

## CADERNOS DO IME – Série Estatística

Universidade do Estado do Rio de Janeiro - UERJ  
ISSN on-line 2317-4536 / ISSN impresso 1413-9022 - v.57, p.27-51, 2024  
DOI: 10.12957/cadest.2024.92391

# MODELAGEM E ESTIMAÇÃO DA DINÂMICA DE PREÇOS DO PETRÓLEO: UMA ANÁLISE DO MODELO DE TRÊS FATORES

Fernando Antonio Lucena Aiube  
Faculdade de Ciências Econômicas, Universidade do Estado do Rio de Janeiro  
[faiube@gmail.com](mailto:faiube@gmail.com)

Edison Americo Huarsaya Tito  
Instituto de Matemática e Estatística, Universidade Estadual do Rio de Janeiro  
[edison.tito@ime.uerj.br](mailto:edison.tito@ime.uerj.br)

Javier Gutierrez Castro  
Departamento de Engenharia de Produção e Sistemas, Universidade Federal de Santa Catarina  
[javier.gutierrez@ufsc.br](mailto:javier.gutierrez@ufsc.br)

### **Resumo**

*O presente trabalho investiga a dinâmica dos preços do petróleo Brent utilizando um modelo de três fatores, com base no modelo proposto por Bhar e Lee (2011). O estudo emprega dados semanais de contratos futuros de petróleo Brent, abrangendo o período de janeiro de 2006 a agosto de 2022, e utiliza a metodologia do Filtro de Kalman para a estimação dos parâmetros do modelo e dos fatores latentes não observáveis. Os resultados indicam a presença de dois fatores de curto prazo com distintas velocidades de reversão à média e um fator de equilíbrio de longo prazo com dinâmica de passeio aleatório. A análise destaca as vantagens do modelo de três fatores em relação a modelos mais parcimoniosos, particularmente sua capacidade aprimorada de capturar a complexa estrutura a termo dos preços futuros e de decompor os prêmios de risco associados a diferentes componentes da dinâmica dos preços. Discute-se a trajetória dos fatores latentes estimados em relação a eventos históricos e aprofunda-se a interpretação dos prêmios de risco. O estudo conclui que a estrutura de três fatores oferece uma representação robusta e flexível para a análise dos preços do petróleo, fornecendo subsídios valiosos para pesquisadores e analistas de mercado.*

**Palavras-chave:** Preços futuros; Modelo de três fatores; Filtro de Kalman; Commodities.

## 1. Introdução

O petróleo é uma das commodities mais importantes e estrategicamente relevantes na economia global. Suas flutuações de preço exercem um impacto profundo em diversos setores, desde os custos de transporte e produção industrial até as políticas macroeconômicas e a inflação em nível mundial (HAMILTON, 2009; KILIAN, 2009). Para o Brasil, um país com significativa produção e um mercado consumidor relevante, a compreensão da dinâmica dos preços do petróleo é crucial tanto para o planejamento energético e decisões de investimento por parte das empresas quanto para a formulação de políticas públicas, como evidenciado pelas recorrentes discussões sobre a política de preços da Petrobras e seus impactos socioeconômicos.

A modelagem dos preços do petróleo, no entanto, é uma tarefa complexa devido à miríade de fatores que influenciam sua oferta e demanda. Eventos geopolíticos, decisões da Organização dos Países Exportadores de Petróleo (OPEP) e seus aliados (OPEP+), o ritmo do crescimento econômico global, avanços tecnológicos em exploração e produção, e até mesmo o desenvolvimento de fontes de energia alternativas e políticas de transição energética contribuem para a notória volatilidade e para os padrões de comportamento, por vezes complexos, observados nos preços (FATTOUH, 2010; ADELMAN, 1990). Nesse contexto, o desenvolvimento de modelos econométricos robustos que possam capturar as principais características estilizadas dos preços do petróleo, como reversão à média, volatilidade variante no tempo e a estrutura a termo dos preços futuros, torna-se fundamental.

A literatura sobre modelos em commodities, e do petróleo em particular, evoluiu consideravelmente nas últimas décadas. Modelos iniciais, frequentemente baseados em processos de um único fator, como o Movimento Geométrico Browniano (MGB) ou modelos de reversão à média como o de Vasicek (1977), mostraram-se limitados em sua capacidade de explicar a complexa estrutura a termo dos preços futuros, que reflete as expectativas do mercado sobre os preços futuros em diferentes horizontes de tempo (SCHWARTZ, 1997). Modelos de múltiplos fatores surgiram como uma alternativa mais promissora, permitindo uma representação mais rica da dinâmica dos preços. Dentre estes, o modelo de dois fatores de Schwartz e Smith (2000) tornou-se uma referência, decompondo o preço à vista (spot) em um componente de equilíbrio de longo prazo e um desvio de curto prazo. Esta abordagem permitiu uma melhor caracterização do fenômeno

do *convenience yield* – o benefício de manter o estoque físico da commodity – e sua influência na relação entre preços spot e futuros.

Contudo, mesmo modelos de dois fatores podem não capturar todas as nuances da dinâmica dos preços, especialmente em mercados caracterizados por diferentes horizontes de reversão à média ou pela influência de múltiplos componentes estocásticos no curto prazo. A curva de futuros do petróleo frequentemente exhibe formas complexas, como *humps* ou múltiplas inflexões, que são desafiadoras para modelos mais simples. Nesse sentido, o modelo de três fatores, como o proposto por Bhar e Lee (2011) – uma extensão do trabalho de Schwartz e Smith (2000) – oferece uma estrutura ainda mais flexível. Ao incorporar dois fatores de curto prazo com distintas velocidades de reversão e um fator de equilíbrio de longo prazo, este modelo possui o potencial de melhor se ajustar aos dados observados de preços futuros e fornecer uma compreensão mais detalhada dos componentes que direcionam os preços do petróleo.

Este artigo tem como objetivo principal estimar e analisar o modelo de três fatores para os preços do petróleo Brent, utilizando dados semanais de contratos futuros. Especificamente, busca-se: (i) calibrar os parâmetros do modelo de Bhar e Lee (2011) utilizando a metodologia do Filtro de Kalman; (ii) interpretar os fatores latentes estimados e seus componentes dinâmicos, correlacionando-os com eventos históricos relevantes; (iii) analisar os prêmios de risco associados a cada fator; e (iv) discutir as vantagens desta abordagem multifatorial em relação a modelos mais parcimoniosos na representação da dinâmica dos preços do petróleo e da sua estrutura a termo. A análise se concentrará na capacidade do modelo em descrever os dados históricos, sem a intenção de realizar previsões de preços.

O artigo está organizado da seguinte forma: a Seção 2 apresenta uma revisão da literatura sobre modelos em commodities, com ênfase nos modelos de fatores para o petróleo e na evolução histórica dessas abordagens. A Seção 3 detalha a especificação do modelo de três fatores e a metodologia de estimação baseada no Filtro de Kalman, incluindo a derivação do preço futuro. A Seção 4 descreve os dados utilizados e apresenta os resultados da estimação dos parâmetros do modelo. A Seção 5 discute os resultados obtidos, com foco na interpretação dos fatores, incluindo uma análise de suas trajetórias históricas, dos prêmios de risco, e nas vantagens do modelo de três fatores. Finalmente, a Seção 6 apresenta as conclusões do estudo, limitações e sugestões para pesquisas futuras.

## 2. Revisão da literatura sobre modelos de commodities

A modelagem da dinâmica dos preços de commodities, em especial do petróleo, tem sido um campo fértil de pesquisa em finanças e economia por várias décadas. A compreensão dessas dinâmicas é crucial para diversas aplicações, como a avaliação de projetos de investimento no setor, gestão de risco, formulação de políticas energéticas e negociação de derivativos. Historicamente, a literatura tem evoluído de modelos mais simples para estruturas mais complexas, buscando capturar as características empíricas observadas nos mercados, como a reversão à média, a volatilidade e a estrutura a termo dos preços futuros.

