2
Metodologia

2.1
Contextualização

Dois são os modelos mais conhecidos e apontados por Diebold e Li (2006) no que se refere à modelagem da estrutura a termo: os modelos de não arbitragem e os modelos de equilíbrio. Esse modelos são muito utilizados para avaliação de contratos de derivativos relacionados às taxas de juros. Os modelos tradicionais de não arbitragem buscam o ajuste perfeito dessa em um determinado instante de tempo, impedindo, assim, qualquer possibilidade de arbitragem, não estando focados na dinâmica ou previsão das ETTJ. Já os modelos de equilíbrio tem como objetivo modelar a dinâmica da taxa instantânea, utilizando tipicamente modelos afins, o que os relaciona com a previsão, apesar dos resultados empíricos de previsões dentro da amostra, não serem muito significativos. Duffie e Kan (1996), Heath, Jarrow e Morton (1992), Ho e Lee (1986), Hull e White (1990) devem ser consultados para uma discussão mais detalhada sobre os modelos de não abitragem e Cox, Ingersoll e Ross (1985) e Vasicek (1977) para os modelos de equilíbrio. Se insere também nessa literatura, trabalhos referentes a modelagem conjunta das ETTJ com variáveis macroeconômicas como apresentado no trabalho de Ang e Piazzesi (2003) e ainda, modelos mais tradicionais centrados nas relações de cointegração geradas a partir do *spread* entre as taxas longas e curtas, que são os modelos de correção de erros, como o proposto por Hall, Anderson e Granger (1992).

O modelo de Dielbold e Li (2006), que será o referencial teórico para este trabalho, se destaca na literatura descrita acima por conter elementos tanto dos modelos de equilibrio, quanto dos modelos de não arbitragem. No que se refere a primeira classe, possui a flexibilidade de ajustar as taxas observadas em um dado instante de tempo, característica proveniente do modelo de Nelson e Siegel, utilizado pelos autores como base. Já para os modelos de não arbitragem, apesar de não compreender as restrições que previnem as oportunidades de arbitragem, esse modelo descreve o movimento dos fatores que guiam a trajetória das curvas, sendo assim capaz de modelar a dinâmica das taxas. Essas características tornam esse modelo bastante atrativo ao nosso próposito, principalmente porque, além do objetivo dos autores ter sido estudar modelos que fossem bons geradores de previsão fora da amostra, esse possui propriedades desejáveis, tais quais: a capacidade de capturar os diferentes formatos assumidos pela curva de juros ao longo do tempo, a descrição simples do movimento conjunto das curvas, a construção de um modelo parcimonioso e a aderência aos fatos estilizados observados para a estrutura a termo norte-americana que foi objeto da modelagem.

Desta forma, esse trabalho está centrado na combinação do modelo proposto por Diebold e Li, que foi uma adaptação da forma funcional proposta por Nelson e Siegel (1987), com extensões propostas em Diebold, Li e Yue (2008), e inovações que serão discutidas adiante.

2.2
O modelo de Dielbod e Li (DL)

O modelo de Diebold e Li é construído sobre o modelo sugerido por Nelson e Siegel (NS), e neste modelo a *forward rate curve* observada em um dado período do tempo t é aproximada pela seguinte função:

$$\mathrm{f_t}(\tau) = \beta_{1,\mathrm{t}} + \beta_{2,\mathrm{t}}\, \mathrm{e}^{-\lambda_t \tau} + \beta_{3,\mathrm{t}} \lambda_\mathrm{t} \mathrm{e}^{-\lambda_t \tau} \qquad [2.1]$$

onde τ representa a maturidade e $f_t(\tau)$ a taxa *forward*. A função que aproxima a curva de juros correspondente a essa formulação para as taxas *forward* seria:

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

Onde $\beta_{1,\mathrm{t}}, \beta_{2,\mathrm{t}}\, \mathrm{e}\, \beta_{3,\mathrm{t}}$ representam os fatores comuns (*loadings*, como é denominado no modelo de fatores) que teriam uma lei de movimento própria, que por sua vez, resultariam na dinâmica da própria estrutura a termo. A discussão referente ao comportamento dos *loadings* é importante, já que a estimação desses será um dos focos desse trabalho. Desta forma, os efeitos de choques nesses fatores são:

• Um choque positivo (negativo) no primeiro fator gera um deslocamento paralelo para cima (baixo) da curva de juros.

