. É exatamente o sucesso na reprodução dos *loadings* empíricos que Diebold e Li exploram a fim de construir o seu modelo de previsão, conforme veremos a seguir.

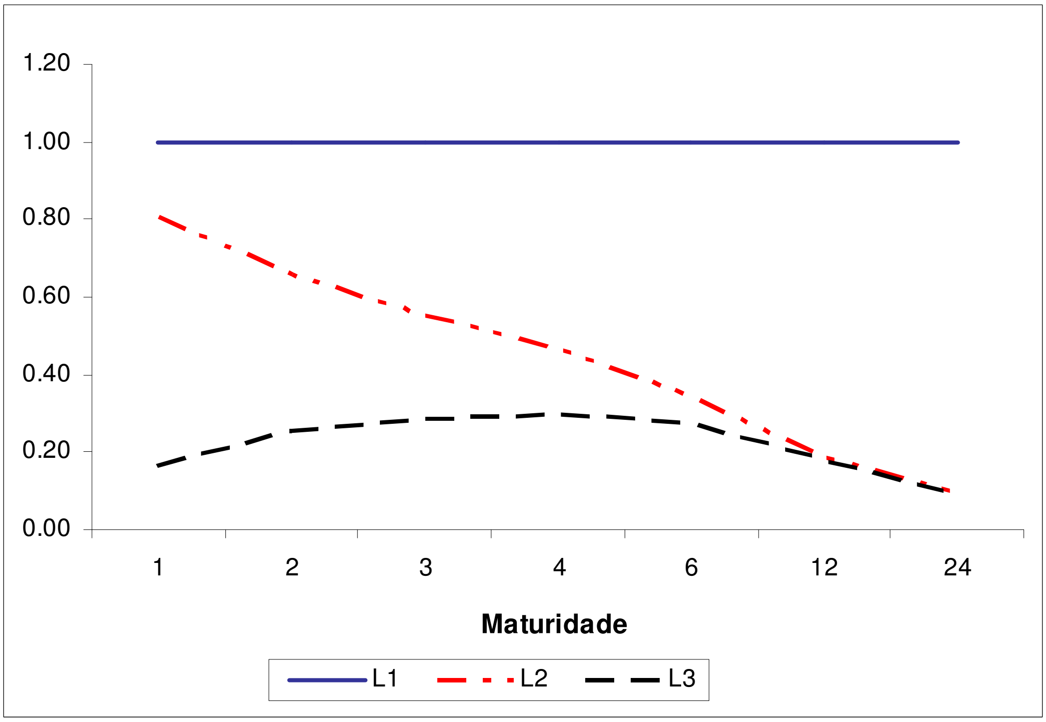

Figura 4 - Perfis correspondentes a cada *factor loading*. Os fatores foram gerados com λ_t constante e igual a 0.448321.

Diebold e Li afirmam que o modelo de Nelson e Siegel é capaz de reproduzir os principais fatos estilizados observados para a estrutura a termo norte-americana, a saber:

• A curva de juros média é crescente e côncava. Na abordagem de Nelson e Siegel, a curva de juros média está associada aos valores médios de $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$[18].

• A curva de juros pode assumir diversas formas ao longo do tempo, quais sejam, crescente, decrescente, *humped* e *inverted humped*. Ora, dependendo dos valores específicos assumidos pelos fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$, é possível que a estrutura a termo assuma qualquer um desses formatos.

• A dinâmica das taxas é mais persistente que a dos *spreads*[19]. A persistência das taxas pode ocorrer se a dinâmica do fator $\beta_{1,t}$ for persistente, enquanto que a baixa persistência dos *spreads* pode ser observada se a dinâmica do fator $\beta_{2,t}$ for pouco persistente.

• A curva de juros é tal que as taxas de curto prazo são mais voláteis do que as taxas de longo prazo. No contexto de Nelson e Siegel as taxas curtas dependem de $\beta_{1,t}$ e $\beta_{2,t}$, enquanto que as taxas longas dependem quase exclusivamente de $\beta_{1,t}$. Por conseguinte, a assertiva inicial será verdadeira se o fator $\beta_{1,t}$ tiver a menor variância incondicional.

• As taxas longas são mais persistentes que as taxas curtas. Como as taxas longas dependem quase exclusivamente de $\beta_{1,t}$ e as taxas curtas dependem de $\beta_{1,t}$ e $\beta_{2,t}$, a assertiva inicial será verdadeira se o fator $\beta_{1,t}$ for o mais persistente.

---

[18] O método utilizado por Diebold e Li para estimar os fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ será examinado a seguir.

[19] Diferença entre taxas relativas a prazos diferentes.