### 2.1. Modelos de um fator

Os primeiros modelos estocásticos para apreçamento de ativos e derivativos de commodities frequentemente se baseavam em um único fator de incerteza. O Movimento Geométrico Browniano (MGB), popularizado por Samuelson (1965) para a análise de preços futuros, assume que os logaritmos dos preços seguem um passeio aleatório com um *drift* constante. Embora simples e tratável analiticamente, o MGB implica que a variância dos preços cresce linearmente com o tempo e não incorpora a característica de reversão à média, um fenômeno amplamente documentado para os preços da maioria das commodities (PINDYCK, 1999; BESSEMBINDER *et al.*, 1995). A reversão à média sugere que os preços tendem a retornar a um nível de equilíbrio de longo prazo após choques de oferta ou demanda, refletindo a resposta da produção e do consumo a desvios de preço.

Para incorporar a reversão à média, modelos baseados no processo de Ornstein-Uhlenbeck, como o proposto por Vasicek (1977) para taxas de juros e adaptado por Schwartz (1997) para commodities, foram introduzidos. Nestes modelos, o preço (ou seu logaritmo) reverte a um nível médio constante com uma determinada velocidade. Embora capturem a reversão à média, os modelos de um fator com média constante ainda enfrentam dificuldades em explicar a complexa estrutura a termo dos preços futuros, especialmente a ocorrência simultânea de *contango* (preços futuros acima do preço spot esperado) e *backwardation* (preços futuros abaixo do preço spot esperado) em diferentes segmentos da curva.

Uma inovação importante foi a introdução explícita do conceito de *convenience yield* estocástico por Brennan e Schwartz (1985) e posteriormente elaborado por Gibson

e Schwartz (1990). O *convenience yield* representa os benefícios não monetários de manter o estoque físico da commodity (como a capacidade de evitar custos de parada de produção ou de lucrar com gargalos temporários de oferta). Se o *convenience yield* é suficientemente alto, pode levar a uma situação de *backwardation*. Modelar o *convenience yield* como um processo estocástico, também seguindo uma dinâmica de reversão à média, permitiu uma melhor representação da estrutura a termo. No entanto, um único fator de incerteza (seja o preço spot ou o *convenience yield*) ainda se mostrava restritivo para capturar toda a dinâmica conjunta do preço spot e da forma da curva de futuros.

## 2.2. Modelos de dois fatores

Reconhecendo as limitações dos modelos de um fator, a literatura evoluiu para modelos de múltiplos fatores. Um dos modelos de dois fatores mais influentes e amplamente adotados é o de Schwartz e Smith (2000), que propõem que o logaritmo do preço spot do petróleo é a soma de dois componentes estocásticos: (a) um fator de equilíbrio de longo prazo ( $\chi_t$ ), que segue um movimento geométrico browniano (ou um processo com tendência), representando as mudanças persistentes no nível de equilíbrio dos preços em função de elementos como avanços tecnológicos, esgotamento de reservas e alterações estruturais na demanda; e (b) um desvio de curto prazo ( $\xi_t$ ), que segue um processo de Ornstein-Uhlenbeck, revertendo a zero, e que captura as flutuações temporárias em torno do nível de equilíbrio de longo prazo, provocadas por choques transitórios de oferta e demanda, restrições logísticas ou variações sazonais.

A vantagem do modelo de Schwartz e Smith (2000) reside na sua capacidade de gerar uma estrutura a termo mais rica e consistente com os dados observados. A interação entre o componente de longo prazo (não estacionário) e o de curto prazo (estacionário e com reversão à média) permite ao modelo explicar diferentes formatos da curva de futuros, incluindo *contango* e *backwardation*, e até mesmo curvas com um *hump*. Este modelo tornou-se uma referência para a indústria e para a academia. Outras abordagens de dois fatores incluem, por exemplo, modelos que especificam dinâmicas para o preço spot e para o *convenience yield* separadamente (PILIPOVIC, 2007), ou para o preço spot e a taxa de juros de curto prazo, embora a última seja menos comum para commodities onde a taxa de juros tem um papel secundário em relação aos fundamentos de oferta e demanda da própria commodity. Cortazar e Naranjo (2006) também propuseram um

modelo de dois fatores (generalizável para  $N$  fatores) onde um representa o preço de equilíbrio de longo prazo e o outro o *convenience yield*, ambos seguindo processos de reversão à média.

### 2.3. Modelos de três fatores

Embora os modelos de dois fatores representem um avanço significativo, pesquisas subsequentes e a observação empírica dos mercados de futuros de petróleo sugeriram que fatores adicionais poderiam ser necessários para capturar completamente a dinâmica dos preços e, especialmente, a forma da estrutura a termo. A curva de futuros pode exibir múltiplas inflexões ou diferentes velocidades de reversão para diferentes maturidades, o que pode ser difícil de acomodar com apenas dois fatores estocásticos.

Nesse contexto, Bhar e Lee (2011) estenderam o modelo de Schwartz e Smith (2000) para um modelo de três fatores. No modelo proposto por Bhar e Lee (2011), o logaritmo do preço spot é expresso como a soma de: (a) dois fatores de curto prazo ( $X_{1,t}$  e  $X_{2,t}$ ), cada um seguindo um processo de Ornstein-Uhlenbeck, com possíveis diferenças nas velocidades de reversão à média ( $\kappa_1$  e  $\kappa_2$ ) e nas volatilidades ( $\sigma_1$  e  $\sigma_2$ ). A inclusão de dois fatores de curto prazo proporciona maior flexibilidade na modelagem das flutuações transitórias e na configuração da curva de futuros em seus segmentos mais curtos; e (b) um fator de equilíbrio de longo prazo ( $X_{3,t}$ ), que segue um processo com tendência (semelhante ao  $\chi_t$  de Schwartz e Smith), representando a trajetória de longo prazo dos preços.

A principal vantagem teórica do modelo de três fatores de Bhar e Lee (2011) é sua capacidade aprimorada de ajustar a estrutura a termo dos preços futuros. Ao permitir duas dinâmicas distintas de curto prazo, o modelo pode capturar, por exemplo, um componente de reversão muito rápida (associado a choques de curtíssimo prazo) e outro de reversão mais lenta (associado a choques de médio prazo), ambos orbitando em torno de uma tendência de longo prazo. Essa flexibilidade adicional é particularmente útil em mercados voláteis e complexos como o do petróleo. Outros modelos de três fatores podem envolver diferentes combinações de variáveis estocásticas; por exemplo, Schwartz (1997) discute a possibilidade de um modelo de três fatores para o preço spot, o *convenience yield* e a taxa de juros de curto prazo. Litzenberger e Rabinowitz (1995) também propuseram um modelo de três fatores para explicar a estrutura a termo, focando nos fluxos de dividendos (ou *convenience yields*).

## 2.4. Outras abordagens

Além dos modelos de fatores em tempo contínuo, que são o foco deste trabalho, a literatura também explorou outras abordagens para modelar os preços das commodities e suas volatilidades. Modelos da família GARCH (*Generalized Autoregressive Conditional Heteroskedasticity*), introduzidos por Engle (1982) e Bollerslev (1986), são amplamente utilizados para capturar a volatilidade variante no tempo e o agrupamento de volatilidade (*volatility clustering*) observados nas séries de retornos de commodities. Modelos com saltos (*jump-diffusion*), baseados no trabalho de Merton (1976), foram adaptados para commodities para acomodar movimentos abruptos e descontínuos de preços, frequentemente causados por eventos inesperados como crises geopolíticas ou desastres naturais (CLEWLOW e STRICKLAND, 2000). Modelos de volatilidade estocástica, como o de Heston (1993) adaptado para commodities, permitem que a própria volatilidade siga um processo estocástico, oferecendo uma representação mais rica da dinâmica da incerteza.

Mais recentemente, com o avanço das técnicas computacionais, modelos baseados em aprendizado de máquina e inteligência artificial têm sido explorados para a previsão de preços de commodities, embora a interpretabilidade econômica desses modelos possa ser um desafio em comparação. No entanto, os modelos de fatores, especialmente aqueles que podem ser estimados via Filtro de Kalman a partir de dados de futuros, permanecem populares devido à sua fundamentação teórica na ausência de arbitragem, sua interpretabilidade econômica em termos de componentes de curto e longo prazo, e sua capacidade de caracterizar a estrutura a termo. Este trabalho se concentra na aplicação e análise do modelo de três fatores de Bhar e Lee (2011), buscando avaliar sua capacidade de descrever a dinâmica dos preços do petróleo Brent e destacar suas vantagens em relação a estruturas mais simples.