• O formato da função $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}$, associado ao segundo fator, que denominaremos $F(\tau)$ é decrescente para qualquer valor positivo de λ_t o que faz com que um choque positivo (negativo) neste fator responda de forma a aumentar (diminuir) menos os retornos dos títulos com maturidades mais longas do que os retornos dos títulos de maturidades mais curtas. O parâmetro λ_t, que pode ser variante no tempo, controla a velocidade de decaimento da função F .

• O formato de "U" invertido da função $\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}$ associada ao terceiro fator, que denominademos $G(\tau)$ atinge um máximo em um τ^* intermediário, fazendo com que um choque positivo (negativo) nesse fator responda de forma a aumentar (diminuir) mais os retornos dos títulos intermediários do que os retornos dos títulos de prazo curtos e longos. Destaca-se que a variável τ varia negativamente com λ_t.

Esse entendimento do modelo de Nelson e Siegel permite associar a decomposição da estrutura a termo nos fatores nível, inclinação e curvatura, identificados nos trabalhos de Litterman e Sheinkman (1991) e Knez, Litterman e Sheinkman (1994). A reprodução dos fatores empíricos foi explorado por DL, e essa modelagem será utilizada nesse estudo a fim de construir o modelo de previsão.

Diebold e Li afirmam em seu artigo, que essa estrutura descrita acima para descrever a curva de juros 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 NS, a curva de juros média está associada aos valores médios de $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$.

• A curva de juros pode assumir diversas formas ao longo do tempo, quais sejam: crescente, decrescente, parábola e parábola invertida. Dependendo dos valores específicos assumidos pelos fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$, a estrutura a termo pode assumir qualquer um desses formatos.

• O movimento das taxas é mais persistente que o dos *spreads*. A persistência das taxas pode ser reproduzida caso o fator $\beta_{1,t}$ seja persistente, enquanto que a baixa persistência dos *spreads* pode ser reproduzida caso o fator $\beta_{2,t}$ não seja tão persistente.

• A curva de juros é tal, que as taxas de curto prazo são mais voláteis 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 menor variância incondicional.

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

Destaca-se que Diebold e Li (2006) demonstram essas relações do modelo de NS, através de uma amostra composta pelas taxas de juros calculadas a partir dos preços do *treasuries*, com observações referentes ao período de janeiro de 1985 a dezembro de 2000, englobando títulos com vencimento em 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 108 e 120 meses. Além da reprodução dos fatos estilizados, e em linha com a existência dos fatores nível, inclinação e curvatura para explicar a trajetória dos juros, a utilização desse *framework* permite construir modelos parcimoniosos, que em geral, se revelam eficientes no exercício de previsão. No entanto, em sua essência essa estrutura descrita é estática, e é nesse contexto que surge uma das principais contribuições de DL. Esses autores conferiram dinamismo ao modelo interpretando as sequências geradas pela estimação dos três fatores, em cada instante de tempo, formando assim, séries temporais que demostram o comportamento desses fatores latentes se modificando ao longo do tempo. Para isso, recomendam a utilização desse arcabouço da seguinte forma:

• A curva de juros seja ajustada pelo modelo com três fatores descrito em [2.1].

• O parâmetro λ_t seja determinado de forma que a função G(τ) atinja seu máximo na maturidade média, neste caso para τ =30 meses[1] e seja constante.

[1] A sequência de maturidades estudadas pelos autores é 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 108 e 120 meses, onde o centro é T=30 meses.

• Pelo fato de o parâmetro λ_t ser igual a um número conhecido a priori,o que faz com que as funções F(τ) e G(τ) sejam perfeitamente definidas, torna-se factível a estimação dos fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ por mínimos quadrados ordinários.

