



Universidade do Estado do Rio de Janeiro

Centro de Ciências Sociais

Faculdade de Ciências Econômicas

Bruna Carolina Fiuza Ferreira

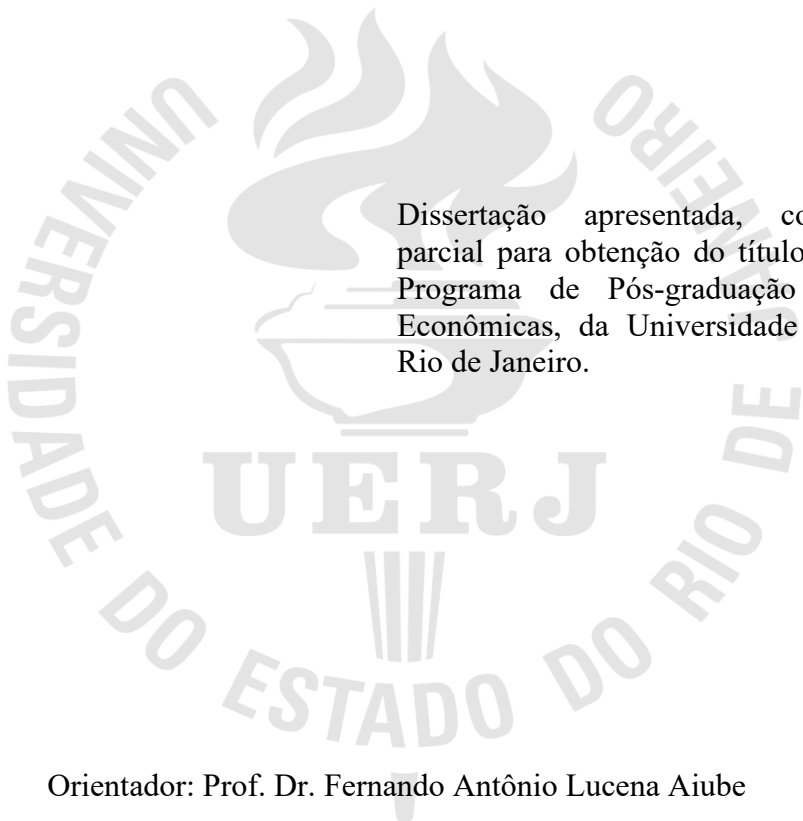
Comparação entre modelos de preços de commodities: dois e três fatores estocásticos

Rio de Janeiro

2019

Bruna Carolina Fiuza Ferreira

Comparação entre modelos de preços de commodities: dois e três fatores estocásticos



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-graduação em Ciências Econômicas, da Universidade do Estado do Rio de Janeiro.

Orientador: Prof. Dr. Fernando Antônio Lucena Aiube

Rio de Janeiro

2019

F363 Ferreira, Bruna Carolina Fiuza
Tese Comparação entre modelos de preços de commodities: dois e três
fatores estocásticos / Bruna Carolina Fiuza Ferreira – 2019.
52 f.

Orientador: Prof. Dr. Fernando Antonio Lucena Aiube.
Dissertação (mestrado) – Universidade do Estado do Rio de Janeiro,
Faculdade de Ciências Econômicas..

1. Bolsa de mercadorias – Teses. 2. Mercado Futuro - Teses. I. Aiube,
Fernando Antonio Lucena. II. Universidade do Estado do Rio de Janeiro.
Faculdade de Ciências Econômicas. III. Título.

CDU 336.764.2

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Data

Bruna Carolina Fiuza Ferreira

Comparação entre modelos de preços de commodities: dois e três fatores estocásticos

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre ao Programa de Pós-graduação em Ciências Econômicas, da Universidade do Estado do Rio de Janeiro.

Aprovada em 30 de agosto de 2019.

Banca Examinadora:

Prof. Dr. Fernando Antonio Lucena Aiube (Orientador)
Faculdade de Ciências Econômicas – UERJ

Prof. Dr. Andrea Ugolini
Faculdade de Ciências Econômicas – UERJ

Prof. Dr. Ariel Levy
Universidade Federal Fluminense

Prof. Dr. Edison Americo Huarsaya Tito
Pontifícia Universidade Católica do Rio de Janeiro

Prof. Dr. Tara Keshar Nanda Baidya
Universidade do Grande Rio Professor José de Souza Herdy

Rio de Janeiro

2019

DEDICATÓRIA

Ao meu filho,
João Victor.

AGRADECIMENTOS

Primeiramente gostaria de agradecer imensamente ao meu orientador Prof. Dr. Fernando Aiube, pela pessoa e profissional que é. Sua dedicação, esforço e paciência foram fundamentais durante a realização desse sonho.

A minha família pelo apoio e esforço empenhado para que eu pudesse superar cada obstáculo em meu caminho e chegar até aqui.

Aos professores Dr. Andrea Ugolini, Dr. Ariel Levy, Dr. Edison Americo Huarsaya Tito e Dr. Tara Keshar Nanda Baidya, membros da banca de Qualificação e Defesa de Mestrado, pelos comentários, sugestões e interesse em contribuir para o desenvolvimento deste projeto.

Por fim, aos meus amigos: Pedro Noronha, Raphael Sousa, Leonardo Galvão, Emily Xavier, Lucas Parente e Ana Pires pelo amor incondicional.

If you only do what you can do, you will never be more than you are now.

Master Shifu-Kung Fu Panda 3

RESUMO

FERREIRA, B.C.F. **Comparação entre modelos de preços de commodities: dois e três fatores estocásticos**. 2019. 39 f. Dissertação (Mestrado em Ciências Econômicas) – Faculdade de Ciências Econômicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Visando aprimorar a previsão de preços de commodities e auxiliar na tomada de decisões dos agentes desse mercado, este estudo propõe um modelo de três fatores estocásticos, expandindo o trabalho de Gibson e Schwartz (1990). Além do preço à vista e do retorno de conveniência, a média do retorno de conveniência é tratada como um fator estocástico. A fim de implementar o modelo, foram coletados dados semanais de preços futuros da *New York Stock Exchange* entre 1996 e 2018. Além disso, a metodologia do filtro de *Kalman* foi empregada para estimar os parâmetros e as séries temporais das variáveis não observáveis. Finalmente, comparamos os modelos de dois e três fatores e o modelo proposto apresentou um poder explicativo superior ao modelo tradicional de dois fatores, capturando com maior fidelidade a estrutura a termo dos preços futuros. Essa performance aprimorada se configura como uma ferramenta para *traders*, gestores de risco e formuladores de políticas, auxiliando na tomada de decisões estratégicas, mitigação de riscos e otimização de resultados.

Palavras – chave: Modelagem de preços de commodities. Mercado Futuro. *Gibson e Schwartz*.

ABSTRACT

FERREIRA, B.C.F. **Comparison of commodity price models: two and three stochastic factors.** 2019. 39 f. Dissertação (Mestrado em Ciências Econômicas) –Faculdade de Ciências Econômicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Aiming to refine commodity price forecasting and assist market agents in decision-making, this study proposes a three-factor stochastic model, extending the work of Gibson and Schwartz (1990). In addition to spot prices and convenience returns, the mean of convenience returns is treated as a stochastic factor. To implement the model, weekly futures price data from the New York Stock Exchange between 1996 and 2018 were collected. Additionally, the Kalman filter methodology was employed to estimate the parameters and time series of the unobserved variables. Finally, we compared the two-factor and three-factor models, and the proposed model exhibited superior explanatory power compared to the traditional two-factor model, more accurately capturing the term structure of futures prices. This enhanced performance positions it as a valuable tool for traders, risk managers, and policymakers, aiding in strategic decision-making, risk mitigation, and outcome optimization.

Keywords: Commodity price modeling. Future markets. *Gibson e Schwartz.*

LISTA DE ILUSTRAÇÕES

Tabela 1-	Estatísticas das séries de preços de contratos futuros do petróleo: 3/1/96 a 6/6/2018.....	24
Gráfico 1-	Preços (US\$/Barril) do contrato F1 do petróleo WTI: 3/1/96 a 6/6/18.....	25
Tabela 2-	Estimação dos hiperparâmetros do modelo de 2 fatores.....	26
Gráfico 2-	Preços F1 observados para o modelo de dois fatores.....	27
Gráfico 3-	Variáveis de estado X_t e δ_t	27
Tabela 3-	Estimação dos hiperparâmetros do modelo de 3 fatores.....	28
Gráfico 4-	Preços F1 observados e filtrados para o modelo de três fatores.....	29
Gráfico 5-	Variáveis de estado X_t , δ_t e α_t	30
Gráfico 6-	Estrutura a termo nos modelos de 2 e 3 fatores: 27/07/2005.....	33
Gráfico 7-	Estrutura a termo nos modelos de 2 e 3 fatores: 27/04/2011.....	33
Gráfico 8-	Estrutura a termo nos modelos de 2 e 3 fatores: 01/10/2018.....	34

SUMÁRIO

INTRODUÇÃO	10
1 REVISÃO BIBLIOGRÁFICA	13
2 MODELOS DE DOIS FATORES	18
3 MODELOS DE TRÊS FATORES	21
4 EXPERIMENTOS NUMÉRICOS	24
4.1. Dados	24
4.2 Resultados empíricos	25
4.3 Comparando os modelos	31
5 APLICAÇÃO EM HEDGE DE COMPROMISSO FORWARD	35
CONCLUSÃO	37
REFERÊNCIAS	38
APÊNDICE A - Modelo De Dois Fatores	40
APÊNDICE B -Modelo de Três Fatores	46

INTRODUÇÃO

A modelagem e estimativa do comportamento estocástico dos preços das commodities permite que produtores e consumidores, além de especuladores financeiros, possam implementar estratégias sólidas de investimento e de gestão de risco.

Geralmente a literatura em finanças considera que o preço à vista das commodities é uma variável não observável. De fato, o preço à vista, resultante de uma transação no mercado físico, não fica definido de imediato. O preço de fechamento ou liquidação (*settlement price*) de uma transação, ocorre quando do recebimento do produto pelo comprador. Por exemplo, se uma transação ocorre na data t , a *commodity* deve ser transportada do seu ponto de origem ao destino. A operação será liquidada em data posterior, ou seja, $t + \Delta t$. Portanto, o valor de liquidação estará associado à data $t + \Delta t$ quando da entrega.

Mais ainda, o valor final da transação não é uma variável prontamente observável pelo mercado, somente é do conhecimento dos envolvidos na negociação do produto. Os preços das negociações são divulgados periodicamente por agências especializadas que coletam as informações junto aos agentes que negociam a *commodity*.

Em geral, o preço de uma *commodity*, divulgado corriqueiramente pela imprensa, é o preço do primeiro contrato futuro negociado, ou seja, com vencimento mais próximo. Assim ocorre com o petróleo, a *commodity* mais relevante para a economia mundial, e com todas as demais que são negociadas em mercados futuros.

O preço futuro da *commodity*, negociado em bolsas, é conhecido diariamente pelos canais de informação do mercado. Além de ser amplamente acessível, o preço futuro é formado competitivamente, mediante inúmeras transações diárias nas bolsas de mercados futuros.

Tendo em vista a relevância do preço à vista para os agentes que negociam a *commodity*, por exemplo no gerenciamento do risco embutido na produção e comercialização, a literatura em finanças relacionada às commodities busca estimá-lo. Para tal, a metodologia empregada, de forma recorrente, tem sido o uso de variáveis de espaço de estado para descrever os preços futuros. A forma de espaço de estado constrói a relação entre um vetor de séries temporais observáveis (preços futuros para diferentes vencimentos) e um vetor de séries não observáveis. As componentes desse vetor não observadas são chamadas de variáveis de estado, estas são geradas por um processo estocástico. A versão em tempo discreto desse

processo é denominada equação de transição. A partir de então, em modelos Gaussianos e afins, o Filtro de *Kalman* é aplicado recursivamente para calcular o estimador ótimo para o vetor de estado no tempo t .

Os modelos de preços das commodities desenvolvidos por Brennan e Schwartz (1985) e Paddock et al. (1988) consideraram apenas uma variável (fator) estocástica, o preço à vista. Embora o modelo de um fator tenha a vantagem de ser muito tratável, ele possui algumas propriedades indesejáveis, como exibir uma estrutura de volatilidade constante dos retornos de preços futuros, ou seja, os preços futuros para vencimentos diferentes estariam perfeitamente correlacionados. Evidências empíricas sugerem, no entanto, que a estrutura de volatilidade dos preços futuros é uma função decrescente da maturidade como documentado por Bessembinder et al. (1995).

Em Gibson e Schwartz (1990), os autores desenvolveram um modelo usando dois fatores estocásticos: o preço à vista e o retorno de conveniência, onde o preço à vista segue um processo geométrico *Browniano* e o retorno de conveniência segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*.

O primeiro fator considerado, o preço à vista, contribui de forma indiscutível para a natureza dos preços futuros. O segundo fator, o retorno de conveniência, foi documentado inicialmente por Kaldor (1939) e descreve a articulação entre os níveis atuais de estoque e a escassez futura de uma mercadoria. A teoria do armazenamento postula uma relação inversa entre o nível de estoque e o retorno de conveniência, o que sugere que esta variável será constante somente sob premissas muito restritivas. Em Gibson e Schwartz (1989) os autores refutam essa suposição no caso do petróleo bruto, mostrando que a tendência de reversão à média, assim como a variabilidade de suas mudanças, requer uma representação estocástica do rendimento de conveniência.