## 3. Modelo

Para capturar a complexa dinâmica dos preços do petróleo e sua estrutura a termo, este estudo adota o modelo de três fatores proposto por Bhar e Lee (2011), que é uma extensão do influente modelo de dois fatores de Schwartz e Smith (2000). A estimação dos parâmetros deste modelo, que envolve variáveis não observáveis (latentes), é realizada utilizando a abordagem do Filtro de Kalman. Esta seção detalha a especificação do modelo de três fatores e o procedimento de estimação.

### 3.1. O Modelo de três fatores para os preços do petróleo

Seja  $(\Omega, \mathcal{F}, \mathbb{P})$  o espaço de probabilidade, onde  $\Omega$  é o espaço amostral,  $\mathcal{F}$  é a  $\sigma$ -álgebra e  $\mathbb{P}$  é a medida de probabilidade. Admita que o mercado é livre de arbitragem e também completo. Portanto, existe uma única medida martingale equivalente (MME)  $\mathbb{Q}$ .

O modelo de três fatores assume que o logaritmo natural do preço à vista (spot) do petróleo, denotado por  $\ln(X_t)$ , é a soma de três componentes estocásticos não observáveis: dois fatores de curto prazo,  $X_{1,t}$  e  $X_{2,t}$ , e um fator de equilíbrio de longo prazo,  $X_{3,t}$ . Portanto, tem-se:

$$\ln X_t = X_{1,t} + X_{2,t} + X_{3,t} \quad (1)$$

As dinâmicas destes fatores sob a medida  $\mathbb{P}$  são:

$$dX_{1,t} = -\kappa_1 X_{1,t} dt + \sigma_1 dB_{1,t} \quad (2)$$

$$dX_{2,t} = -\kappa_2 X_{2,t} dt + \sigma_2 dB_{2,t} \quad (3)$$

$$dX_{3,t} = \mu dt + \sigma_3 dB_{3,t} \quad (4)$$

Onde  $t \in [0, T_j]$  e  $T_j$  é o vencimento do  $j$ -ésimo contrato futuro;  $\kappa_1, \kappa_2$  são as velocidades de reversão à média;  $\sigma_1, \sigma_2, \sigma_3$  são parâmetros de volatilidade (constantes) dos respectivos fatores;  $\mu$  é o *drift* do fator de longo prazo; e  $dB_{i,t}$ ,  $i = 1, 2, 3$ , são incrementos do processo Browniano padrão (ou processo de Wiener padrão) correlacionados, com  $dB_{i,t} dB_{k,t} = \rho_{i,k} dt$ , para  $i \neq k$ .

Para precificar contratos futuros é necessário transitar da medida física  $\mathbb{P}$  para uma medida martingale equivalente  $\mathbb{Q}$  (medida neutra ao risco). Sob a medida  $\mathbb{Q}$ , os preços futuros descontados são martingais, e o *drift* esperado de qualquer ativo negociável é a taxa de juros livre de risco. A mudança de medida é realizada ajustando os *drifts* dos processos estocásticos pelos respectivos preços de mercado do risco.

Sob a MME  $\mathbb{Q}$ , as dinâmicas dos fatores são:

$$dX_{1,t} = (-\kappa_1 X_{1,t} - \alpha_1) dt + \sigma_1 d\tilde{B}_{1,t} \quad (5)$$

$$dX_{2,t} = (-\kappa_2 X_{2,t} - \alpha_2) dt + \sigma_2 d\tilde{B}_{2,t} \quad (6)$$

$$dX_{3,t} = (\mu - \alpha_3) dt + \sigma_3 d\tilde{B}_{3,t} \quad (7)$$

Onde  $\alpha_1$  e  $\alpha_2$  são os preços de mercado do risco associados aos fatores de curto prazo  $X_{1,t}$  e  $X_{2,t}$ , respectivamente. Eles refletem o prêmio exigido pelos investidores para

manterem o risco associado a esses fatores não observáveis;  $\alpha_3$  é o preço de mercado do risco associado ao fator de equilíbrio de longo prazo  $X_{3,t}$ ; e  $d\tilde{B}_{i,t}$ ,  $i = 1, 2, 3$ , são incrementos de processos Brownianos (ou processos de Wiener padrão) sob a medida  $\mathbb{Q}$ . A notação  $d\tilde{B}_{i,t}$  será utilizada para indicar explicitamente que estes processos estão definidos sob a MME  $\mathbb{Q}$ , mantendo a mesma estrutura de correlação  $\rho_{i,k}$  que os  $dB_{i,t}$  sob a medida física  $\mathbb{P}$ .

O preço de um contrato futuro de petróleo no tempo  $t$  com vencimento em  $T_j$ , denotado por  $F_{t,T_j}$ , é dado pelo valor esperado do preço spot no vencimento,  $S_{T_j}$ , sob a medida  $\mathbb{Q}$ , condicionado à informação disponível em  $t$ , ou seja,  $F_{t,T_j} = E^{\mathbb{Q}}(S_{T_j} | \mathcal{F}_t)$ .

Utilizando as dinâmicas sob  $\mathbb{Q}$  e a relação  $\ln X_t = X_{1,t} + X_{2,t} + X_{3,t}$ , pode-se derivar uma expressão para o logaritmo do preço futuro:

$$\ln F_{t,T_j} = X_{1,t}e^{-\kappa_1(T_j-t)} + X_{2,t}e^{-\kappa_2(T_j-t)} + X_{3,t} + H(T_j - t) \quad (8)$$

Sendo  $T_j - t$  o tempo até o vencimento e  $\mu^* = \mu - \alpha_3$  o *drift* ajustado ao risco do fator de longo prazo, assim, o termo  $H(T_j - t)$  é dado por:

$$\begin{aligned} H(T_j - t) = & -\frac{\alpha_1}{\kappa_1} \left(1 - e^{-\kappa_1(T_j-t)}\right) - \frac{\alpha_2}{\kappa_2} \left(1 - e^{-\kappa_2(T_j-t)}\right) \\ & + \mu^*(T_j - t) + \frac{\sigma_1^2}{4\kappa_1} \left(1 - e^{-2\kappa_1(T_j-t)}\right) + \frac{\sigma_2^2}{4\kappa_2} \left(1 - e^{-2\kappa_2(T_j-t)}\right) \\ & + \frac{\sigma_1\sigma_3\rho_{1,3}}{\kappa_1} \left(1 - e^{-\kappa_1(T_j-t)}\right) + \frac{\sigma_2\sigma_3\rho_{2,3}}{\kappa_2} \left(1 - e^{-\kappa_2(T_j-t)}\right) \\ & + \frac{\sigma_1\sigma_2\rho_{1,2}}{\kappa_1 + \kappa_2} \left(1 - e^{-(\kappa_1+\kappa_2)(T_j-t)}\right) \end{aligned} \quad (9)$$

$H(T_j - t)$  é um termo determinístico que incorpora o impacto dos *drifts* ajustados ao risco e das variâncias/covariâncias dos fatores sobre o preço futuro. Note que os termos  $\alpha_1/\kappa_1$  e  $\alpha_2/\kappa_2$  representam as médias de longo prazo dos fatores  $X_{1,t}$  e  $X_{2,t}$  sob a medida  $\mathbb{Q}$ , ajustadas pelo preço de mercado do risco.

O modelo completo é, portanto, definido pelo conjunto de parâmetros  $\Lambda = \{\kappa_1, \sigma_1, \alpha_1, \kappa_2, \sigma_2, \alpha_2, \sigma_3, \mu^*, \mu, \rho_{1,2}, \rho_{1,3}, \rho_{2,3}, \sigma_{s_j}\}$ . O parâmetro  $\sigma_{s_j}$  é o desvio padrão da diferença entre o logaritmo dos preços futuros observados ( $F_{t,T_j}$ ) e os preços modelados.