• Diante da evidência empírica de que os fatores $\beta_{1,t}$, $\beta_{2,t}$ e $\beta_{3,t}$ não possuem correlação cruzada significativa, as séries temporais destes devem ser modeladas como processos estocásticos univariados que seguem a seguinte lei de movimento:

$$\beta_{i,t+h} = c_i^h + \gamma_i^h \beta_{i,t} + \varepsilon_{i,t+1} \qquad [2.3]$$

onde $\beta_{i,t}$ e $\beta_{i,t+h}$ são, respectivamente, os valores assumidos pelo *i-ésimo* fator nos instantes t e t+h e, através de uma regressão de $\beta_{i,t}$ em uma constante e em $\beta_{i,t-h}$, estimamos os parâmetros c_i^h e γ_i^h . Essa é a maneira mais livre de gerar previsões para os fatores, porque os parâmetros c e γ podem se ajustar a h. No entanto, podemos utilizar o processo AR(1) convencional que impõe restrições a esses parâmetros.

Desta forma, o modelo de previsão fora da amostra para a estrutura a termo da taxa de juros estaria definido e a curva prevista dependeria apenas da previsão dos fatores, que pode ser observado abaixo:

$$\mathrm{E}[\mathrm{y}_{\mathrm{t+h}}(\tau) \mid \mathrm{t}] = \mathrm{E}[\beta_{1,\mathrm{t+h}} \mid \mathrm{t}] + \mathrm{E}[\beta_{2,\mathrm{t+h}} \mid \mathrm{t}]\ \frac{1 - \mathrm{e}^{-\lambda_t \tau}}{\lambda_t \tau} + \mathrm{E}[\beta_{3,\mathrm{t+h}} \mid \mathrm{t}](\ \frac{1 - \mathrm{e}^{-\lambda_t \tau}}{\lambda_t \tau} - \mathrm{e}^{-\lambda_t \tau})$$

$$\hat{y}_{t+h}(\tau) = \hat{\beta}_{1.t+h} + \hat{\beta}_{2.t+h}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \hat{\beta}_{3.t+h}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right) \qquad [2.4]$$

Onde $\hat{\beta}_{i,t+h} = \hat{c}_i + \hat{\gamma}_i \beta_{i,t}$ $i = 1,2,3$. 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).

Os autores aplicam a estratégia delineada acima a uma subamostra para testar a previsão fora da amostra, onde o principal critério de avaliação é o erro quadrático médio. Os horizontes de previsão considerados correspondem a h=1, 6 e 12 períodos à frente e as maturidades para o qual calculam as previsões são 3, 12, 36, 60 e 120 meses. Neste exercício, os autores concluem que o modelo proposto é capaz de alcançar resultados satisfatórios, mais especificamente, afirmam que para a previsão um período à frente (h=1) o resultado é decepcionante, já que comparado a um passeio aleatório, a diferença entre eles não é estatisticamente significativa. No entanto, para a previsão seis e doze períodos à frente (h=6 e h=12) definem os resultados como animadores, principalmente para h=12, onde o modelo supera os demais competidores[2] para todos os prazos de vencimento considerados e as diferenças observadas são estatisticamente significativas na maioria dos casos. O contexto geral sugere que a performance preditiva do modelo é aceitável e melhora à medida que aumenta o horizonte de previsão.

2.3
O modelo de Diebold, Li e Yue (DLY)

O desenvolvimento realizado por Diebold, Li e Yue (2008) consiste em aplicar o *framework* de Nelson e Siegel dinamizados por Diebold e Li, na construção de um modelo hierárquico para a análise conjunta das curvas de juros de alguns países desenvolvidos, sendo eles: EUA, Reino Unido, Alemanha e Japão. Os autores buscam mostrar que o movimento das curvas de juros desses países são conduzidas fundamentalmente por fatores (nível e inclinação), que por sua vez são guiados por dois tipos de forças, as comuns e as idiossincráticas. As primeiras, também denominadas fatores globais, são aquelas que afetam simultaneamente a dinâmica das curvas de juros dos diferentes países, enquanto as forças idiossincráticas ou locais são aquelas que afetam cada país especificamente, sem agir sobre os demais.