Brennan e Schwartz (1985) definem o rendimento de conveniência como o benefício que possui um proprietário do produto físico, mas não é devido ao proprietário de um contrato para a entrega futura da mercadoria. Se a demanda é superior à oferta, o preço à vista da mercadoria é mais elevado do que o preço futuro e os detentores da commodity física se beneficiam de possuir o bem subjacente. Portanto, o rendimento de conveniência deve ser maior em valor absoluto do que os custos marginais (ou o custo de carregamento) e, nesse caso, o custo de carregamento líquido é negativo. Esta teoria é elaborada por Working (1949), Brennan (1958) e Pindyck (2001) e vincula o retorno de conveniência às variações de curto prazo no preço à vista.

Além do preço à vista e do retorno de conveniência, a incerteza na dinâmica dos preços das commodities pode ser conferida a outros fatores. Schwartz (1997) compara três modelos em relação a aderência aos preços de contratos futuros existentes. O primeiro tem apenas um fator estocástico: o preço à vista. O segundo modelo é essencialmente o modelo definido em Gibson e Schwartz (1990), onde os fatores são o preço à vista e o retorno de conveniência. O terceiro modelo inclui a taxa de juro americana como um fator estocástico.

Assim como nos trabalhos desenvolvidos anteriormente, os preços futuros constituem as variáveis de observação (vetor de séries temporais observáveis) enquanto o preço à vista, o retorno de conveniência e a taxa de juros são as variáveis de estado ou de transição. Os modelos que consideravam dois e três fatores mostraram uma aderência satisfatória a estrutura a termo dos preços e das volatilidades, enquanto, o de um fator mostrou pouca aderência.

Neste trabalho nos propomos a desenvolver uma metodologia para recuperar o preço à vista a partir de contratos futuros com diferentes maturidades. Incluímos no modelo de Gibson e Schwartz (1990) um terceiro fator estocástico. Além do preço à vista e o retorno de conveniência, propomos que a média para qual o retorno de conveniência reverte, parâmetro α , tenha um comportamento estocástico, seguindo um processo de reversão à média. O objetivo é obter um modelo que se ajuste melhor aos preços futuros observados.

Para tanto fez-se necessária a derivação da equação do preço futuro, ou seja, a solução do sistema de equações das variáveis estocásticas que definem os três fatores. Nesta solução o logaritmo dos preços futuros é uma função linear das variáveis subjacentes, permitindo que o modelo seja colocado na forma de espaço de estado. Utilizando-se de preços de contratos futuros de petróleo com diferentes maturidades aplicamos o filtro de *Kalman* para estimar os parâmetros e a série temporal das variáveis de estado não observáveis. A partir daí prosseguimos com a comparação dos modelos de dois e três fatores.

Este trabalho será dividido da seguinte forma: o Capítulo 1 explora a literatura sobre modelagem de preços de commodities e estimativa via filtro de *Kalman*. O capítulo 2 foi dedicado à derivação e ao detalhamento do modelo de dois fatores desenvolvido em Gibson e Schwartz (1990). No capítulo 3 apresentamos a derivação do modelo proposto, com três fatores estocásticos. O capítulo 4 implementamos os dois modelos, apresentamos os dados utilizados, os principais resultados para cada um dos modelos e a comparação entre eles. O capítulo 5 visa mostrar a aplicabilidade dos modelos em estratégias de *hedge* para compromissos *forward*. Por fim, concluímos com os principais resultados e possíveis pesquisas futuras.

1 REVISÃO BIBLIOGRÁFICA

Nesta seção apresentamos uma breve revisão da literatura dos modelos de fatores estocásticos aplicados na precificação de commodities. Esses modelos diferem-se principalmente no papel desempenhado pelo rendimento de conveniência e no número de fatores usados para descrever a incerteza. Os modelos iniciais assumiram um rendimento de conveniência constante e descreveram a incerteza a partir de um único fator, cuja dinâmica estocástica era um movimento *Browniano*.

Um dos trabalhos pioneiros, que compõe esta abundante literatura, é o artigo de Brennan e Schwartz (1985). Afim de considerar a natureza estocástica do preço das commodities, os autores desenvolveram um modelo de avaliação de projetos de investimentos assumindo que o preço à vista segue um processo geométrico *Browniano*, baseando-se na conjectura de que o preço à vista do petróleo tem uma distribuição estacionária log-normal. A relação entre o preço à vista e os preços futuros é desenvolvida considerando a taxa de juros constante e o retorno de conveniência proporcional ao preço à vista. Tal modelo foi aplicado ao gerenciamento (abertura, fechamento e abandono definitivo) de uma mina de cobre.

Em Gibson e Schwartz (1989) os autores analisaram as propriedades de séries temporais dos rendimentos de conveniência e mostraram que a tendência de reversão à média e a variabilidade, no caso do petróleo bruto, requer uma representação estocástica dessa variável. O artigo de Gibson e Schwartz (1990) desenvolve e testa um modelo de preços de commodities com dois fatores estocásticos: o preço à vista e o retorno de conveniência. Diferenciando-se da maioria dos trabalhos desenvolvidos até então, que se baseavam na suposição da existência de uma única fonte de incerteza.

A dinâmica dos fatores é descrita por um processo estocástico conjunto, onde o preço à vista segue um processo geométrico *Browniano* e o retorno de conveniência segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*. Foi realizada uma regressão para as variáveis estocásticas do modelo e os parâmetros foram estimados a partir de preços semanais de contratos futuros de petróleo negociados na *New York Mercantile Exchange* (NYMEX).

No modelo, os preços de contratos futuros relacionam-se às variáveis de estado a partir de uma equação diferencial parcial (EDP) de segunda ordem, cuja solução analítica não era conhecida à época. Dessa forma o prêmio de risco no mercado do retorno de conveniência não foi estimado diretamente. O parâmetro foi estimado através da minimização do erro

quadrático médio, comparando os resultados encontrados com os dados do mercado. Os autores concluem que o preço de um contrato futuro aumentará, tudo o mais constante, com um aumento no preço à vista do petróleo. Porém, uma vez que tal aumento está associado a um maior retorno de conveniência, o acréscimo será atenuado, mesmo para contratos de maturidade mais curta.

Schwartz (1997) compara o desempenho de três modelos de comportamento estocástico dos preços das commodities. O primeiro considera que o preço à vista segue um processo geométrico *Browniano*, sendo o único fator estocástico. O segundo modelo considera a mesma dinâmica para o preço à vista e assume que o rendimento de conveniência segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*, assim como em Gibson e Schwartz (1990). Finalmente, no terceiro, a taxa de juros instantânea também segue um processo de reversão à média.

A metodologia de espaço de estado foi empregada, onde os preços futuros constituem as variáveis de observação enquanto o preço à vista, o retorno de conveniência e taxa de juros são as variáveis de estado ou de transição. Através da aplicação do filtro de *Kalman* os parâmetros foram estimados para duas commodities, petróleo e cobre, e um metal precioso, ouro. O artigo analisa as implicações dos modelos para a estrutura a termo dos preços futuros, das volatilidades e sua aplicabilidade em *hedge* de compromissos *forward*. Conclui-se que a inclusão de um segundo e terceiro fator permite uma melhor aderência na estrutura a termo de preços e volatilidades. Outra conclusão importante é que tanto para o petróleo quanto para o cobre a análise revela forte reversão à média, já para o ouro, o autor sugere que os modelos são imprecisos.

Schwartz e Smith (2000) desenvolveram seu trabalho a partir da decomposição do preço à vista em dois fatores estocásticos. O primeiro fator captura os desvios de curto prazo no preço à vista e segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*. O segundo fator captura as mudanças no preço de equilíbrio e sua dinâmica é um processo geométrico *Browniano*. Embora não considere explicitamente as mudanças no rendimento de conveniência ao longo do tempo, dada uma combinação apropriada de parâmetros, esse modelo de curto prazo / longo prazo equivale ao desenvolvido em Gibson e Schwartz (1990), com a vantagem de que o entendimento de mudanças de preço de curto e longo prazo é mais intuitivo.

Os parâmetros foram calculados a partir de preços de contratos futuros de petróleo. Os autores utilizam o filtro de *Kalman* na estimação do modelo, que é aplicado a alguns ativos hipotéticos ligados à *commodity*.

Nos artigos de Schwartz (1997) e Schwartz e Smith (2000) fica bem caracterizado que a inclusão de um fator estocástico, que representa o retorno de conveniência, contribui para uma melhor aderência dos modelos aos preços futuros observados.

Pilipovic (1997) supõe que o preço à vista de curto prazo segue um processo estocástico de reversão média e reverte para o preço de equilíbrio de longo prazo, este é considerado log-normalmente distribuído, seguindo um processo geométrico *Browniano*. O autor desenvolve uma forma funcional para o rendimento de conveniência e faz uso das condições de fronteira e de não arbitragem para desenvolver uma solução analítica do modelo que é aplicado aos mercados de energia, gás natural e petróleo.

Pindyck (1999) analisa séries históricas de preços de petróleo bruto, gás natural e carvão. Em relação ao petróleo, o autor conclui que tanto a tendência quanto o nível dos preços são estocásticos, portanto, uma modelagem onde o preço à vista reverte para uma tendência estocástica tem um melhor desempenho para fins de previsão. A estimação dos parâmetros procede com o uso do filtro de *Kalman*.

Dias e Rocha (1999) modela o preço do petróleo a partir de um processo estocástico de reversão à média de um fator com saltos. O trabalho sugere que o preço do recurso evolui revertendo para uma média, mas pode apresentar mudanças bruscas (ou saltos) com determinada probabilidade. O preço tende a convergir para o seu nível de equilíbrio de longo prazo à medida que os efeitos de tais choques desaparecem.

Sørensen (2002) introduz no modelo de Schwartz e Smith (2000) parâmetros de sazonalidade afim de analisar o comportamento das commodities agrícolas (milho, soja e trigo). A partir da formulação de espaço de estado e a estimação via filtro de *Kalman*, o autor conclui que os parâmetros sazonais são significativamente diferentes de zero e contribuem para ajustar o modelo aos dados de futuros de commodities agrícolas.

Lucia e Schwartz (2002) comparam modelos de um e dois fatores estocásticos para o preço à vista da eletricidade com relação a sua capacidade de explicar o comportamento das curvas de preço futuro e *forward*. Além dos fatores estocásticos, inclui-se um componente determinístico que leva em consideração as regularidades nos preços à vista da *commodity*. Os autores concluem por um melhor desempenho do modelo de dois fatores em relação ao de um fator, além da importância de considerar o padrão sazonal dos preços para explicar a estrutura a termo dos preços futuros.

Em Manoliu e Tompaidis (2002) os autores definem o preço à vista como uma função de variáveis subjacentes que seguem processos estocásticos e uma variável determinística que descreve a sazonalidade nos preços. Usando dados de contratos futuros de gás natural, dois

modelos são estimados com a formulação via espaço de estado e implementados com o uso do filtro de *Kalman*. O primeiro modelo é composto por um fator determinístico e um fator estocástico, que segue um processo de reversão à média. O segundo inclui no modelo anterior um outro fator estocástico, que segue um processo *Browniano*. A estimação permitiu a captura do padrão de sazonalidade presente nos dados e conclui-se por um melhor ajuste do modelo de dois fatores em relação aos erros de previsão nos preços futuros e na capacidade de capturar a estrutura de correlação observada.

Geman e Nguyen (2005) propõe uma abordagem de variável de estado para a dinâmica dos preços da soja, usando sucessivamente dois e três fatores. No primeiro modelo os fatores são o preço à vista e a média de curto prazo de seu componente não sazonal. As sazonalidades no preço à vista e na volatilidade são controladas por meio de funções determinísticas. No segundo modelo a escassez (definida como o inverso dos estoques) é incluída como uma terceira variável de estado. Conclui-se que a volatilidade é uma função crescente da escassez e que a introdução dessa variável no modelo de fatores melhora o ajuste comparativamente ao modelo de dois fatores. A estimação dos parâmetros foi feita com a implementação do filtro de *Kalman*.

Cortazar e Naranjo (2006) analisam a capacidade de um modelo gaussiano de N -fatores explicar o comportamento estocástico dos preços futuros do petróleo. Ao contrário das abordagens tradicionais de agregação de dados para um conjunto de vencimentos, o modelo é aplicado para todas as transações diárias de preço de futuros de petróleo disponíveis no período de análise. Um procedimento de estimativa via filtro de *Kalman*, que permite condições incompletas dos dados, é usado para calibrar o modelo. Os autores concluíram que são necessários pelo menos três fatores para explicar a estrutura a termo dos preços futuros e quatro fatores para explicar adequadamente a estrutura a termo da volatilidade.

Hikspoors e Jaimungal (2007) desenvolvem uma abordagem para o preço de commodities energéticas considerando um processo de dois fatores estocásticos de reversão à média e uma função determinística que captura os padrões de sazonalidade, posteriormente há a inclusão de saltos no modelo. Assume-se que o preço à vista segue um processo de reversão à média para um nível de equilíbrio estocástico. A abordagem é similar ao modelo de dois fatores proposto em Pilipovick (1997), diferenciando-se pelo processo estocástico assumido para o preço de equilíbrio.

Pelo lado empírico, Bessembinder et al. (1995) usa dados de preços de contratos futuros com horizontes de entrega variáveis para testar se os investidores esperam um comportamento de reversão nos preços dos ativos. O autor apresenta evidências de um

comportamento de reversão para os preços das commodities agrícolas e do petróleo bruto. O mesmo acontece em Pindyck (1999); nesse trabalho, o autor examina o comportamento de longo prazo dos preços do petróleo, do carvão e do gás natural com o intuito de identificar a dinâmica estocástica da evolução dos preços e conclui que estes estão de fato revertendo, mas a taxa de reversão à média é lenta e flutua ao longo do tempo.