### 3.2. Especificação do filtro de Kalman

Nesta seção, apresenta-se a especificação do filtro de Kalman utilizado no procedimento de estimação. Uma vez que o modelo esteja estimado e o vetor  $\mathbf{\Lambda}$  seja conhecido, o processo do preço à vista pode ser simulado. O uso do filtro de Kalman em econometria pode ser encontrado em Harvey (1989), Durbin e Koopman (2001) e Shumway e Stoffer (2006), entre outros. Seguindo a mesma notação de Shumway e Stoffer (2006),  $\mathbf{x}_{t+1}^t = E(\mathbf{x}_{t+1}|\mathbf{Y}_t)$ , onde  $\mathbf{Y}_t$  é o conjunto de informações das observações  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t\}$  no instante  $t$ . Mais especificamente, considere  $\mathbf{y}_t$  uma série temporal multivariada representada por um vetor  $q \times 1$  de variáveis observáveis, e  $\mathbf{x}_t$  um vetor de estado  $p \times 1$  de variáveis não observáveis (ou variáveis latentes). Desta forma,  $q$  representa o número de variáveis observáveis e  $p$  o número de variáveis latentes. Ambas as variáveis  $\mathbf{y}_t$  e  $\mathbf{x}_t$  estão relacionadas através de uma matriz  $q \times p$ ,  $\mathbf{A}_t$ , descrita como:

$$\mathbf{y}_t = \mathbf{A}\mathbf{x}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t \quad (10)$$

Onde  $\mathbf{d}_t$  é um vetor  $q \times 1$ ,  $\boldsymbol{\varepsilon}_t \sim N_q(\mathbf{0}, \mathbf{R})$ , e  $\mathbf{R}$  é uma matriz de covariância  $q \times q$ . A equação (10) é chamada de equação de medição ou equação de observação. As variáveis não observáveis estão relacionadas a um processo autorregressivo através de uma matriz  $p \times p$ ,  $\Phi$ , e são expressas como:

$$\mathbf{x}_{t+1} = \Phi\mathbf{x}_t + \mathbf{c}_t + \Theta\mathbf{v}_t \quad (11)$$

Onde  $\mathbf{c}_t$  é um vetor  $p \times 1$ ,  $\Theta$  é uma matriz  $p \times m$ ,  $\mathbf{v}_t \sim N_m(\mathbf{0}, \mathbf{Q})$ , e  $\mathbf{Q}$  é uma matriz de covariância  $m \times m$ . Desta forma,  $m$  é a ordem da distribuição normal multivariada. A equação (11) é chamada de equação de transição. O vetor inicial  $\mathbf{x}_0$  é definido como  $\mathbf{x}_0 \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ . Tanto os erros  $\boldsymbol{\varepsilon}_t$  quanto  $\mathbf{v}_t$  são independentes de  $\mathbf{x}_0$ .

As equações (10) e (11) definem a formulação em espaço de estados da especificação canônica do filtro de Kalman.

Considerando  $T^* = qn$ , onde  $q$  é número de variáveis observáveis (ou o número de contratos) multiplicado pelo tamanho da amostra para cada contrato, pode-se calcular a função de verossimilhança do modelo como:

$$L(\mathbf{\Lambda}; \mathbf{y}) = \prod_{t=1}^{T^*} f_Y(\mathbf{y}_t | \mathbf{Y}_{t-1}) \quad (12)$$

Onde  $f_Y(\cdot | \mathbf{Y}_{t-1})$  é a função densidade multivariada no instante  $t$ , condicional à informação disponível até  $\mathbf{Y}_{t-1}$ .

Os hiperparâmetros do modelo são estimados através da estimação por máxima verossimilhança (MLE) do erro de predição  $\varepsilon_t$ , ou seja:

$$\hat{\Lambda} = \arg \max_{\Lambda} L(\Lambda; \mathbf{y}) \quad (13)$$

Utilizando as equações (10)-(11), calcula-se o vetor de estado previsto  $\mathbf{x}_t^{t-1}$ , o vetor de estado filtrado  $\mathbf{x}_t^t$ , e as variâncias correspondentes  $\mathbf{P}_t^{t-1}$  e  $\mathbf{P}_t^t$ . Uma vez que a convergência do algoritmo seja alcançada, os valores suavizados do estado podem ser obtidos através do algoritmo de suavização de Kalman. Para mais detalhes sobre o algoritmo de suavização de Kalman, veja em Shumway e Stoffer (2006).

#### 4. Dados e resultados da estimação do modelo

Esta seção descreve os dados utilizados para a estimação do modelo de três fatores para os preços do petróleo e apresenta os resultados da estimação pelo Filtro de Kalman.

##### 4.1. Descrição dos dados

Os dados utilizados neste estudo compreendem os preços de fechamento semanais de contratos futuros de petróleo bruto do tipo Brent. O Brent é uma referência (benchmark) crucial para o mercado global de petróleo, influenciando o apetrechamento de uma parcela significativa da produção mundial. A escolha de dados semanais busca um equilíbrio entre capturar a dinâmica dos preços e evitar o ruído excessivo presente em dados de frequência mais alta (diária), além de mitigar problemas de liquidez que podem afetar contratos com vencimentos mais distantes em dados diários.

A amostra consiste em um painel completo com seis contratos futuros com diferentes prazos de vencimento. Para cada contrato, foram coletadas 863 observações semanais, cobrindo o período de 25 de janeiro de 2006 a 3 de agosto de 2022. A utilização de múltiplos contratos futuros é essencial para a estimação de modelos de fatores, pois a estrutura a termo (a relação entre os preços futuros e seus prazos de vencimento) contém informações valiosas sobre os fatores latentes que direcionam os preços spot. Para evitar distorções causadas pela baixa liquidez ou turbulência próxima ao vencimento dos contratos, os preços de cada série de futuros foram rolados para o contrato subsequente três dias úteis antes da data de expiração do contrato vigente.

A Tabela 1 apresenta as principais estatísticas descritivas dos preços futuros (em US\$/barril) para os seis contratos utilizados na amostra.

Tabela 1 – Estatísticas descritivas dos preços futuros do petróleo Brent (US\$/barril)

|               | $F_{t,T_1}$ | $F_{t,T_5}$ | $F_{t,T_{10}}$ | $F_{t,T_{15}}$ | $F_{t,T_{20}}$ | $F_{t,T_{24}}$ |
|---------------|-------------|-------------|----------------|----------------|----------------|----------------|
| Mínimo        | 20,37       | 29,38       | 33,49          | 35,86          | 37,69          | 39,16          |
| Máximo        | 144,36      | 146,11      | 146,92         | 145,99         | 145,11         | 144,50         |
| Média         | 77,02       | 77,20       | 77,12          | 76,82          | 76,42          | 76,10          |
| Desvio Padrão | 25,24       | 23,74       | 22,55          | 21,63          | 20,85          | 20,31          |
| Assimetria    | 0,29        | 0,35        | 0,39           | 0,41           | 0,43           | 0,45           |
| Curtose       | -0,96       | -0,87       | -0,77          | -0,70          | -0,63          | -0,56          |

Nota:  $F_{t,T_j}$  refere-se ao preço futuro no tempo  $t$  para o contrato com  $j$ -ésima maturidade

Fonte: os autores

Observa-se na Tabela 1 que o preço médio dos contratos futuros durante o período analisado situou-se em torno de US\$ 77,00 por barril, com uma variação considerável, refletindo a volatilidade intrínseca ao mercado de petróleo. O desvio padrão tende a diminuir para contratos com vencimentos mais longos, o que é consistente com a teoria de que a incerteza de curto prazo é maior do que a de longo prazo (efeito Samuelson, 1965). A assimetria positiva indica que a distribuição dos preços possui uma cauda direita mais longa, sugerindo que grandes altas de preços, embora menos frequentes, podem ocorrer. Todas as séries de log-retornos (as estatísticas não são apresentadas por questão de brevidade) são caracterizadas por excesso de curtose (leptocurtose), indicando a presença de eventos extremos.

#### 4.2. Resultados da estimação do modelo

As equações (5)-(9) escritas na forma espaço-estado permite a estimação do Filtro de Kalman. Os hiperparâmetros estimados são apresentados na Tabela 2, juntamente com seus erros padrão.