A estrutura inicial do modelo é uma versão simplificada da equação [2.2], onde os fatores $\beta_{1,t}$ e $\beta_{2,t}$ são identificados como l_{jt} e s_{jt} e a taxa de juros para as operações com maturidade igual a τ períodos para cada país, são como descrito

---

[2] Os modelos competidores são: passeio aleatório, Nelson e Siegel com fatores AR(1) dinâmicos, Regressão com taxas *forwad* de Fama e Bliss, Regressão com a curva *forwad* de Cochrane-Piazzesi, Modelos auto-regressivos univariados AR(1) com taxas de juros em nível, VAR com taxas de juros em nível, VAR com taxas de juros em primeira diferença e o modelo de regressão na inclinação.

por:

$$y_t^j(\tau) = l_{jt} + s_{jt}\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} + \zeta_t^j(\tau) \qquad [2.5]$$

onde j denota os diferentes países e $\zeta_t^j(\tau)$ denota a componente do erro de ajuste[3] com desvio padrão igual a $\sigma_j(\tau)$.

A justificativa dos autores para modelar a curva de juros com os fatores l_{jt} e s_{jt}, se referindo somente ao nível e a inclinação respectivamente, é dada da seguinte forma:

- Assumem o parâmetro λ_t igual para todos os países e invariante com o tempo, isto é $\lambda_t^j = \lambda, \forall j, t$. Segundo Diebold e Li (2006), a perda em generalidade em assim assumir, é pequena, principalmente, porque esse parâmetro é determinado quando a função que descreve a curvatura atinge o seu máximo (como descrito no tópico 2.2).

- A baixa precisão na estimação do fator curvatura, que se deveu a ausência de dados relativos a maturidades muito curtas ou muito longas das curvas de juros, na maioria dos países compreendidos no estudo.

- Segundo mostrado por DL, a curvatura não possui ligações claras com fundamentos macroeconômicos.[4]

Neste artigo, a curva de juros de cada país é determinada por [2.5], mas passam a considerar que os fatores l_{jt} e s_{jt} possuem leis de movimento próprias, onde a sua dinâmica é governada por forças globais L_t e S_t e locais $\varepsilon_{1,t}^j$ e $\varepsilon_{2,t}^j$ como descritas abaixo:

$$l_{jt} = \alpha_j^l + \theta_j^l L_t + \varepsilon_{jt}^l \qquad [2.6]$$

$$s_{jt} = \alpha_j^s + \theta_j^s S_t + \varepsilon_{jt}^s \qquad [2.7]$$

Onde α_j^l e α_j^s são constantes, θ_j^l e θ_j^s são os *loadings* dos fatores globais, ε_{jt}^l e ε_{jt}^s são as forças específicas. As suposições referentes a essas equações são:

---

[3] Na equação utilizada por Diebold, Li e Yue para a modelagem da curva de juros, aparece o componente que se refere ao erro de ajuste, pelo fato da equação compor um modelo de espaço de estado.

[4] Simplificação que não é defendida por outros autores, como no trabalho seminal de **Litterman e Sheinkman (1991)**, onde modelam os fatores latentes: nível, inclinação e curvatura associando-os a variáveis macroeconômicas.

• Devido à inclusão da constante nas equações [2.6] e [2.7], assume-se que o componente idiossincrático possui média igual a zero.

• Assume que as inovações referentes aos fatores globais tem desvio padrão igual a 1 (σ^n=1 para n=l,s), porque a magnitude dos fatores globais e seus coeficientes não são identificados separadamente.

• Adota-se que o *loading* do fator global referente aos EUA é positivo (θ_{US}^n > 0, para n=l,s), para identificar os sinais dos fatores e dos coeficientes referentes aos mesmos.

Deve-se reforçar que a hipótese fundamental do modelo de DLY está centrada nas equações [2.6] e [2.7], onde cada fator, l_{jt} e s_{jt}, depende de uma força global, que afeta todos os países conjuntamente, e de determinantes idiossincráticos que afetam cada país individualmente.