Apesar das evidências empíricas sugerirem que o preço à vista do petróleo evolui de acordo com um movimento de reversão à média, Postali e Picchetti (2006) argumentam que é possível usar o Movimento Geométrico *Browniano* como uma proxy para modelar a dinâmica de preços e aproveitar sua facilidade operacional sem que isso incorra em erros de avaliação significativos. A rejeição de tal processo ocorre somente com longas séries face a baixa velocidade de reversão. Por outro lado, a rejeição pode ocorrer quando são omitidas as quebras estruturais.

Os modelos de preços para commodities mencionados até o momento possuem condições adequadas para a estimação. Trata-se de modelos Gaussianos e afins. Nestas circunstâncias o filtro de *Kalman* é a metodologia adequada para a estimação. As referências clássicas sobre filtro de *Kalman* na econometria são os textos de Harvey (1989), Hamilton (1994) e Durbin e Koopman (2002). A metodologia tem aplicação variada em economia, por exemplo em Pasricha (2006).

Em finanças o filtro de *Kalman* tem sido aplicado para estimar variáveis de estado de modelos de preços de commodities, como por exemplo nos trabalhos desenvolvidos por Schwartz (1997), Schwartz e Smith (2000), Manoliu e Tompaidis (2002) e Sørensen (2002), já citados acima. Além destes, podemos acrescentar Babbs and Nowman (1999) que utilizam o filtro de *Kalman* na estimação do modelo de Vacisek para taxa de juros.

Os modelos Gaussianos foram estendidos para abrigar mais adequadamente o comportamento empírico dos preços. Por exemplo, a inclusão de saltos torna o modelo não gaussiano. Neste caso, o filtro de *Kalman* não é aplicado em sua forma canônica. Villaplana (2003) usa o modelo de dois fatores de Schwartz e Smith (2000) e inclui a componente de saltos no fator de curto prazo, o modelo é aplicado aos preços da energia elétrica. Aiube et al. (2008) segue esta modelagem e utiliza o filtro de partículas na estimação, aplicando o modelo aos preços futuros do petróleo, conclui-se que o componente de saltos melhora aderência aos dados empíricos. De forma análoga, a inclusão da volatilidade como uma variável estocástica faz com que o modelo perca as propriedades Gaussianas. Javaheri et al. (2003) apresentam o modelo de volatilidade estocástica fazendo uso do filtro de partículas.

2 MODELO DE DOIS FATORES

Neste capítulo apresentamos o modelo desenvolvido por Gibson e Schwartz (1990). Os autores desenvolveram uma abordagem de precificação de commodities considerando o uso de dois fatores. Diferente dos modelos desenvolvidos até então, que se baseavam na suposição de que o preço à vista seria a única fonte de incerteza na dinâmica dos preços de commodities, o comportamento estocástico dos preços é descrito por um processo conjunto: além do preço à vista, inclui-se um rendimento de conveniência instantâneo estocástico.

O preço à vista segue um processo geométrico *Browniano* e o retorno de conveniência segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*. Derivamos uma solução de forma fechada para os preços, de tal forma que o logaritmo dos preços futuros está linearmente relacionado aos fatores subjacentes

Optou-se pela utilização de tal modelo, pois as variáveis estocásticas estão diretamente modeladas e estão bem fundamentadas na literatura relacionada às commodities. Embora outros modelos de fatores também pudessem ser aplicados ao presente trabalho como, por exemplo, o modelo de dois fatores de Schwartz (2000).

Dado o espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$, que descreve a incerteza na economia. Ω é o espaço de eventos, $\mathcal{F} = \{\mathcal{F}\}_{0 \leq t \leq T}$ representa a filtração natural dos *Brownianos* associados às incertezas dos processos e $\mathbb{P}$ a medida de probabilidade real ou histórica. No modelo desenvolvido por Gibson e Schwartz (1990) a dinâmica dos preços é descrita por duas equações diferenciais estocásticas para as variáveis (ou fatores): preço à vista, S_t , e retorno de conveniência, δ_t , cujos processos estão definidos abaixo, sob a medida real de probabilidade $\mathbb{P}$.

$$dS_t = (\mu - \delta_t)S_t dt + \sigma_1 S_t dB_{1,t} \quad (1)$$

O primeiro fator estocástico S_t segue um processo geométrico *Browniano* baseado no pressuposto de que o preço à vista do petróleo tem uma distribuição estacionária log-normal. Definindo $X_t = \ln S_t$ e usando o Lema de Itô, temos que X_t tem a dinâmica abaixo:

$$dX_t = \left(\mu - \delta_t - \frac{1}{2} \sigma_1^2 \right) dt + \sigma_1 dB_{1,t} \quad (2)$$

Em que μ representa o retorno total de longo prazo esperado pelo proprietário da *commodity* (valorização do preço mais o retorno de conveniência), $\sigma_1 > 0$ é a volatilidade de X_t , $dB_{1,t}$ é o incremento do processo *Browniano* padrão para X_t e δ_t é o retorno de conveniência instantâneo cuja dinâmica é dada pelo processo:

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_2 dB_{2,t} \quad (3)$$

O segundo fator δ_t evolui segundo um processo de reversão a média do tipo *Ornstein-Uhlenbeck*. Na equação (3) temos que $\kappa > 0$ é a velocidade de reversão a média e α é a média de longo prazo do retorno de conveniência. $\sigma_2 > 0$ é a volatilidade de δ_t e $dB_{2,t}$ é o incremento do processo *Browniano* padrão para δ_t .

Além disso, os incrementos dos processos *Brownianos* padrão $dB_{1,t}$ de S_t e $dB_{2,t}$ de δ_t são correlacionados por ρ , tal que $dB_{1,t}dB_{2,t} = \rho dt$.

Seja $t \in [0, T]$ em que T representa o vencimento do contrato futuro, o preço futuro pode ser escrito como função das duas variáveis estocásticas e da data de vencimento do contrato futuro, ou seja $F(S, \delta, T)$. Ou ainda, como função dos dois fatores de risco, ou seja, $F_{t,T}(S_t, \delta_t)$.

Assumindo que o preço de um contrato futuro $F_{t,T}$ é uma função duas vezes continuamente diferenciável de S_t e δ_t , podemos usar Lema de Itô para mostrar que a solução que define o preço de um contrato futuro entregue no momento T é dada a partir da solução da equação diferencial parcial (EDP):

$$\begin{aligned} \frac{1}{2} \sigma_1^2 S^2 F_{SS} + \sigma_1 \sigma_2 S F_{S\delta} + \frac{1}{2} \sigma_2^2 F_{\delta\delta} \rho + (r - \delta) S F_S \\ + (\kappa(\alpha - \delta) - \lambda \sigma_2) F_\delta - F_\tau = 0 \end{aligned} \quad (4)$$

Sujeito à condição limite do terminal $F(S, \delta, 0) = S$.

Onde λ denota o preço de mercado por unidade de risco do rendimento de conveniência e τ é o tempo restante para o vencimento do contrato ($T - t$). A solução analítica não era conhecida à época e os autores propõem uma solução numérica definindo proxies para as variáveis de estado, não diretamente observáveis.

Abstraindo a incerteza da taxa de juros e na ausência de arbitragem, podemos definir a Medida Martingal Equivalente- MME, $\mathbb{Q}$. No Apêndice A propomos a solução do sistema de equações diferenciais estocásticas que definem a dinâmica das variáveis S_t e δ_t sob a medida de probabilidade $\mathbb{Q}$. Sabendo que $F_{0,T} = E^{\mathbb{Q}}(S_T | \mathcal{F}_0)$ e usando as equações (2) e (3) acima, demonstra-se que:

$$\ln F_{0,T} = \ln S_0 - \delta_0 \frac{1 - e^{-\kappa T}}{\kappa} + A(T) \quad (5)$$

Onde:

$$A(T) = f_T + \left(r - \hat{\alpha} + \frac{\sigma_2^2}{2\kappa} - \frac{\sigma_1 \sigma_2 \rho}{\kappa} \right) T + \left(\kappa \hat{\alpha} + \sigma_1 \sigma_2 \rho - \frac{\sigma_2^2}{k} \right) \frac{1 - e^{-\kappa T}}{\kappa^2} + \frac{\sigma_2^2}{4\kappa^3} (1 - e^{-2\kappa T}) - f_0$$

e ainda $\hat{\alpha} = \alpha - \frac{\lambda}{\kappa}$.

Nesse modelo, a *commodity* é tratada como um ativo que paga um dividendo estocástico δ_t . Assim, o desvio ajustado ao risco será $r - \delta_t$. Como o risco de rendimento de conveniência não pode ser protegido, o processo de rendimento de conveniência ajustado ao risco terá um preço de mercado de risco associado a ele, λ .

3 MODELO DE TRÊS FATORES

A partir da suposição de que o retorno de conveniência reverte para uma média estocástica, incluímos um terceiro fator no modelo de Gibson e Schwartz (1990), nesse modelo o parâmetro α segue um processo estocástico de reversão a média do tipo *Ornstein-Uhlenbeck*. Os processos estão definidos abaixo, sob a medida real de probabilidade $\mathbb{P}$.

$$dS_t = (\mu - \delta_t) S_t dt + \sigma_1 S_t dB_{1t} \quad (6)$$

$$d\delta_t = \kappa_1(\alpha_t - \delta_t)dt + \sigma_2 dB_{2t} \quad (7)$$

$$d\alpha_t = \kappa_2(\bar{\alpha} - \alpha_t) + \sigma_3 dB_{3t} \quad (8)$$

Na equação (6), assim como na equação (2), μ representa o retorno total de longo prazo esperado pelo proprietário da *commodity*, $\sigma_1 > 0$ é a volatilidade de X_t , $dB_{1,t}$ é o incremento do processo *Browniano* padrão para X_t e δ_t é o retorno de conveniência instantâneo.

Na equação (7) temos que $\kappa_1 > 0$ é a velocidade de reversão a média e α_t , que neste modelo será tratada como uma variável estocástica, é a média de longo prazo do retorno de conveniência. $\sigma_2 > 0$ é a volatilidade e $dB_{2,t}$ é o incremento do processo *Browniano* padrão para δ_t .

Na equação (8) κ_2 é a velocidade de reversão à média e $\bar{\alpha}$ é a média de longo prazo de α_t , $\sigma_3 > 0$ é a volatilidade e $dB_{3,t}$ é o incremento do processo *Browniano* padrão para α_t .

Além disso, os incrementos dos processos *Brownianos* padrão $dB_{1,t}$, $dB_{2,t}$ e $dB_{3,t}$ são correlacionados por ρ_i , tal que: $dB_{2,t}dB_{3,t} = \rho_1 dt$, $dB_{1,t}dB_{2,t} = \rho_2 dt$ e $dB_{1,t}dB_{3,t} = \rho_3 dt$

No Apêndice B propomos uma solução do sistema de equações diferenciais estocásticas que definem a dinâmica das variáveis S_t e δ_t e α_t sob a Medida Martingal Equivalente- MME. Sabendo que $F_{0,T} = E^{\mathbb{Q}}(S_T | \mathcal{F}_0)$, demonstra-se que:

$$\begin{aligned} \ln F_{0,T} = \ln S_0 - \frac{\delta_0}{k_1} (1 - e^{-k_1 T}) \\ - \frac{k_1}{k_1 - k_2} \alpha_0 \left[\frac{1}{k_2} (1 - e^{-k_2 T}) - \frac{1}{k_1} (1 - e^{-k_1 T}) \right] + C(T) \end{aligned} \quad (9)$$

Onde:

$$\begin{aligned}
C(T) = & \left(r - \hat{\alpha} + \frac{\lambda_2}{k_1} \right) T + \frac{1}{k_1} \left(\hat{\alpha} - \frac{\lambda_2}{k_1} \right) (1 - e^{-k_1 T}) \\
& + \frac{k_1}{k_1 - k_2} \hat{\alpha} \left[\frac{1}{k_2} (1 - e^{-k_1 T}) - \frac{1}{k_1} (1 - e^{-k_1 T}) \right] \\
& + \frac{1/2 \sigma_3^2 k_1^2}{(k_1 + k_2)^2} \left[\frac{T}{(k_1 + k_2)^2} + \frac{T^2}{k_1 + k_2} + \frac{T^3}{3} + \frac{4}{(k_1 + k_2)^3} (1 - e^{-(k_1 + k_2)T}) \right. \\
& + \left. \frac{1}{2(k_1 + k_2)^2} (1 - e^{-2(k_1 + k_2)T}) - \frac{2}{(k_1 + k_2)^2} t e^{-(k_1 + k_2)T} \right] \\
& + \frac{\frac{1}{2} \sigma_2^2}{k_1^2} \left[T + \frac{1}{2k_1} (1 - e^{-k_1 T}) - \frac{2}{k_1} (1 - e^{-k_1 T}) \right] \\
& + \frac{\rho_2 \sigma_2 \sigma_3}{k_1 + k_2} \left[\frac{T^2}{2} + \frac{T}{k_1} e^{-k_1 T} - \frac{1}{k_1^2} + \frac{1}{k_1^2} e^{-k_1 T} + \frac{T}{k_1 + k_2} \right. \\
& + \left. \frac{1}{(k_1 + k_2)^2} (1 - e^{-(k_1 + k_2)T}) - \frac{1}{k_1(k_1 + k_2)} (1 - e^{-k_1 T}) \right. \\
& - \left. \frac{1}{(k_1 + k_2)(2k_1 + k_2)} (1 - e^{-(2k_1 + k_2)T}) \right] \\
& - \frac{\rho_3 \sigma_1 \sigma_3}{k_1 + k_2} \left[\frac{T^2}{2} + \frac{T}{k_1 + k_2} - \frac{1}{(k_1 + k_2)^2} (1 - e^{-(k_1 + k_2)T}) \right] \\
& - \frac{\rho_1 \sigma_1 \sigma_2}{k_1} \left[T - \frac{1}{k_1} (1 - e^{-k_1 T}) \right]
\end{aligned}$$

Uma das dificuldades na implementação empírica de modelos de preços de commodities é que frequentemente os fatores desses modelos não são diretamente observáveis. Usualmente, o contrato futuro com vencimento mais próximo é usado como proxy para o preço à vista. Já para estimar o retorno de conveniência, geralmente, são usados dois preços futuros com vencimentos diferentes. A forma de espaço de estado é o procedimento apropriado para lidar com situações nas quais as variáveis não são diretamente observáveis, mas são conhecidas por serem geradas por um processo de Markov.