Tabela 2 – Resultados da estimação dos parâmetros do modelo de três fatores

| Parâmetro    | Valor     | Erro padrão |
|--------------|-----------|-------------|
| $\kappa_1$   | 2,202***  | 0,153       |
| $\sigma_1$   | 0,143***  | 0,044       |
| $\alpha_1$   | 0,276     | 0,314       |
| $\kappa_2$   | 0,491***  | 0,011       |
| $\sigma_2$   | 0,297***  | 0,024       |
| $\alpha_2$   | -0,238*** | 0,094       |
| $\sigma_3$   | 0,247***  | 0,008       |
| $\mu^*$      | -0,034*** | 0,003       |
| $\mu$        | -0,024    | 0,189       |
| $\rho_{1,3}$ | 0,260***  | 0,036       |
| $\rho_{1,2}$ | 0,275***  | 0,136       |
| $\rho_{2,3}$ | 0,102***  | 0,026       |

Nota: \*\*\*, indica nível de significância de 1%.

Fonte: os autores

Analisando os resultados da Tabela 2, observa-se que os parâmetros de reversão à média ( $\kappa_1$ ,  $\kappa_2$ ) para os fatores de curto prazo são ambos positivos e estatisticamente significativos. O valor estimado de  $\kappa_1$  (2,202) é consideravelmente maior que o de  $\kappa_2$  (0,491), o que sugere que o primeiro fator de curto prazo ( $X_{1,t}$ ) apresenta uma velocidade de reversão à média muito superior à do segundo fator ( $X_{2,t}$ ). Essa diferença indica a existência de dinâmicas distintas no curto prazo: enquanto ( $X_{1,t}$ ) pode capturar choques de curtíssimo prazo com efeitos transitórios, ( $X_{2,t}$ ) parece refletir desequilíbrios de médio prazo, que se dissipam mais lentamente.

Em relação às volatilidades dos fatores ( $\sigma_1$ ,  $\sigma_2$ ,  $\sigma_3$ ), todos os parâmetros estimados são positivos e estatisticamente significativos. A volatilidade associada ao segundo fator de curto prazo  $\sigma_2$  (0,297) é a mais elevada entre os três fatores, seguida pela volatilidade do fator de longo prazo  $\sigma_3$  (0,247) e pela do primeiro fator de curto prazo  $\sigma_1$  (0,143). Esses resultados sugerem que os choques de médio prazo, ainda que mais persistentes, têm impacto imediato mais intenso do que os choques de curtíssimo ou longuíssimo prazo.

Quanto aos preços de mercado do risco ( $\alpha_1$ ,  $\alpha_2$ ), observa-se que  $\alpha_1$  (0,276), associado ao primeiro fator de curto prazo, não é estatisticamente significativo. Em contrapartida,  $\alpha_2$  (-0,238), relacionado ao segundo fator de curto prazo, é negativo e estatisticamente significativo. Esse resultado implica que os agentes de mercado podem estar dispostos a pagar um prêmio para se proteger contra os riscos associados ao segundo fator, ou que esse fator está correlacionado negativamente com outras fontes sistemáticas de risco, o que gera um prêmio negativo.

O *drift* do fator de longo prazo também merece destaque. O valor ajustado ao risco, representado por  $\mu^* = \mu - \alpha_3$ , foi estimado como negativo (-0,034) e estatisticamente significativo. Isso sugere uma tendência de queda, embora moderada, no componente de equilíbrio de longo prazo dos preços do petróleo sob a medida neutra ao risco. Já o *drift* sob a medida física,  $\mu$  (-0,024), também foi negativo, porém sem significância estatística. Essa ausência de significância é comum em modelos financeiros de séries temporais, dado que a detecção robusta de tendências estruturais de longo prazo frequentemente exige amostras muito extensas (CAMPBELL, LO e MACKINLAY, 1997).

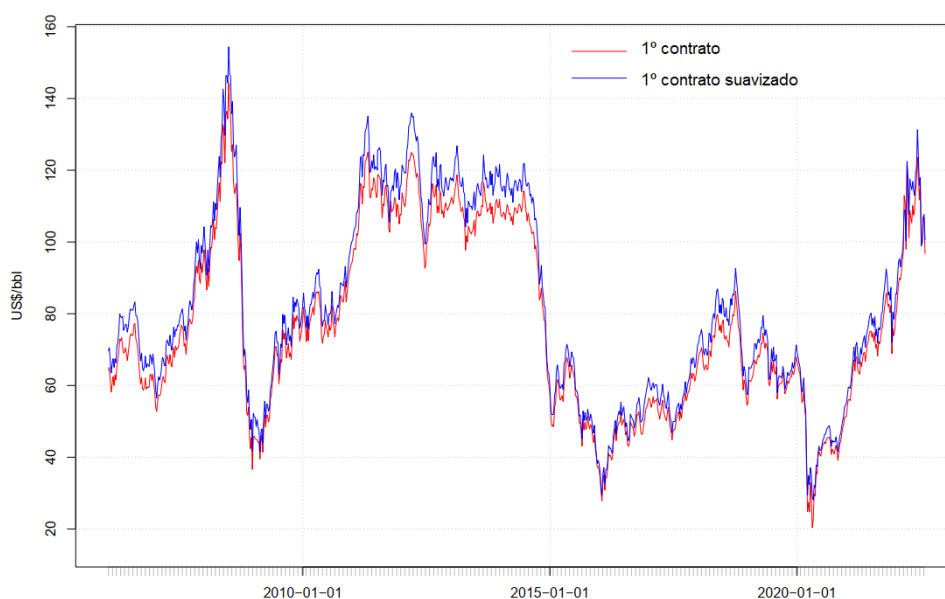
Por fim, os coeficientes de correlação ( $\rho_{1,2}$ ,  $\rho_{1,3}$ ,  $\rho_{2,3}$ ) entre os incrementos dos processos de Wiener que impulsionam os fatores latentes são todos positivos e estatisticamente significativos. A correlação entre os dois fatores de curto prazo  $\rho_{1,2}$  (0,275) sugere que choques que afetam um deles tendem a afetar o outro na mesma direção. As correlações entre os fatores de curto prazo e o fator de longo prazo  $\rho_{1,3}$  (0,260) e  $\rho_{2,3}$  (0,102) também indicam que os componentes estocásticos de diferentes horizontes temporais não são totalmente independentes, o que é coerente com a natureza interligada dos choques no mercado de petróleo.

Os erros padrão para os termos de erro de medição  $\sigma_{s_j}$ , (assumidos como parte da matriz  $\mathbf{R}$  e estimados implicitamente, omitidos aqui para fins de brevidade), possuem valores muito pequenos, indicando um bom ajuste do modelo aos preços futuros observados nos diferentes vencimentos.

Para visualizar o ajuste do modelo, a Figura 1 exibe a série temporal do preço do primeiro contrato futuro e os preços respectivos suavizados oriundos da estimação e do filtro de suavização (*smoothing*).

A inspeção visual da Figura 1 sugere que o modelo de três fatores, com os parâmetros estimados, consegue acompanhar de perto a trajetória dos preços futuros observados, capturando tanto as tendências gerais quanto as flutuações de mais alta frequência.

Figura 1 – Preço observado do primeiro contrato futuro de petróleo Brent e o preço suavizado



Fonte: Os autores, 2025

## 5. Análise e discussão dos resultados

Os resultados da estimação do modelo de três fatores para os preços do petróleo Brent, apresentados na Seção 4, fornecem *insights* valiosos sobre a complexa dinâmica que rege este mercado. Esta seção se aprofunda na interpretação desses resultados, com foco especial em como a estrutura de três fatores contribui para uma melhor representação das características observadas nos preços do petróleo e na sua estrutura a termo. Adicionalmente, analisa-se a trajetória histórica dos fatores latentes estimados e discute-se mais a fundo os prêmios de risco.

### 5.1. Interpretação dos fatores latentes e seus parâmetros

A principal contribuição do modelo de Bhar e Lee (2011) reside na decomposição do logaritmo do preço spot em três fatores latentes com distintas características dinâmicas: dois fatores de curto prazo ( $X_{1,t}$  e  $X_{2,t}$ ) e um fator de equilíbrio de longo prazo ( $X_{3,t}$ ).