Diebold e Li demonstram essa concordância analisando uma amostra de taxas de juros calculadas a partir dos preços de *treasuries* coletados nos arquivos pertencentes à base de dados "Fama CRSP *Treasury Bill*". A amostra (de freqüência mensal) é composta por observações retiradas do período que vai de janeiro de 1985 a dezembro de 2000 e engloba títulos com vencimento em 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 e 120 meses. A Tabela 11 traz os valores médios de $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ calculados com base nessa amostra, assim como os seus desvios padrão e as autocorrelações de ordem 1, 12 e 30 (denotadas por $r(1)$, $r(12)$ e $r(30)$, respectivamente). Nota-se que o primeiro fator é o mais persistente de todos, pois apresenta autocorrelações que, em conjunto, têm a maior magnitude. O primeiro fator também é o menos volátil, pois o seu desvio padrão é menor do que o calculado para os fatores dois e três. Os autores também demonstram que, para os valores médios de $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$, a curva de juros média é crescente e aproximadamente côncava. Vale comentar também que os fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ recuperados por Diebold e Li possuem uma alta correlação com as medidas empíricas de nível, inclinação e curvatura comumente utilizadas[20]. Esse resultado é bastante desejável na medida em que o *framework* de Nelson e Siegel não poderia ser considerado adequado se os fatores dele decorrentes (que dependem das formas funcionais pré-especificadas para os *loadings*) não se assemelhassem aos fatores decorrentes do que os agentes econômicos entendem por medidas de nível, inclinação e curvatura.

Tabela 11 - Estatísticas dos fatores comuns para a curva de juros norte-americana.

Fator	Média	Desvio-Padrão	Mínimo	Máximo	r_1	r_{12}	r_{30}	ADF
Beta 1	7.5790	1.5240	4.4270	12.0880	0.9570	0.5110	0.4540	-2.4100
Beta 2	-2.0980	1.6080	-5.6160	0.9190	0.9690	0.4520	-0.0820	-1.2050
Beta 3	-0.1620	1.6870	-5.2490	4.2340	0.9010	0.3530	-0.0060	-3.516***

Fonte: Diebold e Li (2006).

20 Para o mercado norte-americano, define-se nível como sendo o *yield* de um *zero coupon bond* com prazo de vencimento de 10 anos, inclinação como sendo a diferença entre os *yields* de dois *zero counpon bonds* com prazos de vencimento de 10 anos e 3 meses, e curvatura como sendo igual a duas vezes o *yield* de um *zero coupon bond* de 2 anos subtraído da soma dos *yields* de *zeros* de 3 meses e 10 anos. As correlações entre os fatores empíricos assim definidos e os estimados são altas, da ordem de 0.97. Verificou-se também que, dentre os fatores empíricos, o primeiro (terceiro) é o mais (menos) persistente.

As considerações acima enfatizam o sucesso da formulação de Nelson e Siegel em reproduzir os principais fatos estilizados registrados para o comportamento da curva de juros norte-americana e a sua compatibilidade com os resultados encontrados por Litterman e Scheinkman, a saber, que a dinâmica da curva de juros depende fundamentalmente de três fatores comuns (nível, inclinação e curvatura). Logo, é natural supor que a formulação de Nelson e Siegel pode ser utilizada na construção de um modelo econométrico que descreva o comportamento da estrutura a termo e preveja a sua trajetória futura. Espera-se que essa observação seja válida também para dados brasileiros, pois já demonstramos acima que as características das curvas de juros brasileira e norte-americana são aproximadamente as mesmas (logo, se a formulação de Nelson e Siegel é aceitável para dados norte-americanos, deve ser aceitável também para dados brasileiros).

Diebold e Li também argumentam a favor da utilização do arcabouço de Nelson e Siegel dizendo que ele permite construir modelos parcimoniosos e que esses, em geral, se revelam mais eficientes em exercícios de previsão. Segundo os autores, a parcimônia é um atributo desejável porque modelos com muitos parâmetros geram incerteza significativa (pois os parâmetros em excesso devem ser estimados) e incorrem em *overfitting*, ou seja, "explicam" a contento o comportamento dos dados não porque são "corretos" mas sim porque são suficientemente complexos. Os autores finalizam dizendo que a utilização do arcabouço de Nelson e Siegel emana de uma "... *broad interpretation of the shrinkage principle, which has a firm foundation in Bayes-Stein theory, in empirical intuition and in an accumulated track record of good performance*". Segundo Diebold e Li,

> *"... here we interpret the shrinkage principle as the insight that the imposition of restrictions, which will of course degrade in-sample fit, may nevertheless be useful for out-of-sample forecasting, even if the restrictions are false".*

Dessa maneira, as restrições inerentes ao arcabouço de Nelson e Siegel (que exigem que os *loadings* em (12) venham de formas funcionais pré-especificadas) podem fazer com que um modelo econométrico que se baseie nelas acabe exibindo uma *performance* preditiva melhor.

Agora que já sabemos os fundamentos da estratégia proposta por Diebold e Li é possível passar para os seus detalhes. Os autores recomendam que:

• O modelo com os três fatores comuns descrito na eq.(12) seja utilizado para "ajustar" a curva de juros.

• O parâmetro λ_t seja considerado constante e calibrado de tal forma que a função $G(\tau)$ atinja seu máximo em $\tau = 30$, prazo que os autores definem como de maturidade média[21].

• Os fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ sejam estimados por mínimos quadrados ordinários (opção factível porque λ_t é igual a um número definido *a priori*, o que faz com que os *factor loadings* sejam perfeitamente conhecidos).