A forma de espaço de estado se aplica a uma série temporal multivariada de variáveis observáveis. Neste caso, preços de contratos futuros com diferentes maturidades, relacionadas às variáveis de estado (não observadas) através de uma equação de medição (ou observação) que pode ser escrita por:

$$F_{t,T} = \hat{F}_{t,T} + \epsilon_t \quad (10)$$

Onde $F_{t,T}$ é um vetor coluna com o número de linhas igual ao número de contratos observados contendo os preços em cada instante de tempo t . $\hat{F}_{t,T}$ representa o exponencial da equação (5) para o modelo de dois fatores e equação (9) para o modelo de três fatores. ϵ_t é um vetor de distúrbios não serialmente correlacionados com média zero, tal que $\epsilon_t \sim N(0, s_{\epsilon i}^2)$, com i representando cada contrato futuro que compõe o vetor.

As equações de transição são uma versão em tempo discreto do processo estocástico das variáveis de estado: no modelo de dois fatores as variáveis de estado são S_t e δ_t , cujas dinâmicas estão descritas nas equações (1) e (3). Já para o modelo de três fatores as equações de transição derivam dos processos descritos em (6), (7) e (8). Nos apêndices A e B descrevemos os modelos de dois e três fatores na forma de espaço de estado.

Uma vez que o modelo tenha sido lançado na forma de espaço de estado o filtro de *Kalman* pode ser aplicado para estimar os parâmetros e a série temporal das variáveis de estado não observáveis.

O filtro de *Kalman* é um procedimento recursivo para calcular o estimador ótimo do vetor de estado no tempo t , com base nas informações disponíveis no tempo t , e permite que a estimativa do vetor de estado seja continuamente atualizada à medida que novas informações se tornam disponíveis. Quando as perturbações e o vetor de estado inicial são normalmente distribuídos, o filtro de *Kalman* permite que a função de verossimilhança seja calculada e os parâmetros são obtidos maximizando a função de verossimilhança, o que fornece a base para testes estatísticos.

4 EXPERIMENTOS NUMÉRICOS

Neste capítulo apresentamos os experimentos numéricos realizados. Para tanto os modelos foram estimados com dados semanais de preços futuros do petróleo. Os modelos foram comparados com relação à aderência entre os preços futuros observados e os preços futuros filtrados, isto é, obtidos a partir da aplicação do filtro de *Kalman*.

4.1 Dados

Os dados utilizados correspondem a observações semanais dos preços futuros para o petróleo. Considerou-se a informação do preço de fechamento de quarta-feira. Na ocorrência de feriado neste dia, utilizou-se o preço do dia útil mais próximo. Todos os dados foram obtidos através da plataforma *Bloomberg*.

A amostra completa do petróleo WTI abrange o período de 3/1/1996 a 6/6/2018. Foram utilizados cinco contratos de futuros (N=5) na estimação, F1, F5, F10, F15 e F20, onde F1 é o contrato com vencimento mais próximo, ou seja, de maior liquidez, F5 é o segundo contrato mais próximo do vencimento, e assim por diante. O vencimento dos contratos futuros no NYMEX ocorre no terceiro dia útil anterior ao dia 25 de cada mês.

A Tabela 1 apresenta algumas propriedades estatísticas das séries de preços.

Tabela 1—Estatísticas das séries de preços de contratos futuros do petróleo: 3/1/96 a 6/6/2018

	F1	F5	F10	F15	F20
Média	54.30	54.75	54.55	54.21	53.90
Máximo	143.57	144.97	144.96	143.87	142.86
Mínimo	11.16	12.06	12.72	13.27	13.76
DP	29.70	30.07	30.20	30.16	30.04
Assimetria	0.46	0.38	0.33	0.29	0.27
Curtose	-0.88	-0.98	-1.05	-1.10	-1.12
Obs.	1171	1171	1171	1171	1171

Fonte: A autora, 2019.

O Gráfico 1 mostra a série completa do primeiro contrato futuro F1 do painel de dados.

Gráfico 1 –Preços (US\$/Barril) do contrato F1 do petróleo WTI: 3/1/96 a 6/6/18.



Fonte: A autora, 2019.

4.2 Resultados empíricos

Nesta seção apresentamos os resultados da calibração dos modelos de dois e três fatores estocásticos. O modelo de dois fatores é definido pelas equações (1), (3) e (5) reescritas abaixo:

$$dS_t = (\mu - \delta_t)S_t dt + \sigma_1 S_t dB_{1,t}$$

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_2 dB_{2,t}$$

$$\ln F_{0,T} = \ln S_0 - \delta_0 \frac{1 - e^{-\kappa T}}{\kappa} + A(T)$$

Na Tabela 2 apresentamos as estimativas de parâmetros, assim como os erros padrões para o modelo de dois fatores. A partir dessa tabela, pode-se observar que a maioria dos parâmetros é estatisticamente significativa, o que mostra a confiabilidade do modelo aplicado ao mercado de petróleo. O único parâmetro não significativamente diferente de zero é o de tendência μ . Este fato está de acordo com a observação feita em Schwartz e Smith (2000), onde os autores observam a dificuldade de se obter um resultado coerente para este

hiperparâmetro. Além disso a amplitude do ruído na série de preços do quinto e décimo quinto contrato futuro não é estatisticamente significativa (parâmetros $\sigma_{\epsilon_5}^2$ e $\sigma_{\epsilon_{15}}^2$).

Pela estatística t calculada, rejeitamos a hipótese de que o parâmetro α , ou seja, a média de longo prazo do retorno de conveniência, que é tratada como variável estocástica no modelo proposto, seja igual a zero.

Tabela 2—Estimação dos hiperparâmetros do modelo de 2 fatores

Parâmetro	Modelo 2 fatores		
	Valor	Erro padrão	Estatística t
μ	0,0759588	0,0915885	0,8293482
κ	0,6467875	0,0090523	71,450440
α	0,4078165	0,0652142	6,2534910
σ_1	0,4039931	0,0089542	45,117620
σ_2	0,1703894	0,0046890	36,338160
ρ	0,4318257	0,0522248	8,2686000
λ_{F1}	0,2469379	0,0402758	6,1311680
$\sigma_{\epsilon_1}^2$	0,0013048	0,0000008	1692,7260
$\sigma_{\epsilon_5}^2$	0,0000003	0,0000070	0,0376360
$\sigma_{\epsilon_{10}}^2$	0,0000225	0,0000000	2003,0150
$\sigma_{\epsilon_{15}}^2$	0,0000000	0,0000000	0,0026299
$\sigma_{\epsilon_{20}}^2$	0,0000612	0,0000000	1719,8820
AIC	-42540,36		

Fonte: A autora, 2019.

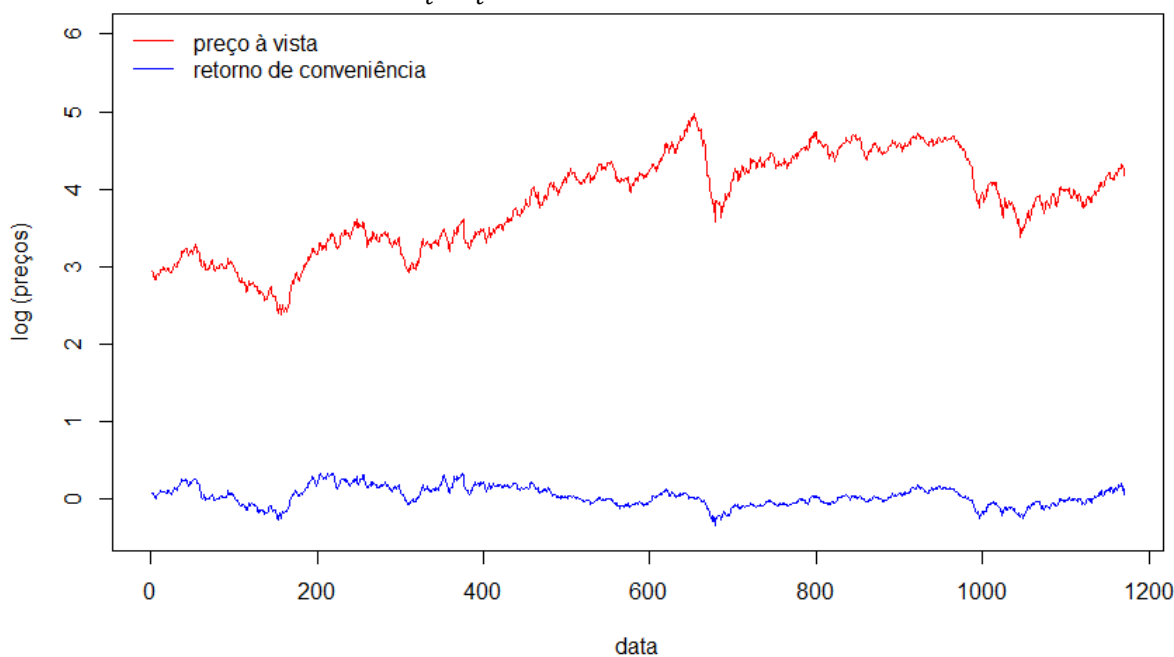
O Gráfico 2 mostra os preços filtrados e observados para o primeiro contrato futuro, F1. A variável filtrada segue apropriadamente a variável observada.

Gráfico 2 – Preços F1 observados para o modelo de dois fatores



Fonte: Autor, 2019.

O Gráfico 3 mostra as duas variáveis de estado S_t e δ_t estimadas. Pode-se observar que o preço à vista, S_t , é mais volátil do que o retorno de conveniência δ_t , o que está de acordo com os hiperparâmetros estimados σ_1 e σ_2 .

Gráfico 3—Variáveis de estado X_t e δ_t 

Fonte: A autora, 2019.

O modelo de três fatores é definido pelas equações (6), (7), (8) e (9) reescritas abaixo:

$$dS_t = (\mu - \delta_t) S_t dt + \sigma_1 S_t dB_{1t}$$

$$d\delta_t = \kappa_1(\alpha_t - \delta_t)dt + \sigma_2 B_{2t}$$

$$d\alpha_t = \kappa_2(\bar{\alpha} - \alpha_t) + \sigma_3 dB_{3t}$$

$$\ln F_{0,T} = \ln S_0 - \frac{\delta_0}{k_1}(1 - e^{-k_1 T}) - \frac{k_1}{k_1 - k_2} \alpha_0 \left[\frac{1}{k_2}(1 - e^{-k_2 T}) - \frac{1}{k_1}(1 - e^{-k_1 T}) \right] + C(T)$$

A Tabela 3 apresenta os resultados da calibragem do modelo de três fatores. Assim como no modelo de dois fatores a maioria dos hiperparâmetros é estatisticamente significativa. Nesse modelo o parâmetro λ_i (preço do risco no mercado do retorno de conveniência e da média do retorno de conveniência) é não estatisticamente significativa. Além disso, o parâmetro μ é estatisticamente significativa para um nível de 10%.

Tabela 3- Estimação dos hiperparâmetros do modelo de 3 fatores

Parâmetro	Modelo 3 fatores		
	Valor	Erro padrão	Estatística t
μ	0,002088	0,001066	1,958162
κ_2	2,927107	0,060447	48,42410
κ_3	0,5867198	0,012117	48,42179
$\bar{\alpha}$	-0,000726	0,000225	-3,230833
σ_1	0,357275	0,010471	34,12149
σ_2	0,605436	0,028511	21,23498
σ_3	0,106695	0,003707	28,78076
ρ_1	0,600704	0,032335	18,57770
ρ_2	-0,112278	0,007814	-14,36812
ρ_3	-0,437285	0,029614	-14,76635
λ_1	0,041743	0,125135	0,333583
λ_2	-0,031347	0,025737	-1,217972
$\sigma_{\epsilon_1}^2$	0,000311	0,000003	105,5703
$\sigma_{\epsilon_5}^2$	0,000051	0,000000	333,6229
$\sigma_{\epsilon_{10}}^2$	0,000000	0,000000	2,142704
$\sigma_{\epsilon_{15}}^2$	0,000004	0,000000	-1149,236
$\sigma_{\epsilon_{20}}^2$	0,000015	0,000000	-556,9992
AIC	-45991,29		

Fonte: A autora, 2019.

O Gráfico 4 mostra os preços filtrados e observados para o primeiro contrato futuro, F1, para o modelo de três fatores. Note que, assim como no modelo de dois fatores, a variável filtrada segue apropriadamente a variável observada.