A estimação revelou velocidades de reversão à média significativamente diferentes para os dois fatores de curto prazo:  $\kappa_1$  (2,202) e  $\kappa_2$  (0,491). Isso implica que  $X_{1,t}$  possui uma meia-vida (tempo para reverter metade do caminho de volta à média) de aproximadamente  $\ln(2)/2,202 \approx 0,31$  anos (cerca de 4 meses), enquanto  $X_{2,t}$  tem uma meia-vida de  $\ln(2)/0,491 \approx 1,41$  anos. Modelos com apenas um fator de curto prazo (como o de Schwartz e Smith (2000), que possui um único  $\xi_t$ ) são forçados a ajustar uma única velocidade de reversão para todos os tipos de choques transitórios. No entanto, o mercado de petróleo é sujeito a uma variedade de distúrbios, desde interrupções operacionais de curtíssimo prazo em refinarias ou oleodutos, que podem ser rapidamente resolvidas, até desequilíbrios sazonais ou flutuações na demanda de médio prazo devido a ciclos econômicos menores. O modelo de três fatores, ao permitir duas velocidades de reversão, é capaz de dissociar esses efeitos.  $X_{1,t}$  pode estar capturando os choques mais efêmeros e de rápida dissipação, enquanto  $X_{2,t}$  pode representar desvios mais persistentes, mas ainda transitórios, do equilíbrio de longo prazo. Essa capacidade de modelar diferentes persistências no curto prazo é crucial para explicar a forma da curva de futuros, especialmente em seus segmentos mais curtos, onde a sensibilidade a diferentes tipos de notícias e eventos é maior. Por outro lado, a volatilidade de  $X_{2,t}$  ( $\sigma_2 = 0,297$ ) é maior que a de  $X_{1,t}$  ( $\sigma_1 = 0,143$ ), indicando que os choques de médio prazo (representados por

$X_{2,t}$ ) têm uma magnitude de impacto instantâneo maior, embora revertam mais lentamente.

O fator de equilíbrio de longo prazo ( $X_{3,t}$ ), com sua dinâmica de passeio aleatório com *drift* ( $\mu^*$  sob a medida  $\mathbb{Q}$ ), representa o componente não estacionário dos preços do petróleo. A estimativa de  $\mu^* = -0,034$ , estatisticamente significativa, sugere uma leve tendência de declínio no preço de equilíbrio de longo prazo sob a medida neutra ao risco durante o período analisado. Embora o *drift* sob a medida física ( $\mu = -0,024$ ) não tenha sido significativo, a própria natureza de “passeio aleatório” de  $X_{3,t}$  implica que choques neste fator são permanentes, alterando a trajetória de longo prazo dos preços. A presença deste fator é fundamental para que o modelo possa gerar curvas de futuros que não sejam necessariamente planas no longo prazo. Em modelos que possuem apenas fatores de reversão à média (*mean-reverting*), a curva de futuros tenderia a convergir para uma constante no longo prazo. O fator  $X_{3,t}$  permite que o “nível de ancoragem” da curva de futuros seja ele próprio estocástico e sujeito a uma tendência.

O preço de mercado do risco  $\alpha_1$  não foi significativo, sugerindo que o risco associado ao fator  $X_{1,t}$  (choques de curtíssimo prazo) pode não ser sistematicamente precificado pelos investidores, ou que sua estimação é mais ruidosa. Por outro lado,  $\alpha_2 = -0,238$ , significativo, para o fator  $X_{2,t}$  (choques de médio prazo) implica um prêmio de risco negativo. Um prêmio de risco negativo para um fator de commodity pode surgir se esse fator tiver propriedades de hedge contra outros riscos no mercado (por exemplo, se ele for negativamente correlacionado com o retorno do mercado agregado) ou se a oferta da commodity for relativamente inelástica e os produtores estiverem dispostos a aceitar preços futuros menores para garantir a venda de sua produção (teoria da normal backwardation de Keynes). A diferença entre  $\mu^*$  (-0,034) e  $\mu$  (-0,024) permite inferir o preço de mercado do risco para o fator de longo prazo  $X_{3,t}$ , resultando em  $\alpha_3 = \mu - \mu^* \approx 0,010$ . Um  $\alpha_3$  positivo indicaria que os investidores exigem um prêmio para suportar o risco de longo prazo nos preços do petróleo. Assim, a capacidade de estimar prêmios de risco distintos para diferentes fatores latentes é outra vantagem de modelos multifatoriais. Isso permite uma representação mais realista de como os investidores percebem e precificam os diferentes tipos de incerteza que afetam os preços das commodities.

As correlações positivas estimadas entre todos os fatores ( $\rho_{1,2}, \rho_{1,3}, \rho_{2,3}$ ) indicam que os diferentes componentes da dinâmica dos preços não são independentes. Choques

positivos no fator de longo prazo ( $X_{3,t}$ ) tendem a ser acompanhados por choques positivos nos fatores de curto prazo ( $X_{1,t}$  e  $X_{2,t}$ ), e vice-versa. Isso é economicamente plausível: por exemplo, um anúncio de uma grande descoberta de petróleo (afetando  $X_{3,t}$  negativamente no longo prazo) pode, no curtíssimo prazo, não ter impacto ou até mesmo aliviar pressões altistas momentâneas. No entanto, de forma geral, os fatores tendem a se mover em conjunto, embora com diferentes magnitudes e persistências. A correlação  $\rho_{1,2} = 0,275$  entre os dois fatores de curto prazo sugere que, embora distintos, eles compartilham algumas fontes comuns de variação.

## 5.2. Trajetória histórica das variáveis latentes estimadas

Uma vantagem crucial da estimação via Filtro de Kalman é a capacidade de extrair as trajetórias temporais dos fatores latentes não observáveis ( $X_{1,t}$ ,  $X_{2,t}$  e  $X_{3,t}$ ) a partir dos dados dos preços futuros.

### 5.2.1. Fator de equilíbrio de longo prazo ( $X_{3,t}$ )

Espera-se que a trajetória de  $X_{3,t}$  reflita as mudanças mais persistentes e estruturais nas expectativas sobre os preços do petróleo. Durante a crise financeira de 2008–2009, por exemplo, a queda abrupta nos preços do petróleo provavelmente se refletiria em uma queda acentuada em  $X_{3,t}$ , indicando uma revisão para baixo nas expectativas de demanda de longo prazo e, possivelmente, nos custos marginais de produção.

A ascensão do petróleo de xisto (*shale oil*), especialmente entre 2010 e 2014, representou uma mudança estrutural na oferta global, em grande parte devido ao aumento da produção nos Estados Unidos. Esse fator teria exercido uma pressão baixista ou, ao menos, moderadora sobre  $X_{3,t}$ , ao alterar as perspectivas de equilíbrio de longo prazo para os preços do petróleo.

No período pós-2014, a decisão da OPEP de não cortar a produção em resposta ao crescimento da oferta de fora do cartel contribuiu para uma nova queda nos preços do petróleo. Essa movimentação deveria ser visível como uma tendência de queda ou, no mínimo, de estagnação na trajetória de  $X_{3,t}$ , refletindo a percepção de que o novo equilíbrio estrutural era mais baixo.

A pandemia de COVID-19, em 2020, provocou um colapso abrupto na demanda global por petróleo, e essa disrupção teria levado a uma queda drástica em  $X_{3,t}$ . Esse

movimento representaria a incerteza quanto à recuperação econômica e às mudanças nos padrões de consumo, como o trabalho remoto e a menor mobilidade global, afetando as expectativas de longo prazo.

Já no período de 2021 a 2022, com a recuperação econômica pós-pandemia e, especialmente, os choques de oferta e as preocupações com a segurança energética decorrentes da guerra na Ucrânia, houve provavelmente uma elevação em  $X_{3,t}$ . Essa dinâmica indicaria um novo patamar de preços de equilíbrio de longo prazo percebido pelo mercado, influenciado tanto por fatores geopolíticos quanto por pressões estruturais sobre a oferta.

### **5.2.2. Fatores de curto prazo ( $X_{1,t}$ e $X_{2,t}$ ):**

Os fatores de curto prazo  $X_{1,t}$  e  $X_{2,t}$  capturam as flutuações em torno da trajetória do fator de equilíbrio de longo prazo  $X_{3,t}$ . O fator  $X_{1,t}$ , com sua rápida reversão à média, tende a apresentar picos e vales de curta duração. Ele reflete principalmente a reação a notícias e eventos com impacto imediato e transitório, como interrupções temporárias na produção, divulgações inesperadas de dados semanais de estoque ou oscilações cambiais de curto prazo que afetam o poder de compra.