Neste *framework*, como mostrado em Diebold e Li (2006), os fatores globais de cada país e os componentes locais seguem processos auto-regressivos de ordem 1:

$$\begin{bmatrix} L_t \\ S_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} L_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} U_t^l \\ U_t^s \end{bmatrix} \quad [2.8]$$

$$\begin{bmatrix} \varepsilon_{jt}^l \\ \varepsilon_{jt}^s \end{bmatrix} = \begin{bmatrix} \varphi_{j,11} & \varphi_{j,12} \\ \varphi_{j,21} & \varphi_{j,22} \end{bmatrix} \begin{bmatrix} \varepsilon_{j,t-1}^l \\ \varepsilon_{j,t-1}^s \end{bmatrix} + \begin{bmatrix} u_t^l \\ u_t^s \end{bmatrix} \quad [2.9]$$

Em relação a essas equações são assumidas as seguintes hipóteses para a estimação do modelo:

• Os choques U e u possuem média zero e são independentes entre si.

• O vetor de choques $\begin{bmatrix} U_t^l & U_t^s \end{bmatrix}^{tr}$ possui matriz de variância e covariância Σ_U igual à matriz identidade, uma vez que esses choques não apresentam qualquer tipo de correlação serial.

• Os choques relativos ao componente idiossincrático, dado pelo vetor $\begin{bmatrix} u_t^l & u_t^s \end{bmatrix}^{tr}$ não apresentam qualquer tipo de correlação serial e as perturbações relativos ao país j são independentes das perturbações relativas a outros países, assim as matrizes de variância e covariância Σ_u são diagonais.

• Nas equações [2.8] e [2.9], $\phi_{mn} = 0, \forall m \neq n$ e $\varphi_{mn} = 0, \forall m \neq n$. Essa hipótese é bastante importante, porque além de reduzir a quantidade de parâmetros a serem estimados, indica também que a dinâmica de um fator, tanto ele global, como específico, é influenciado apenas por ele mesmo defasado. Isso pode ser explicado mais claramente da seguinte forma: a dinâmica do fator S_t não é influenciada pelo movimento das demais forças L_t, ε^l_{jt}, ε^s_{jt} e sim, apenas por S_{t-1} associado a um choque U^s_t.

O processo de estimação utilizado pelos autores é o método em dois estágios:

- Primeiro Estágio: Estimação do modelo dado pela equação [2.5] por mínimos quadrados ordinários, com $\lambda = 0{,}0609$, obtendo separadamente os fatores l_{jt} e s_{jt} (nível, inclinação) para cada país.
- Segundo Estágio: Tratando os fatores l_{jt} e s_{jt} estimados no estágio anterior como observados e, usando o filtro de Kalman, estimaram o modelo de espaço de estado composto pelas equações de medição [2.6], [2.7] e os sistemas de equação de estado [2.8] e [2.9]. Nessa etapa obtiveram as trajetórias dos super fatores[5] que guiam à dinâmica dos componentes nível e inclinação.

Diebold, Li e Yue destacam três resultados principais. O primeiro refere-se à capacidade do modelo de evidenciar a hipótese inicial, de que existem fatores globais que conduzem a dinâmica da taxa de juros nos diferentes países, ressaltando também que esses são economicamente importantes parecendo estar associados a fundamentos macroeconômicos globais, como inflação e nível de atividade. DLY mostram que os fatores l_{jt} e s_{jt} possuem a eles associados, super fatores que são responsáveis de forma relevante pelo seu movimento, revelando o poder explicativo de 91%[6] para a trajetória do nível e 50% para da inclinação. O segundo resultado que os autores destacam, é que a variação na estrutura a termo americana explicada pelo fator global é menor em todas as maturidades, o que é consistente com a independência relativa do mercado americano. Por fim, o

[5] São assim denominados super fatores, pelo fato de em uma escala hierárquica aparecem como os principais componentes, porque guiam o movimento dos fatores, que por sua vez conduzem a dinâmica da curva de juros

[6] Esse resultado pode ser verificado na Tabela 3, página 355 do referido artigo.

terceiro refere-se à capacidade de explicação do fator global ser crescente com a maturidade, o que na prática significa que o fator nível torna-se mais importante quanto maior a maturidade, o que enfatiza a importância das expectativas de inflação nos preços títulos de longo prazo.