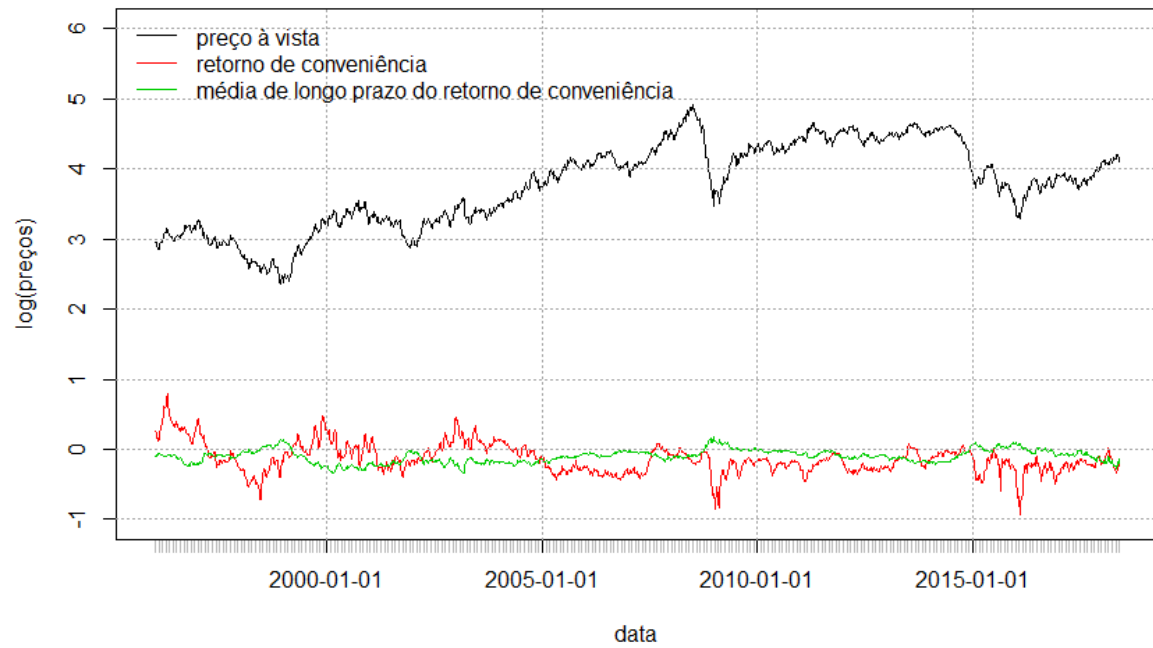
Gráfico 4—Preços F1 observados e filtrados para o modelo de três fatores



Fonte: A autora, 2019.

O Gráfico 5 mostra as três variáveis de estado S_t , δ_t e α_t estimadas. A variável com maior volatilidade é o preço à vista, X_t , o que está de acordo com os hiperparâmetros estimados.

Gráfico 5 – Variáveis de estado X_t , δ_t e α_t



Fonte: A autora, 2019.

4.3 Comparando os modelos

A fim de verificar a aderência dos modelos estimados aos dados empíricos, calculamos os erros entre preço futuro observado ($F_{t,Tj}$) e preço futuro filtrado ($\hat{F}_{t,Tj}$), isto é, obtido a partir da filtragem juntamente com os parâmetros ótimos do modelo. As medidas de erro são dadas pelas fórmulas:

- (i) raiz do erro quadrático médio (*root-mean-square error*):

$$RMSE = \left[\frac{1}{T} \sum_{t=1}^T (F_{t,Tj} - \hat{F}_{t,Tj})^2 \right]^{0,5}, \quad j = 1, \dots, N \quad (11)$$

- (ii) erro absoluto médio (*mean absolute error*):

$$MAE = \frac{1}{T} \sum_{t=1}^T |F_{t,Tj} - \hat{F}_{t,Tj}|, \quad j = 1, \dots, N \quad (12)$$

- (iii) erro absoluto percentual médio (*mean absolute percentage error*):

$$MAPE = \frac{1}{T} \sum_{t=1}^T \frac{|F_{t,Tj} - \hat{F}_{t,Tj}|}{\hat{F}_{t,Tj}}, \quad j = 1, \dots, N \quad (13)$$

A comparação entre os modelos de 2 e de 3 fatores através dos erros expressos estão apresentados na Tabela 4. O erro percentual médio das observações filtradas para o modelo de 2 fatores é de 0,69%, enquanto para o modelo de 3 fatores é de 0,55%. O maior erro observado é de 2,49% e 1,14% referentes ao primeiro contrato, F1, para o modelo de 2 e de 3 fatores, respectivamente. O contrato F15 apresentou menor erro no modelo de 2 fatores. Para o contrato F10 o modelo de 3 fatores, apresentou menores erros.

Tabela 4—Comparação entre os modelos

Séries de preços futuros	Modelo de 2 fatores			Modelo de 3 fatores		
	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>
F1	1,603011	1,153197	0,024801	0,844801	0,632750	0,011411
F5	0,001881	0,001359	0,000028	0,589876	0,447970	0,007860
F10	0,251669	0,179666	0,003639	0,002702	0,001775	0,000035
F15	0,000000	0,000000	0,000000	0,178418	0,141273	0,002419
F20	0,382216	0,279513	0,006022	0,448316	0,356560	0,006104
Média	0,4477554	0,322746	0,006898	0,412822	0,3160655	0,005565

Fonte: A autora, 2019.

Comparamos também a qualidade do ajuste dos modelos partir do Critério de Informação de Akaike que descreve principalmente o *trade-off* entre precisão e complexidade do modelo e é expresso por:

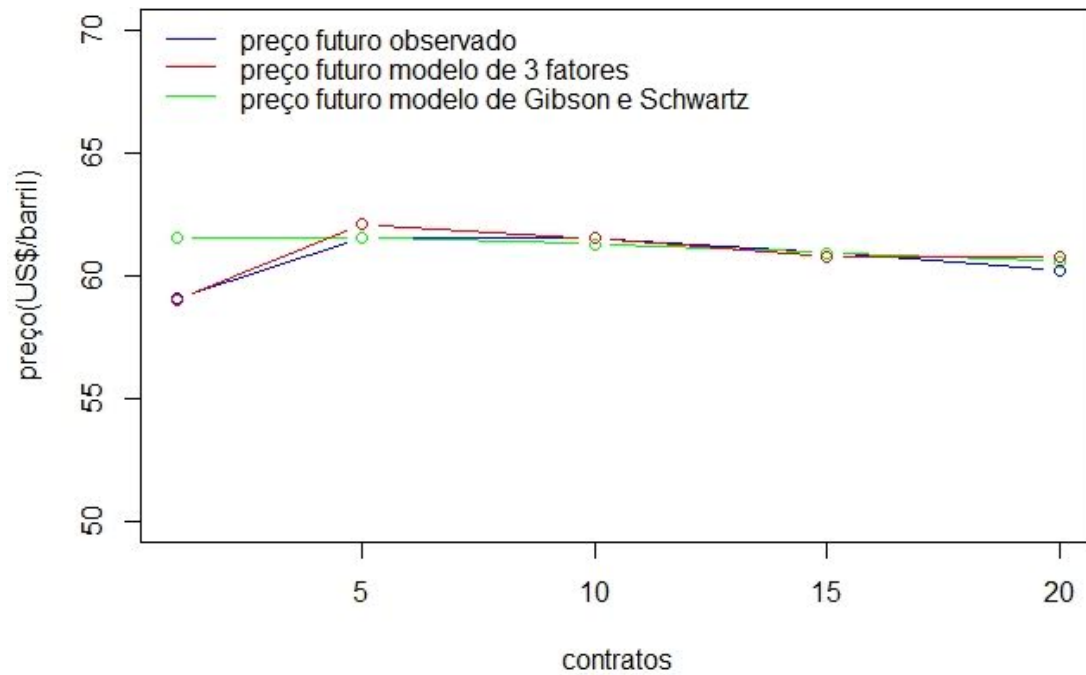
$$AIC = -2 \ln f(x|\hat{\theta}) + 2k \quad (14)$$

onde $\ln f(x|\hat{\theta})$ o logaritmo da máxima verossimilhança e k o número de parâmetros da regressão.

O modelo de três fatores apresentou AIC menor (-45991,29) do que o de dois fatores (-42540,36), indicando que tal modelo é estatisticamente mais apropriado. Além disso, se compararmos os modelos por um critério de aderência, haja vista os valores RMSE, MAE e MAPE calculados, o modelo de três fatores apresentou melhores resultados.

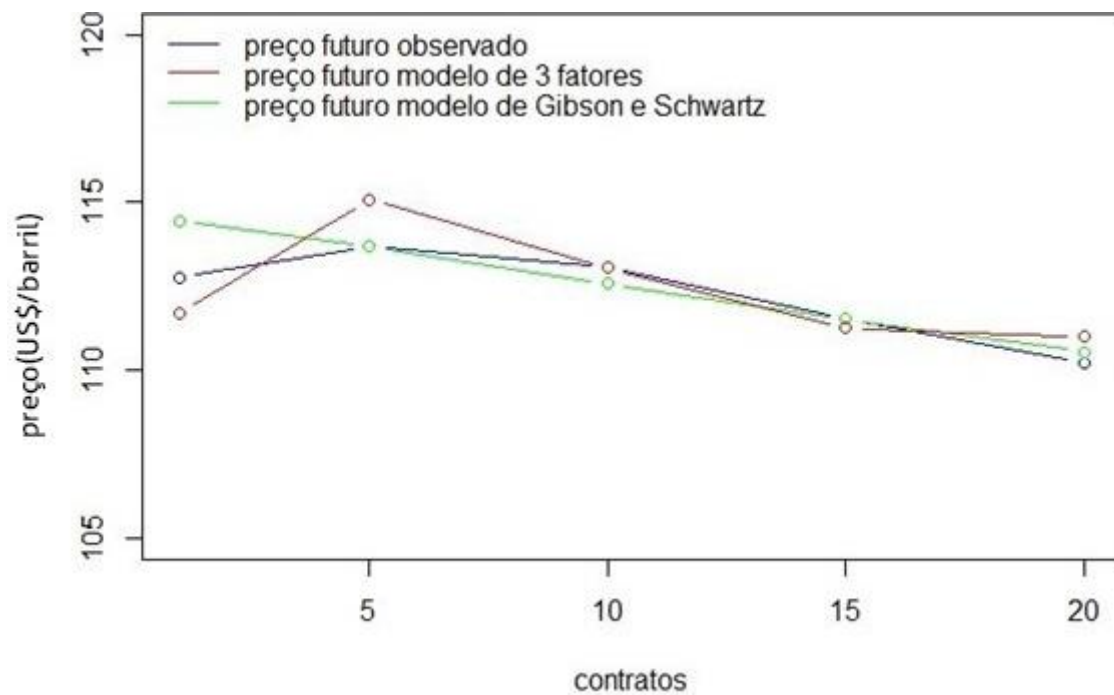
Após a calibragem dos modelos e apresentação dos erros observados na estrutura a termo, consideramos três datas escolhidas aleatoriamente dentro da amostra para verificar se os preços futuros estimados, referente aos contratos F1, F5, F10, F15 e F20, estão aderentes à estrutura a termo na respectiva data. O Gráfico 6 mostra a aderência dos modelos de dois e três fatores à estrutura a termo dos preços futuros estimados para os contratos F1, F5, F10, F15 e F20 em 27/07/2005, o Gráfico 7 mostra a mesma estrutura em 27/04/2011 e no Gráfico 8 em 01/10/2018. Pode-se observar que a estrutura a termo do modelo de 3 fatores está mais aderente do que o modelo de 2 fatores, principalmente para os contratos com um prazo de vencimento mais longo. Note que o primeiro contrato futuro apresentou um erro mais elevado para o modelo de 2 fatores. A partir do quinto contrato a aderência do modelo de 2 fatores à estrutura a termo dos preços melhora. De forma geral, o modelo de 3 fatores apresentou uma melhor aderência à estrutura a termo dos preços no período analisado.

Gráfico 6 – Estrutura a termo nos modelos de 2 e 3 fatores: 27/07/2005



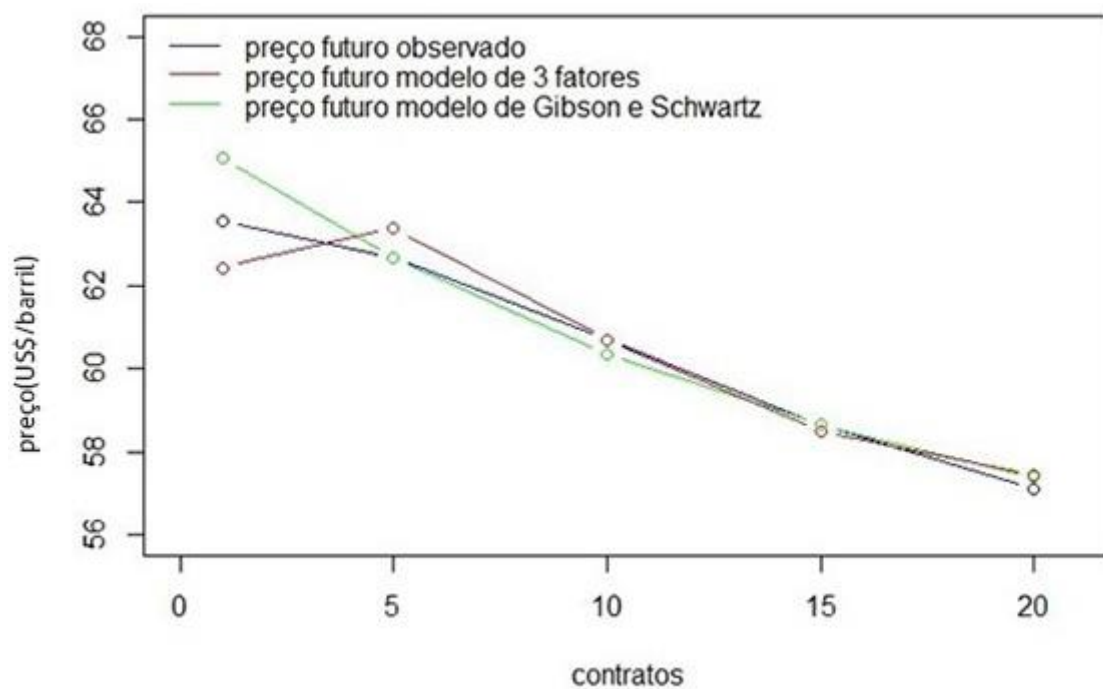
Fonte: A autora, 2019.