Por outro lado,  $X_{2,t}$ , que possui uma reversão mais lenta à média, representa desvios mais persistentes, ainda que não estruturais. Durante a crise de 2008, por exemplo, além da queda observada em  $X_{3,t}$ , o fator  $X_{2,t}$  poderia ter assumido valores significativamente negativos e permanecido assim por vários trimestres, refletindo um pessimismo de médio prazo que levou tempo para se dissipar. De modo similar, em fases de otimismo ou em contextos de gargalos de oferta com duração intermediária,  $X_{2,t}$  poderia tornar-se positivo e manter-se nesse patamar por um período prolongado. A diferença observada entre os preços de mercado do risco  $\alpha_1$  (não significativo) e  $\alpha_2$  (negativo e estatisticamente significativo) sugere que os desvios de médio prazo representados por  $X_{2,t}$  são mais sistematicamente precificados pelos investidores do que os choques de curtíssimo prazo capturados por  $X_{1,t}$ .

### **5.3. Vantagens do modelo de três fatores na explicação da estrutura a termo**

A principal vantagem de um modelo de três fatores sobre modelos de um ou dois fatores reside em sua capacidade aprimorada de ajustar e explicar a estrutura a termo dos

preços futuros, ou seja, a curva formada pelos preços de contratos futuros com diferentes vencimentos.

Modelos de um fator geralmente produzem curvas de futuros monotônicas — crescentes (em *contango*) ou decrescentes (em *backwardation*). Já modelos de dois fatores, como o de Schwartz e Smith (2000), conseguem gerar curvas com um único *hump*, em que os preços de médio prazo se encontram acima ou abaixo dos preços de curto e longo prazo. O modelo de três fatores, por sua vez, ao incorporar duas dinâmicas distintas de curto prazo, tem o potencial de capturar formatos de curva ainda mais complexos, com múltiplas inflexões ou *humps* localizados em diferentes regiões da estrutura a termo, quando presentes nos dados. Isso ocorre porque a contribuição de cada fator para o preço futuro (conforme a equação 8) decai exponencialmente com o tempo até o vencimento, a uma taxa determinada por seu respectivo parâmetro  $\kappa$ . Com dois valores distintos de  $\kappa$  nos fatores de curto prazo, o modelo apresenta maior flexibilidade na forma como esses componentes afetam a estrutura da curva ao longo de diferentes horizontes temporais.

Do ponto de vista empírico, espera-se que um modelo com mais fatores, desde que estatisticamente significativos e economicamente interpretáveis, proporcione um melhor ajuste aos dados observados. No presente estudo, os pequenos erros de medição obtidos na estimação, somados à inspeção visual da Figura 1, sugerem que o modelo de três fatores apresenta bom desempenho na replicação dos preços futuros do Brent. Embora não tenha sido realizada uma comparação formal com modelos mais simples neste trabalho, a literatura, como em Bhar e Lee (2011), frequentemente demonstra a superioridade estatística de modelos com maior dimensionalidade, especialmente em termos de ajuste *in-sample* e métricas como o erro quadrático médio.

Além disso, a estrutura de três fatores permite uma decomposição mais rica da volatilidade ao longo da curva de futuros. As volatilidades dos diferentes segmentos da estrutura a termo podem ser explicadas com base nas volatilidades e correlações dos três fatores subjacentes. Contratos com vencimentos mais curtos são mais sensíveis aos fatores de curto prazo ( $\sigma_1$  e  $\sigma_2$ ), enquanto os vencimentos mais longos são predominantemente influenciados pela volatilidade associada ao fator de longo prazo ( $\sigma_3$ ). Essa capacidade de mapear as fontes de volatilidade por horizonte reforça o poder

analítico do modelo e sua utilidade na compreensão da dinâmica dos preços futuros do petróleo.

#### 5.4. Implicações e limitações

A estimação do modelo de três fatores para o petróleo Brent revelou uma dinâmica complexa, na qual choques de curtíssimo, médio e longo prazo interagem para determinar a trajetória dos preços e a forma da curva de futuros. A possibilidade de distinguir entre componentes com diferentes velocidades de reversão à média e prêmios de risco associados oferece uma ferramenta analítica poderosa para compreender o comportamento dos preços no mercado de petróleo.

Apesar disso, algumas limitações metodológicas devem ser consideradas. Uma delas é o pressuposto de normalidade. O modelo assume que os incrementos dos fatores são normalmente distribuídos, o que permite a aplicação ótima do Filtro de Kalman. No entanto, é amplamente documentado que os retornos do petróleo apresentam caudas pesadas (leptocurtose) e, em alguns casos, assimetrias, características que não são totalmente capturadas por distribuições Gaussianas. A incorporação de saltos nos preços, por meio de extensões do tipo *jump-diffusion*, ou a adoção de volatilidade estocástica, poderia melhorar a representação da dinâmica dos retornos, embora isso implicasse em um aumento considerável na complexidade do modelo e da estimação.

Outra limitação está relacionada à suposição de parâmetros constantes. O modelo assume que parâmetros como  $(\kappa, \sigma, \mu, \alpha, \rho)$  permanecem inalterados ao longo do tempo. Na prática, esses parâmetros podem variar em resposta a eventos de instabilidade ou transformações estruturais nos fundamentos do mercado. Alternativas como modelos com mudança de regime (*regime-switching*) ou especificações com parâmetros variantes no tempo oferecem maior flexibilidade, embora também exijam maior sofisticação computacional.

Adicionalmente, o modelo foi estimado com base em preços futuros, o que levanta uma questão importante sobre a relação entre o preço spot teórico derivado do modelo e os preços spot efetivamente observados no mercado. A presença de problemas de liquidez, ruído ou inconsistências na formação do preço spot pode afetar essa comparação, e merece atenção em aplicações práticas.

Por fim, embora os fatores sejam formalmente decompostos no modelo, a atribuição de significados econômicos claros — como choques de oferta, choques de

demanda agregada ou variações específicas de inventário — não é direta. A identificação econômica desses componentes exigiria uma análise complementar, com a incorporação de variáveis macroeconômicas, dados setoriais ou fundamentos de mercado.

Ainda assim, o modelo de três fatores representa um avanço importante na modelagem da dinâmica dos preços do petróleo. Sua estrutura flexível e teoricamente consistente permite uma representação mais completa dos movimentos observados no mercado, oferecendo suporte analítico valioso para pesquisadores, analistas financeiros e gestores de risco que buscam uma compreensão mais profunda das forças que determinam os preços dessa commodity essencial.

## 6. Conclusões

Este artigo investigou a dinâmica dos preços do petróleo Brent através da estimação e análise de um modelo de três fatores, conforme proposto por Bhar e Lee (2011). Utilizando dados semanais de contratos futuros cobrindo o período de janeiro de 2006 a agosto de 2022, e empregando a metodologia do Filtro de Kalman para a estimação dos parâmetros e dos fatores latentes, o estudo buscou fornecer uma compreensão mais aprofundada dos componentes que direcionam os preços desta commodity crucial.

Os resultados da estimação confirmaram a relevância da estrutura de três fatores. Foi identificada a presença de dois fatores de curto prazo com velocidades de reversão à média distintas ( $\kappa_1 = 2,202$  e  $\kappa_2 = 0,491$ ), indicando que o mercado de petróleo é afetado por choques transitórios de diferentes persistências. O primeiro fator de curto prazo ( $X_{1,t}$ ) exibe uma reversão rápida, capturando possivelmente distúrbios de curtíssimo prazo, enquanto o segundo ( $X_{2,t}$ ) reverte mais lentamente, refletindo desequilíbrios de médio prazo. Adicionalmente, um fator de equilíbrio de longo prazo ( $X_{3,t}$ ), com dinâmica de passeio aleatório, foi identificado como o componente que ancora a trajetória de longo prazo dos preços. Os prêmios de risco associados a esses fatores também apresentaram nuances, com o prêmio para o segundo fator de curto prazo sendo negativo e significativo, enquanto o prêmio para o primeiro não se mostrou estatisticamente relevante no período analisado.