• As séries temporais correspondentes aos fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ sejam modeladas como processos estocásticos univariados autoregressivos de primeira ordem (AR(1)). Sendo assim, as previsões acerca da trajetória futura dos fatores são obtidas através da seguinte especificação:

$$\beta_{i,t+h} = c_i + \gamma_i \beta_{i,t} + \varepsilon_{i,t} \qquad (14)$$

onde $\beta_{i,t}$ e $\beta_{i,t+h}$ denotam os valores assumidos pelo *i*-ésimo fator nos instantes *t* e $t+h$, respectivamente. Os parâmetros c_i e γ_i são estimados através da regressão de $\beta_{i,t}$ em uma constante e em $\beta_{i,t-h}$. A opção por modelos univariados repousa na constatação feita por Diebold e Li de que os fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ não possuem correlação cruzada significativa.

• As previsões para cada *yield* sejam obtidas a partir de:

---

[21] Com efeito, $\tau = 30$ está no "centro" da seqüência de maturidades {3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120}.

$$\hat{y}_{t+h}(\tau)=\hat{\beta}_{1,t+h}+\hat{\beta}_{2,t+h}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right)+\hat{\beta}_{3,t+h}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}-e^{-\lambda_t\tau}\right) \quad (15)$$

onde $\hat{\beta}_{i,t+h}=\hat{c}_i+\hat{\gamma}_i\beta_{i,t}$. Os valores $\hat{c}_i$ e $\hat{\gamma}_i$ denotam as estimativas de c_i e γ_i, enquanto que $\hat{\beta}_{i,t+h}$ denota a previsão *h* passos à frente do valor que o *i*-ésimo fator assume no instante $t+h$ condicionada ao fato do último valor observado ser igual a $\beta_{i,t}$ (o parâmetro *h*, portanto, denota o horizonte de previsão).

É necessário ressaltar que o tratamento dispensado ao parâmetro λ_t e a modelagem dos fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ como meros processos AR(1) estão de acordo com o desejo dos autores de "... *intentionally impose substantial a priori structure, motivated by simplicity, parsimony, and theory*".

Os autores aplicam a estratégia delineada acima a uma amostra com dados mensais de janeiro de 1985 a dezembro de 2000. Uma sub-amostra iniciada em janeiro de 1994 é separada para julgar a qualidade das previsões fora da amostra. O principal critério de avaliação adotado é o erro médio quadrático[22] e os horizontes de previsão considerados correspondem a $h=1$, 6 e 12 períodos à frente. Os *yields* para os quais se calculam previsões estão vinculados aos prazos de 3, 12, 36, 60 e 120 meses até o vencimento. Tanto o modelo sugerido pelos autores quanto os modelos que serviram de *benchmark* para julgar se as inovações sugeridas são bem-sucedidas ou não são estimados recursivamente de janeiro de 1985 até o momento no qual as previsões são calculadas[23].

[22] 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. A média desses erros é que serve de medida para o desempenho preditivo do modelo sugerido por Diebold e Li e dos modelos que eles propõem como base de comparação.

[23] O procedimento é tal que, no primeiro passo, os modelos são estimados com dados de janeiro de 1985 a dezembro de 1994. Em seguida, os modelos são utilizados para gerar previsões para a curva de juros de janeiro de 1995, junho de 1995 e dezembro de 1995, e os erros cometidos são calculados. O próximo passo é uma repetição do primeiro, ou seja, os modelos são estimados de novo com dados de janeiro de 1985 a janeiro de 1995 e, em seguida, eles são utilizados para gerar previsões para a curva de juros de fevereiro de 1995, julho de 1995 e janeiro de 1996. Os erros cometidos são calculados e acumulados àqueles calculados anteriormente. O procedimento se repete até que a amostra utilizada para estimação englobe dados de janeiro de 1985 a dezembro de

Diebold e Li concluem que o modelo proposto é capaz de alcançar resultados satisfatórios. Mais especificamente, os autores afirmam que os resultados são decepcionantes para $h = 1$ ("... *our model's 1-month-ahead forecasting results are in certain respects humbling* ..."), já que a diferença de *performance* entre o seu modelo e um mero *random walk* não é estatisticamente significativa. Para $h = 6$ e $h = 12$, no entanto, os autores sustentam que os resultados são animadores. Com efeito, nas palavras de Diebold e Li, "... *our model's 6-month-ahead forecasting results are noticeably improved, and our model's 12-month-ahead forecasting results are much improved*". Em particular, o desempenho preditivo do modelo para previsões 12 passos à frente supera o desempenho dos demais competidores para todos os prazos de vencimento considerados, e as diferenças observadas são estatisticamente significativas na maioria dos casos. O quadro geral sugere que o desempenho preditivo do modelo de Diebold e Li é aceitável e melhora à medida que o horizonte de previsão aumenta[24].

1999. A análise de desempenho se baseia nas previsões obtidas a partir das 60 repetições executadas.

[24] É necessário ressaltar que Vereda, Lopes e Fukuda (2006) chegam a uma conclusão semelhante ao testarem um modelo inspirado em Evans e Marshall (1998) para gerar previsões acerca do comportamento futuro da estrutura a termo brasileira.