Gráfico 7 – Estrutura a termo nos modelos de 2 e 3 fatores: 27/04/2011



Fonte: O autor, 2019.

Gráfico 8 –Estrutura a termo nos modelos de 2 e 3 fatores: 01/10/2018



Fonte: A autora, 2019.

5 APLICAÇÃO EM *HEDGE* DE COMPROMISSO *FORWARD*

Assim como em Schwartz (1997), investigamos a viabilidade de *hedge* para compromissos *forward* de longo prazo utilizando-se de contratos futuros de curto prazo. O objetivo é construir estratégias de *hedge* dinâmico ideais para proteger adequadamente um compromisso a termo em uma determinada mercadoria. A sensibilidade do valor presente do compromisso *forward* com relação aos fatores subjacentes do modelo deve ser igual à sensibilidade da carteira de contratos futuros usada para proteger o compromisso com relação aos mesmos fatores. Isso implica, portanto, que o número de contratos futuros necessários para o *hedge* é igual ao número de fatores utilizados no modelo. Seja ω_i a posição comprada no contrato com maturidade τ_i necessária para o *hedge* de um compromisso de entrega da *commodity* em τ . No modelo de dois fatores requer-se duas posições no mercado futuro para entrega no tempo t_1 e t_2 .

$$\begin{aligned}\omega_1 F_S(S, \delta, \tau_1) + \omega_2 F_S(S, \delta, \tau_2) &= e^{-r\tau} F_S(S, \delta, \tau) \\ \omega_1 F_\delta(S, \delta, \tau_1) + \omega_2 F_\delta(S, \delta, \tau_2) &= e^{-r\tau} F_\delta(S, \delta, \tau)\end{aligned}\tag{15}$$

A relação só é possível porque a taxa de juros foi assumida constante para ambos os modelos, o que implica que os preços futuros são iguais aos preços *forward*, logo o valor do compromisso *forward* pode ser obtido trazendo as equações (5) e (9) ao valor presente.

O modelo de três fatores requer três posições no mercado futuro para entrega no tempo t_1 , t_2 e t_3 , representando o preço à vista, o retorno de conveniência e a média de longo prazo do retorno de conveniência.

$$\begin{aligned}\omega_1 F_S(S, \delta, \alpha, \tau_1) + \omega_2 F_S(S, \delta, \alpha, \tau_2) + \omega_3 F_S(S, \delta, \alpha, \tau_3) &= e^{-r\tau} F_S(S, \delta, \alpha, \tau) \\ \omega_1 F_\delta(S, \delta, \alpha, \tau_1) + \omega_2 F_\delta(S, \delta, \alpha, \tau_2) + \omega_3 F_\delta(S, \delta, \alpha, \tau_3) &= e^{-r\tau} F_\delta(S, \delta, \alpha, \tau) \\ \omega_1 F_\alpha(S, \delta, \alpha, \tau_1) + \omega_2 F_\alpha(S, \delta, \alpha, \tau_2) + \omega_3 F_\alpha(S, \delta, \alpha, \tau_3) &= e^{-r\tau} F_\alpha(S, \delta, \alpha, \tau)\end{aligned}\tag{16}$$

A cada instante de tempo são computados os valores para F_S , F_δ e F_α e os sistemas em (15) e (16) são resolvidos obtendo-se os valores ω_1 , ω_2 e ω_3 para este instante.

CONCLUSÃO

Esta dissertação propôs uma abordagem de variável de estado para o comportamento estocástico dos preços das commodities. Comparamos modelos de dois e três fatores quanto a sua capacidade de precificar contratos futuros existentes. O modelo de dois fatores foi desenvolvido em Gibson e Schwartz (1990) e a partir da extensão deste, desenvolvemos um modelo de três fatores, onde a média de longo prazo do retorno de conveniência segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*.

Para os dois modelos obtemos soluções de forma fechada para os preços de futuros, de tal forma que o logaritmo do preço futuro é uma função linear dos fatores subjacentes. Os modelos foram lançados na forma de espaço de estado e o filtro de *Kalman* foi aplicado para estimar os parâmetros do modelo e a série temporal das variáveis de estado não diretamente observáveis. Os dados utilizados na estimação são séries de preços semanais de contratos futuros do petróleo WTI entre 1996 e 2018.

Em ambos os modelos a maioria dos parâmetros estimados foram estatisticamente significantes, no entanto o modelo proposto (com três fatores) apresentou menores erros - expressos pelos RMSE, MAE e MAPE -, uma melhor aderência aos dados observados, e um AIC inferior ao modelo de dois fatores.

Futuras pesquisas podem aplicar o modelo a outros ativos. Para aqueles que apresentam um padrão sazonal nos preços, como as commodities agrícolas, por exemplo, recomenda-se a inclusão de parâmetros de sazonalidade. Além disso, o modelo pode ser estendido com a inclusão de novos fatores.

REFERÊNCIAS

- AIUBE, FERNANDO A. L., BAIDYA, TARA K. N. E TITO, EDISON A. H. Analysis of Commodity Prices with the Particle Filter. **Energy Economics**, [s.l.], v. 30, p. 597–605, 2008.
- BABBS, S. H. E NOWMAN, K. B. Kalman Filtering of Generalized Vasicek Term Structure Models. **Journal of Financial and Quantitative Analysis**, [s.l.], v. 34, n. 1, p. 115-130, 1999.
- BESSEMBINDER, H.; COUGHENOUR, J. F.; SMOLLER, M.; SEGUIN, P. J. Mean Reversion in Equilibrium Asset Prices: Evidence from the Futures Term Structure. **Journal of Finance**, , [s.l.], v. 50, n.1, p. 361-375, 1995.
- BRENNAN, M.J. The Supply of Storage. **The American Economic Review**, v. 48, p. 50-72, 1958.
- BRENNAN, M. J.; SCHWARTZ, E.S. Evaluating natural resource investments. **Journal of business**, JSTOR, p. 135-157, 1985.
- CORTAZAR, G.; NARANJO, L. An n-factor Gaussian model of oil futures prices. **Journal of Futures Markets**, Wiley Online Library, v. 26, n. 3, p. 243-268, 2006.
- DIAS, M. A. G.; ROCHA, K. Petroleum Concessions With Extendible Options using Mean Reversion with Jumps to model Oil Prices. **3rd Real Option Conference, Annals, Wassenaar/Leiden**, 1999.
- DURBIN, J. E KOOPMAN, S. J. **Time Series Analysis by State Space Methods**. Oxford: Oxford University Press, 2002.
- GEMAN, H.; NGUYEN, V. Soybean Inventory and Forward Curve Dynamics. **Management Science**, v. 51, n. 7, p. 1076–1091, 2005.
- GIBSON, R.; SCHWARTZ, E. S. Stochastic convenience yield and the pricing of oil contingent claims. **The Journal of Finance**, Wiley Online Library, v. 45, n. 3, p. 959-976, 1990.
- GIBSON, R.; SCHWARTZ, E.S. Valuation of long term oil-linked assets. **Anderson Graduate School of Management**, UCLA, p. 6-89, 1989.
- HAMILTON, J. D. **Time Series Analysis**. Princeton: Princeton University Press, 1994.
- HARVEY, A. **Forecasting, Structural Time Series Models and the Kalman Filter**. Cambridge: Cambridge University Press, 1989.
- HIKspoors, S. ; JAIMUNGAL, S. Energy spot price models and spread options pricing. **International Journal of Theoretical and Applied Finance**, v. 10, n. 7, p. 1111–1135, 2007.

JAVAHERI, A.; LAUTIER, D.; GALLI, A. Filtering in finance. **Willmott Magazine**, [s.l.], v. 5, 2003.

KALDOR N. Speculation and Economic Stability. **The Review of Economic Studies**, [s.l.], v. 7, n. 1, p. 1–27, 1939

LUCIA, J. J.; SCHWARTZ, E. S. Electricity prices and power derivatives: Evidence from the Nordic power exchange. **Review of derivatives research**, Springer, v. 5, n. 1, p.5-50, 2002.

MANOLIU, M.; TOMPAIDIS, S. Energy futures prices: term structure models with Kalman Filter estimation. **Applied mathematical finance**, Taylor & Francis, v. 9, n. 1, p.21-43, 2002.

PADDOCK, J. L.; SIEGEL, D. R.; SMITH J. L. Option Valuation of Claims on Real Assets: The Case of Offshore Petroleum Leases. **Quarterly Journal of Economics**, Oxford, v. 103, n. 3, p. 479-508, 1988.

PASRICHA, G. **Survey of Literature on Covered and Uncovered Interest Parities**. Germany: MPRA Paper, University Library of Munich, 2006.

PILIPOVIĆ, D. **Energy Risk: Valuing And Managing Energy Derivatives**. New York: McGraw-Hill, 1997.

PINDYCK, R. S. The Dynamics of Commodity Spot and Futures Markets: A Primer. **The Energy Journal**, [s.l.], v. 22, n. 3, p. 1-29, 2001.

PINDYCK, R. S. The Long-Run Evolutions of Energy Prices. **The Energy Journal**, [s.l.], v. 20, n. 2, p. 1-27, 1999.

POSTALI, F. A.; PICCHETTI, P. Geometric Brownian Motion and structural breaks in oil prices: A quantitative analysis. **Energy Economics**, [s.l.], v.28, n. 4, p. 506-522, 2006.

SCHWARTZ, E. S. The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging. . **Journal of Finance**, [s.l.], 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**, Informs, v. 46, n. 7, p. 893-911, 2000.

SORENSEN, C. Modeling seasonality in agricultural commodity futures. **Journal of Futures Markets**, Wiley Online Library, v. 22, n. 5, p. 393-426, 2002.

VILLAPLANA, P. Pricing Power Derivatives: A Two-Factor Jump-Diffusion Approach. **EFMA 2004 Basel Meetings Paper**, 2003.

WORKING, H. The Theory of Price of Storage. **American Economic Review**, Pittsburgh, v.39, p. 1254-1262, 1949.

APÊNDICE A - Modelo De Dois Fatores

Neste Apêndice são apresentados os detalhes da solução analítica do conjunto de equações diferenciais estocásticas do modelo de dois fatores para o preço futuro $F_{0,T}$.

Os pressupostos para a derivação abaixo estão no corpo do artigo. Inicialmente apresentamos o modelo sob a medida real de probabilidade.

Seja S_t é o preço à vista, podemos definir X_t tal que:

$$S_t = e^{X_t} \quad (17)$$

onde X_t tem a dinâmica abaixo:

$$dX_t = \left(\mu - \frac{1}{2} \sigma_1^2 \right) dt - \delta_t dt + \sigma_1 dB_{1,t} \quad (18)$$

Na equação (18), a variável δ_t é o retorno de conveniência. Sua dinâmica segue um processo de reversão à média do tipo *Ornstein-Uhlenbeck*, definido abaixo:

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_2 dB_{2,t} \quad (19)$$

As equações (18) e (19) descrevem o modelo de dois fatores sob a medida real de probabilidade $\mathbb{P}$. A solução de (19) pode ser obtida integrando-se entre a data $t = 0$ e t , onde $t \in [0, T]$.

$$\delta_t = \alpha(\delta_0 - \alpha)^{-\kappa t} + \sigma_2 \int_0^t e^{\kappa(u-t)} dB_{2,u} \quad (20)$$

Integrando a equação (18) de 0 a t temos

$$X_t = X_0 + \left(\mu - \frac{1}{2} \sigma_1^2 \right) t - \int_0^t \delta_u du + \sigma_1 \int_0^t dB_{1,u} \quad (21)$$

Integrando a equação (20) e usando o resultado em (21), temos:

$$\begin{aligned}
X_t = X_0 + \left(\mu - \frac{1}{2}\sigma_1^2\right)t - \alpha t + \frac{\alpha - \delta_0}{\kappa}(1 - e^{-\kappa t}) - \frac{\sigma_2}{\kappa} \int_0^t dB_{2,u} \\
+ \frac{\sigma_2}{\kappa} \int_0^t e^{\kappa(u-t)} dB_{2,u} + \sigma_1 \int_0^t dB_{1,u}
\end{aligned} \tag{22}$$

O valor esperado de X_t , sob a medida $\mathbb{P}$ condicionado às informações disponíveis em $t = 0$ é

$$E^{\mathbb{P}}(X_t | \mathcal{F}_0) = X_0 + \left(\mu - \frac{1}{2}\sigma_1^2\right)t - \alpha t + \frac{\alpha - \delta_0}{\kappa}(1 - e^{-\kappa t}) \tag{23}$$

Usando a isometria de Itô podemos calcular a variância de X_t sob a medida $\mathbb{P}$. Como a variância é a mesma em ambas as medidas, apresentaremos este resultado ao tratarmos do modelo sob a medida $\mathbb{Q}$.