A principal contribuição deste estudo, alinhada com a literatura de modelos multifatoriais, reside na demonstração da capacidade aprimorada do modelo de três fatores em representar a complexa dinâmica dos preços do petróleo e, consequentemente,

sua estrutura a termo. As vantagens sobre modelos mais parcimoniosos (um ou dois fatores) são notáveis, como a maior flexibilidade na modelagem da reversão à média: ao permitir duas velocidades distintas de reversão para os componentes de curto prazo, o modelo acomoda de forma mais realista a diversidade de choques que afetam o mercado. Além disso, observa-se um melhor ajuste à estrutura a termo, já que a interação dos três fatores, cada um com sua dinâmica específica, permite capturar formatos mais complexos da curva de futuros, incluindo múltiplas inflexões, o que seria desafiador para modelos com menos fatores. Por fim, destaca-se a decomposição mais rica dos prêmios de risco: a capacidade de estimar prêmios distintos para os diferentes fatores latentes oferece uma visão mais granular de como os investidores precificam as diversas fontes de incerteza no mercado de petróleo.

Visualmente, o modelo demonstrou um bom ajuste aos dados históricos dos preços futuros do Brent, com os preços suavizados pelo modelo acompanhando de perto os preços observados.

Apesar dos *insights* valiosos proporcionados, reconhecemos algumas limitações inerentes à abordagem. O pressuposto de normalidade dos choques e a constância dos parâmetros ao longo do tempo são simplificações que podem não capturar integralmente todas as características do mercado de petróleo, como caudas pesadas nos retornos ou mudanças estruturais.

Para pesquisas futuras, diversos caminhos se apresentam. A incorporação de saltos (*jump-diffusions*) ou processos de volatilidade estocástica poderia aprimorar a captura de movimentos extremos e da dinâmica da volatilidade. A investigação de modelos com parâmetros variantes no tempo ou com mudança de regime (*regime-switching*) poderia oferecer maior flexibilidade para lidar com períodos de instabilidade ou mudanças estruturais no mercado. Adicionalmente, a inclusão de variáveis macroeconômicas ou dados de fundamentos do mercado de petróleo (como níveis de estoque, produção da OPEP, etc.) no processo de estimação poderia auxiliar na interpretação econômica dos fatores latentes e na análise da relação entre os preços do petróleo e a economia global. Uma comparação formal do desempenho do modelo de três fatores com modelos de um e dois fatores, utilizando critérios estatísticos de ajuste e capacidade preditiva (fora do escopo deste trabalho, mas relevante), também constituiria uma contribuição importante.

Em suma, este estudo reforça a utilidade de modelos multifatoriais, especificamente o modelo de três fatores, como uma ferramenta robusta para analisar e entender a dinâmica dos preços do petróleo. A capacidade de decompor os movimentos de preços em componentes com características distintas oferece uma perspectiva mais rica e detalhada, essencial para pesquisadores, analistas e formuladores de políticas que lidam com este mercado fundamental para a economia global e brasileira.

## Referências

- ADELMAN, M. A. Mineral depletion, with special reference to petroleum. **The Review of Economics and Statistics**, v. 72, n. 1, p. 1-10, 1990.
- BESSEMBINDER, H.; COUGHENOUR, J. F.; SEGUIN, P. J.; SMOLLER, M. M. Mean reversion in equilibrium asset prices: Evidence from the futures term structure. **The Journal of Finance**, v. 50, n. 1, p. 361-375, mar. 1995.
- BHAR, R.; LEE, D. Time-varying market price of risk in the crude oil futures market. **Journal of Futures Markets**, v. 31, n. 8, p. 779-807, 2011.
- BOLLERSLEV, T. Generalized autoregressive conditional heteroskedasticity. **Journal of Econometrics**, v. 31, n. 3, p. 307-327, 1986.
- BRENNAN, M. J.; SCHWARTZ, E. S. Evaluating natural resource investments. **The Journal of Business**, v. 58, n. 2, p. 135-157, 1985.
- CAMPBELL, J. Y.; LO, A. W.; MACKINLAY, A. C. **The econometrics of financial markets**. Princeton, NJ: Princeton University Press, 1997.
- CLEWLOW, L.; STRICKLAND, C. **Energy derivatives: pricing and risk management**. London: Lacima Publications, 2000.
- CORTAZAR, G.; NARANJO, L. An N-factor Gaussian model of oil futures prices. **Journal of Futures Markets**, v. 26, n. 3, p. 243-268, 2006.
- DURBIN, J.; KOOPMAN, S. J. **Time series analysis by state space methods**. Oxford: Oxford University Press, 2001.
- ENGLE, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. **Econometrica**, v. 50, n. 4, p. 987-1007, 1982.
- FATTOUH, B. **Oil market dynamics through the lens of the 2002-2009 price cycle**. Oxford: Oxford Institute for Energy Studies, 2010. (Working Paper WPM 39).
- GIBSON, R.; SCHWARTZ, E. S. Stochastic convenience yield and the pricing of oil contingent claims. **The Journal of Finance**, v. 45, n. 3, p. 959-976, 1990.
- HAMILTON, J. D. Causes and consequences of the oil shock of 2007–08. **Brookings Papers on Economic Activity**, v. 40, n. 1 (Spring), p. 215-261, 2009.
- HARVEY, A. C. **Forecasting, structural time series models and the Kalman filter**. Cambridge: Cambridge University Press, 1989.

HESTON, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. **The Review of Financial Studies**, v. 6, n. 2, p. 327-343, 1993.

KILIAN, L. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. **American Economic Review**, v. 99, n. 3, p. 1053-1069, 2009.

LITZENBERGER, R. H.; RABINOWITZ, N. Backwardation in commodity futures markets: Theory and tests. **Journal of Finance**, v. 50, n. 5, p. 1575-1601, 1995.

MERTON, R. C. Option pricing when underlying stock returns are discontinuous. **Journal of Financial Economics**, v. 3, n. 1-2, p. 125-144, 1976.

PILIPOVIC, D. **Energy risk: valuing and managing energy derivatives**. 2. ed. New York: McGraw-Hill, 2007.

PINDYCK, R. S. The long-run evolution of energy prices. **The Energy Journal**, v. 20, n. 2, p. 1-27, 1999.

SAMUELSON, P. A. Proof that properly anticipated prices fluctuate randomly. **Industrial Management Review**, v. 6, n. 2, p. 41-49, Spring 1965.

SCHWARTZ, E. S. The stochastic behavior of commodity prices: Implications for valuation and hedging. **The Journal of Finance**, v. 52, n. 3, p. 923-973, 1997.

SCHWARTZ, E. S.; SMITH, J. E. Short-term variations and long-term dynamics in commodity prices. **Management Science**, v. 46, n. 7, p. 893-911, 2000.

SHUMWAY, R. H.; STOFFER, D. S. **Time series analysis and its applications with R examples**. 2. ed. New York: Springer, 2006.

VASICEK, O. An equilibrium characterization of the term structure. **Journal of Financial Economics**, v. 5, n. 2, p. 177-188, 1977.

## MODELING AND ESTIMATING OIL PRICE DYNAMICS: AN ANALYSIS OF THE THREE-FACTOR MODEL

### Abstract

*This study investigates the dynamics of Brent crude oil prices using a three-factor model, based on the framework proposed by Bhar and Lee (2011). The analysis employs weekly data from Brent crude oil futures contracts, covering the period from January 2006 to August 2022, and applies the Kalman Filter methodology to estimate the model parameters and the unobservable latent factors. The results indicate the presence of two short-term factors with distinct mean-reversion speeds and one long-term equilibrium factor following a random walk process. The analysis highlights the advantages of the three-factor model over more parsimonious alternatives, particularly its enhanced ability to capture the complex term structure of futures prices and to decompose the risk premiums associated with different components of price dynamics. The estimated trajectories of the latent factors are discussed in relation to historical events, and the interpretation of the risk premiums is further explored. The study concludes that the three-factor framework provides a robust and flexible representation for analyzing oil prices, offering valuable insights for researchers and market analysts.*

**Key-words:** Oil pricing, Three-factor model, Kalman filter, Commodities.