Nesse modelo, a *commodity* é tratada como um ativo que paga um dividendo estocástico δ_t . Assim, o desvio ajustado ao risco será $r - \delta_t$. Como o risco de rendimento de conveniência não pode ser protegido, o processo de rendimento de conveniência ajustado ao risco terá um preço de mercado de risco associado a ele. Sob a Medida Martingal Equivalente, o processo estocástico para os fatores pode ser escrito como:

$$dX_t = \left(r - \frac{1}{2}\sigma_1^2\right)dt - \delta_t dt + \sigma_1 d\tilde{B}_{1,t} \tag{24}$$

$$d\delta_t = [\kappa(\alpha - \delta_t) - \lambda]dt + \sigma_2 d\tilde{B}_{2,t} \tag{25}$$

onde λ , considerado constante em $[0, T]$, é o preço de mercado do risco de rendimento de conveniência. A equação (25) pode ser reescrita como:

$$d\delta_t = \kappa(\hat{\alpha} - \delta_t)dt + \sigma_2 dB_{2,t} \tag{26}$$

onde $\hat{\alpha} = \alpha - \frac{\lambda}{\kappa}$, define a média de longo prazo do retorno de conveniência sob a MME.

A solução da equação (26) é dada por:

$$\delta_t = \hat{\alpha} + (\delta_0 - \hat{\alpha})e^{-\kappa t} + \sigma_2 \int_0^t e^{\kappa(u-t)} d\tilde{B}_{2,u} \quad (27)$$

Similarmente à equação (21), escrevemos:

$$X_t = X_0 + \left(r - \frac{1}{2}\sigma_1^2\right)t - \int_0^t \delta_u du + \sigma_1 \int_0^t d\tilde{B}_{1,u} \quad (28)$$

Integrando a equação (27) e introduzindo este resultado na equação (28), chegamos ao resultado para X_t sob a MME, isto é

$$\begin{aligned} X_t = X_0 + \left(r - \frac{1}{2}\sigma_1^2\right)t - \hat{\alpha}t + \frac{\hat{\alpha} - \delta_0}{\kappa}(1 - e^{-\kappa t}) - \frac{\sigma_2}{\kappa} \int_0^t d\tilde{B}_{2,u} \\ + \frac{\sigma_2}{\kappa} \int_0^t e^{\kappa(u-t)} d\tilde{B}_{2,u} + \sigma_1 \int_0^t d\tilde{B}_{1,u} \end{aligned} \quad (29)$$

O valor esperado de X_t sob a MME, condicionado às informações disponíveis na data $t = 0$ é dado por

$$E^{\mathbb{Q}}(X_t|\mathcal{F}_0) = X_0 + \left(r - \frac{1}{2}\sigma_1^2\right)t - \hat{\alpha}t + \frac{\hat{\alpha} - \delta_0}{\kappa}(1 - e^{-\kappa t}) \quad (30)$$

Reescrevemos a equação (31) de forma simplificada onde as integrais estocásticas do segundo membro serão denominadas por I_1, I_2 e I_3 , respectivamente. Usando o resultado da equação (30) chegamos a:

$$X_t = E^{\mathbb{Q}}(X_t|\mathcal{F}_0) - \frac{\sigma_2}{\kappa}I_1 + \frac{\sigma_2}{\kappa}I_2 + \sigma_1 I_3 \quad (31)$$

A variância condicional de X_t sob a MME, envolve o cálculo da variância de cada integral estocástica e as covariâncias entre elas. Portanto, escrevemos

$$\begin{aligned} Var^{\mathbb{Q}}(X_t|\mathcal{F}_0) = \frac{\sigma_2^2}{\kappa^2}Var^{\mathbb{Q}}(I_1|\mathcal{F}_0) + \frac{\sigma_2^2}{\kappa^2}Var^{\mathbb{Q}}(I_2|\mathcal{F}_0) + \sigma_1^2 Var^{\mathbb{Q}}(I_3|\mathcal{F}_0) \\ - 2\frac{\sigma_2^2}{\kappa^2}Cov^{\mathbb{Q}}(I_1, I_2|\mathcal{F}_0) - 2\frac{\sigma_1\sigma_2}{\kappa}Cov^{\mathbb{Q}}(I_1, I_3|\mathcal{F}_0) \end{aligned} \quad (32)$$

$$+2 \frac{\sigma_1 \sigma_2}{\kappa} \text{Cov}^{\mathbb{Q}}(I_2, I_3 | \mathcal{F}_0)$$

Usando a isometria de Itô para o cálculo de variâncias e covariâncias, chegamos ao resultado

$$\begin{aligned} \text{Var}^{\mathbb{Q}}(X_t | \mathcal{F}_0) = & \left(\sigma_1^2 + \frac{\sigma_2^2}{\kappa^2} - \frac{2\sigma_1\sigma_2\rho}{\kappa} \right) t - 2 \frac{\sigma_2^2}{\kappa^3} (1 - e^{-\kappa t}) \\ & + 2 \frac{\sigma_1\sigma_2\rho}{\kappa^2} (1 - e^{-\kappa t}) + \frac{\sigma_2^2}{2\kappa^3} (1 - e^{-2\kappa t}) \end{aligned} \quad (33)$$

Com os resultados acima definiremos a expressão para o preço futuro na data $t = 0$ de um contrato que tem vencimento em T , seja então este preço denotado por $F_{0,T}$. Assumindo uma taxa de juros constante, temos que $F_{0,T} = E^{\mathbb{Q}}(S_T | \mathcal{F}_0)$. Ou também

$$\ln F_{0,T} = \ln[E^{\mathbb{Q}}(S_T | \mathcal{F}_0)] \quad (34)$$

Escrevendo o valor esperado condicional de S_T como

$$E^{\mathbb{Q}}(S_T | \mathcal{F}_0) = E^{\mathbb{Q}}(e^{\ln S_T} | \mathcal{F}_0) \quad (35)$$

Então, a partir das propriedades da distribuição log-normal, temos:

$$E^{\mathbb{Q}}(S_T | \mathcal{F}_0) = \exp \left[E^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) + \frac{1}{2} \text{Var}^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) \right] \quad (36)$$

$$\ln[E^{\mathbb{Q}}(S_T | \mathcal{F}_0)] = E^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) + \frac{1}{2} \text{Var}^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) \quad (37)$$

Levando as equações (30) e (32) na equação (37) e considerando a data terminal $t = T$, temos o resultado:

$$\begin{aligned} \ln[E^{\mathbb{Q}}(S_T | \mathcal{F}_0)] = & X_0 + \left(r - \frac{1}{2} \sigma_1^2 \right) T - \hat{\alpha} T + \frac{\hat{\alpha} - \delta_0}{\kappa} (1 - e^{-\kappa T}) \\ & + \left(\sigma_1^2 + \frac{\sigma_2^2}{\kappa^2} - \frac{2\sigma_1\sigma_2\rho}{\kappa} \right) T - 2 \frac{\sigma_2^2}{\kappa^3} (1 - e^{-\kappa T}) \end{aligned} \quad (38)$$

$$+2 \frac{\sigma_1 \sigma_2 \rho}{\kappa^2} (1 - e^{-\kappa T}) + \frac{\sigma_2^2}{2\kappa^3} (1 - e^{-2\kappa T})$$

Usando a equação (34) e definindo por $A(T)$ a função dos termos que dependem do tempo e demais parâmetros, temos que:

$$\ln F_{0,T} = \ln S_0 - \frac{\delta_0}{\kappa} (1 - e^{-\kappa T}) + A(T) \quad (39)$$

onde

$$A(T) = f_T + \left(r - \hat{\alpha} + \frac{1}{2} \frac{\sigma_2^2}{\kappa^2} - \frac{\sigma_1 \sigma_2 \rho}{\kappa} \right) T + \frac{1 - e^{-\kappa T}}{\kappa^2} \left(\hat{\alpha} \kappa - \frac{\sigma^2}{\kappa} + \sigma_1 \sigma_2 \rho \right) + \frac{\sigma_2^2}{4\kappa^3} (1 - e^{-2\kappa T}) - f_0$$

MODELO NA FORMA ESPAÇO DE ESTADO

A fim de estimar o modelo via filtro de *Kalman*, apresentamos o modelo na forma espaço de estado.

As variáveis de estado, preço à vista e retorno de conveniência, não são diretamente observáveis, apenas o preço futuro é observável. A partir das equações (2), (3) e (5), o modelo foi escrito sob a forma espaço de estado. Esta é a formulação usual para aplicação do filtro de *Kalman* e estimação dos hiperparâmetros, ou seja, dos parâmetros do modelo não variam com o tempo.

Seja N o número de contratos futuros e y_t o logaritmo dos preços futuros observados no mercado. De (5), podemos escrever que a equação de observação é

$$y_t = d_t + Z_t [X_t, \delta_t]^T + \epsilon_t \quad (40)$$

onde

$$y_t = \ln[F_{0,T_i}]$$

$$d_t = A(T_i) \quad (41)$$

$$Z_t = \left[1, -\frac{1 - e^{-\kappa T_i}}{\kappa} \right] \quad (42)$$

Seja ϵ_t é um vetor $N \times 1$, serialmente não correlacionado com média zero e matriz de covariância H_t , ou seja: $E^{\mathbb{P}}(\epsilon_t) = 0$ e $Var^{\mathbb{P}}(\epsilon_t) = H_t$.

As equações de transição do modelo são obtidas das equações (2) e (3) e são escritas por:

$$[X_t, \delta_t]^T = c_t + Q_t[X_{t-1}, \delta_{t-1}]^T + \eta_t \quad (43)$$

onde

$$c_t = \left[\left(\mu - \frac{\sigma_1^2}{2} \right) \Delta t, \kappa \alpha \Delta t \right]^T \quad (44)$$

$$Q_t = \begin{bmatrix} 1 & -\delta t \\ 0 & 1 - \kappa \Delta t \end{bmatrix} \quad (45)$$

onde η_t é serialmente não correlacionado, com médias e variâncias sob a medida $\mathbb{P}$ dadas por:

$$\begin{aligned} E^{\mathbb{P}}(\eta_t) &= 0 \\ Var^{\mathbb{P}}(\eta_t) &= \begin{bmatrix} \sigma_1^2 \Delta t & \rho \sigma_1 \sigma_2 \Delta t \\ \rho \sigma_1 \sigma_2 \Delta t & \sigma_2^2 \Delta t \end{bmatrix} \end{aligned} \quad (46)$$

APÊNDICE B – Modelo de Três Fatores

Neste Apêndice são apresentados os detalhes da solução analítica do conjunto de equações diferenciais estocásticas (6), (7) e (8) que definem os fatores para o preço futuro $F_{0,T}$ no modelo de três fatores. Inicialmente apresentamos o modelo sob a medida real de probabilidade.

$$dS_t = (\mu - \delta_t) S_t dt + \sigma_1 S_t dB_{1,t} \quad (47)$$

$$d\delta_t = \kappa_1(\alpha_t - \delta_t)dt + \sigma_2 B_{2,t} \quad (48)$$

$$d\alpha_t = \kappa_2(\bar{\alpha} - \alpha_t)dt + \sigma_3 dB_{3,t} \quad (49)$$

Seja $X_t = \ln S_t$.

$$(X_t)' = 1/S_t \quad (50)$$

$$(X_t)'' = -1/S_t^2 \quad (51)$$

Usando o Lema de Itô, temos que:

$$dX_t = \frac{\partial X_t}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 X_t}{\partial S_t^2} dS_t^2$$

Logo:

$$dX_t = \left(\mu - \delta_t - \frac{1}{2} \sigma_1^2 \right) dt + \sigma_1 dB_{1,t} \quad (52)$$

Integrando (53) de 0 a t:

$$X_t = X_0 + \left(\mu - \frac{1}{2} \sigma_1^2 \right) t - \int_0^t \delta_z dz + \sigma_1 \int_0^t dB_{1,u} \quad (53)$$

Integrando (49) de 0 a t, temos:

$$\delta_t = \delta_0 e^{-k_1 t} + k_1 \int_0^t e^{k_1(u-t)} \alpha_u du + \sigma_2 \int_0^t e^{k_1 u} dB_{2,u} \quad (54)$$

Integrando (50) de 0 a t, temos:

$$\alpha_t = (\alpha_0 - \bar{\alpha}) e^{-k_2 t} + \bar{\alpha} + \sigma_3 \int_0^t e^{k_2(u-t)} dB_{3,u} \quad (55)$$

Levando (54) em (53), podemos calcular:

$$\begin{aligned} \int_0^t \delta_s ds &= \frac{\delta_0}{k_1} (1 - e^{-k_1 t}) + \bar{\alpha} t + \frac{\bar{\alpha}}{k_1} (e^{-k_1 t} - 1) \\ &\quad + \frac{k_1}{k_1 - k_2} (\alpha_0 - \bar{\alpha}) \left[\frac{-e^{-k_2 t}}{k_2} + \frac{1}{k_2} + \frac{e^{-k_1 t}}{k_1} - \frac{1}{k_1} \right] \\ &\quad + \frac{\sigma_3 k_1}{k_1 + k_2} \int_0^t (t - u) - \frac{1}{k_1 + k_2} (1 - e^{(k_1 + k_2)(u-t)}) dB_{3,u} \\ &\quad + \sigma_2 \int_0^t \frac{1}{k_1} (1 - e^{k_1(u-t)}) dB_{2,u} \end{aligned} \quad (56)$$

Levando (55) em (52) e calculando o valor esperado sob a medida real de probabilidade $\mathbb{P}$.

$$\begin{aligned} E^{\mathbb{P}} &= X_0 + \left(\mu - \frac{1}{2} \sigma_1^2 - \bar{\alpha} \right) t - \frac{1}{k_1} (y_0 - \bar{\alpha}) (1 - e^{-k_1 t}) \\ &\quad - \frac{k_1}{k_2 - k_1} (\alpha_0 - \bar{\alpha}) \left[-\frac{e^{-k_2 t}}{k_2} + \frac{1}{k_2} + \frac{e^{-k_1 t}}{k_2} - \frac{1}{k_1} \right] \end{aligned} \quad (57)$$

Sob a Medida Martingal Equivalente (MME):

$$dX_t = \left(r - \frac{1}{2} \sigma_1^2 \right) dt - \delta_t dt + \sigma_1 d\tilde{B}_{1,t} \quad (58)$$

$$d\delta_t = [\kappa_1 (\alpha_t - \delta_t) - \lambda_1] dt + \sigma_2 d\tilde{B}_{2,t} \quad (59)$$

$$d\alpha_t = [\kappa_2 (\alpha_t^* - \alpha_t)] dt + \sigma_3 d\tilde{B}_{3,t} \quad (60)$$

onde $\alpha^* = \bar{\alpha} - \frac{\lambda_2}{k_2}$

Procedendo da mesma forma que na medida real de probabilidade, temos que:

$$\begin{aligned}
X_t = X_0 &+ \left(r - \frac{1}{2} \sigma_1^2 - \hat{\alpha} + \frac{\lambda_2}{k_1} \right) t - \frac{1}{k_1} \left(\hat{\alpha} - \delta_0 - \frac{\lambda_2}{k_1} \right) (1 - e^{-k_1 t}) \\
&- \frac{k_1}{k_1 - k_2} (\alpha_0 - \hat{\alpha}) \left(\frac{1}{k_2} (1 - e^{-k_2 t}) - \frac{1}{k_1} (1 - e^{-k_1 t}) \right) \\
&- \frac{\sigma_3 \kappa_1}{\kappa_1 + \kappa_2} \int_0^t (t - u) + \frac{1}{k_1 + k_2} (1 - e^{(k_1 + k_2)(u-t)}) d\tilde{B}_{3,u} \\
&- \frac{\sigma_2}{\kappa_1} \int_0^t (1 - e^{k_1(u-t)}) d\tilde{B}_{2,u} + \sigma_1 \int_0^t d\tilde{B}_{1,u}
\end{aligned} \tag{61}$$

$$\begin{aligned}
E^{\mathbb{Q}}(X_t) = X_0 &+ \left(r - \frac{1}{2} \sigma_1^2 - \hat{\alpha} + \frac{\lambda}{k_1} \right) t - \frac{1}{k_1} \left(\hat{\alpha} - \delta_0 - \frac{\lambda_2}{k_1} \right) (1 - e^{-k_1 t}) \\
&- \frac{k_1}{k_1 - k_2} (\alpha_0 - \hat{\alpha}) \left(\frac{1}{k_2} (1 - e^{-k_2 t}) - \frac{1}{k_1} (1 - e^{-k_1 t}) \right)
\end{aligned} \tag{62}$$

Agora podemos escrever que:

$$X_t = E^{\mathbb{Q}}(X_t) - \frac{k_1 \sigma_3}{k_1 + k_2} I_1 - \frac{\sigma_2}{k_1} I_2 + \sigma_1 I_3 \tag{63}$$

onde:

$$I_1 = \int_0^t (t - u) + \frac{1}{k_1 + k_2} (1 - e^{(k_1 + k_2)(u-t)}) d\tilde{B}_{3,u} \tag{64}$$

$$I_2 = \int_0^t (1 - e^{k_1(u-t)}) d\tilde{B}_{2,u} \tag{65}$$

$$I_3 = \int_0^t d\tilde{B}_{1,u} \tag{66}$$

Afim de calcular a variância de X_t , temos que:

$$\begin{aligned}
Var(I_1) = & \frac{t}{(k_1 + K_2)^2} + \frac{t^2}{k_1 + k_2} + \frac{t^3}{3} + \frac{4}{(k_1 + k_2)^3} (1 - e^{-(k_1 + k_2)t}) \\
& + \frac{1}{2(k_1 + k_2)^3} (1 - e^{-2(k_1 + k_2)t}) - \frac{2}{(k_1 + k_2)^2} t e^{-(k_1 + k_2)t}
\end{aligned} \tag{67}$$

$$Var(I_2) = t + \frac{1}{2k_1}(1 - e^{-2k_1t}) - \frac{2}{k_1}(1 - e^{-k_1t}) \quad (68)$$

$$Var(I_3) = t \quad (69)$$

$$\begin{aligned} Cov(I_1, I_2) = \rho_2 \left[\frac{t^2}{2} + \frac{t}{k_1} e^{-k_1t} - \frac{1}{k_1^2} + \frac{1}{k_1^2} e^{-k_1t} + \frac{t}{k_1 + k_2} \right. \\ \left. + \frac{1}{(k_1 + k_2)^2} (1 - e^{-(k_1+k_2)t}) - \frac{1}{k_1(k_1 + k_2)} (1 - e^{-k_1t}) \right. \\ \left. - \frac{1}{(k_1 + k_2)(2k_1 + k_2)} (1 - e^{-(2k_1+k_2)t}) \right] \end{aligned} \quad (70)$$

$$Cov(I_1, I_3) = \rho_3 \left[\frac{t^2}{2} + \frac{t}{k_1 + k_2} + \frac{1}{(k_1 + k_2)^2} (1 - e^{-(k_1+k_2)t}) \right] \quad (71)$$

$$Cov(I_2, I_3) = \rho_1 \left[t - \frac{1}{k_1} (1 - e^{-k_1t}) \right] \quad (72)$$

Agora podemos calcular:

$$\ln F_{0,T} = \ln E^{\mathbb{Q}}(S_T | \mathcal{F}_0) = E^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) + \frac{1}{2} Var^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) \quad (73)$$

onde:

$$\ln E^{\mathbb{Q}}(S_T | \mathcal{F}_0) = E^{\mathbb{Q}}(e^{\ln(S_T)} | \mathcal{F}_0) = \exp [E^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0) + \frac{1}{2} Var^{\mathbb{Q}}(\ln S_T | \mathcal{F}_0)] \quad (74)$$

$$\begin{aligned} Var^{\mathbb{Q}}(X_t | \mathcal{F}_0) = \frac{\sigma_3^2 k_1^2}{(\kappa_1 + \kappa_2)^2} Var^{\mathbb{Q}}(I_1 | \mathcal{F}_0) + \frac{\sigma_2^2}{\kappa_1^2} Var^{\mathbb{Q}}(I_2 | \mathcal{F}_0) + \\ \sigma_1^2 Var^{\mathbb{Q}}(I_3 | \mathcal{F}_0) + 2\rho_2 \frac{\sigma_2 \sigma_3}{\kappa_1 + \kappa_2} Cov^{\mathbb{Q}}(I_1, I_2 | \mathcal{F}_0) - 2\rho_3 \frac{\sigma_1 \sigma_3 k_1}{\kappa_1 + \kappa_2} Cov^{\mathbb{Q}}(I_1, I_3 | \mathcal{F}_0) \\ - 2\rho_1 \frac{\sigma_1 \sigma_2}{\kappa_1} Cov^{\mathbb{Q}}(I_2, I_3 | \mathcal{F}_0) \end{aligned} \quad (75)$$

Usando as equações (75) e (74) na equação (73), temos uma solução para $\ln F_{0,T}$:

$$\begin{aligned} \ln F_{0,T} = X_0 + \left(r - \frac{1}{2} \sigma_1^2 - \hat{\alpha} + \frac{\lambda_2}{k_1} \right) T - \frac{1}{k_1} \left(\hat{\alpha} - \delta_0 - \frac{\lambda_2}{k_1} \right) (1 - e^{-k_1 T}) \\ - \frac{k_1}{k_1 - k_2} (\alpha_0 - \hat{\alpha}) \left(\frac{1}{k_2} (1 - e^{-k_2 T}) - \frac{1}{k_1} (1 - e^{-k_1 T}) \right) \end{aligned} \quad (76)$$

$$\begin{aligned}
& + \frac{1}{2} \frac{\sigma_3^2 \kappa_1^2}{(k_1 + \kappa_2)^2} \left[\frac{T}{(k_1 + \kappa_2)^2} + \frac{T^2}{k_1 + k_2} + \frac{T^3}{3} + \frac{4}{(k_1 + K_2)^3} (1 - e^{-(k_1 + k_2)T}) \right. \\
& \quad \left. + \frac{1}{2(k_1 + K_2)^3} (1 - e^{-2(k_1 + k_2)T}) + \frac{2}{(k_1 + K_2)^2} T e^{-(k_1 + k_2)T} \right] \\
& \quad + \frac{1}{2} \frac{\sigma_2^2}{\kappa_1^2} \left[T + \frac{1}{2k_1} (1 - e^{-k_2 T}) - \frac{2}{k_1} (1 - e^{-k_1 t}) \right] + \frac{1}{2} \sigma_1^2 T \\
& \quad + \frac{\rho_2 \sigma_2 \sigma_3}{\kappa_1 + \kappa_2} \left[\frac{T^2}{2} + \frac{T}{k_1} e^{-k_1 t} - \frac{1}{k_1^2} + \frac{1}{k_1^2} e^{-k_1 t} + \frac{T}{k_1 + k_2} \right. \\
& \quad - \frac{1}{(k_1 + k_2)^2} (1 - e^{-(k_1 + k_2)T}) - \frac{1}{k_1(k_1 + k_2)} (1 - e^{-k_1 T}) \\
& \quad \left. - \frac{1}{(k_1 + k_2)(2k_1 + k_2)} (1 - e^{-(2k_1 + k_2)T}) \right] \\
& \quad - \frac{\rho_3 \sigma_1 \sigma_3 \kappa_1}{\kappa_1 + \kappa_2} \left[\frac{T^2}{2} + \frac{T}{k_1 + k_2} + \frac{1}{(k_1 + k_2)^2} (1 - e^{-(k_1 + k_2)T}) \right] \\
& \quad - \frac{\rho_1 \sigma_1 \sigma_2}{\kappa_1} \left[T - \frac{1}{k_1} (1 - e^{-k_1 t}) \right]
\end{aligned}$$

MODELO NA FORMA ESPAÇO DE ESTADO

Da mesma forma que no modelo de dois fatores, apresentamos o modelo na forma espaço de estado.

Sejam as equações que definem o modelo na medida real:

$$dX_t = \left(\mu - \frac{1}{2} \sigma_1^2 \right) dt - \delta_t dt + \sigma_1 dB_{1,t} \quad (77)$$

$$d\delta_t = \kappa_1 (\alpha_t - \delta_t) dt + \sigma_2 dB_{2,t} \quad (78)$$

$$d\alpha_t = \kappa_2 (\bar{\alpha} - \alpha_t) dt + \sigma_3 dB_{3,t} \quad (79)$$

Discretizando as variáveis de estado:

$$X_t = X_{t-1} + \left(\mu - \frac{1}{2} \sigma_1^2 \right) \Delta t - \delta_{t-1} \Delta t + \eta_1 t \quad (80)$$

$$\delta_t = \delta_{t-1} + \kappa_1 \Delta t \alpha_{t-1} - k_1 \Delta t \delta_{t-1} + \eta_2 t \quad (81)$$

$$\alpha_t = \alpha_{t-1} + \kappa_2 \bar{\alpha} \Delta t - k_2 \Delta t \alpha_{t-1} + \eta_3 t \quad (82)$$

Temos as equações de transição do modelo:

$$\begin{bmatrix} X_t \\ \delta_t \\ \alpha_t \end{bmatrix} = \begin{bmatrix} (\mu - \frac{1}{2}\sigma_1^2)\Delta t \\ 0 \\ \kappa_2\bar{\alpha}\Delta t \end{bmatrix} + \begin{bmatrix} 1 & -\Delta t & 0 \\ 0 & (1 - k_1\Delta t) & k_1\Delta t \\ 0 & 0 & 1 - k_2\Delta t \end{bmatrix} \begin{bmatrix} X_{t-1} \\ \delta_{t-1} \\ \alpha_{t-1} \end{bmatrix} + \begin{bmatrix} \eta_1 t \\ \eta_2 t \\ \eta_3 t \end{bmatrix} \quad (83)$$

Reescrevendo a equação (83), temos que:

$$[X_t \delta_t \alpha_t]^\top = c_t + Q_t [X_{t-1} \delta_{t-1} \alpha_{t-1}]^\top + \eta_t \quad (84)$$

onde η_t é serialmente não correlacionado, com médias e variâncias sob a medida $\mathbb{P}$ dadas por:

$$\begin{aligned} E^{\mathbb{P}}(\eta_t) &= 0 \\ Var^{\mathbb{P}}(\eta_t) &= \begin{bmatrix} \sigma_1^2 \Delta t & \rho_1 \sigma_1 \sigma_2 \Delta t & \rho_3 \sigma_1 \sigma_3 \Delta t \\ & \sigma_2^2 \Delta t & \rho_2 \sigma_2 \sigma_3 \Delta t \\ & & \sigma_3^2 \Delta t \end{bmatrix} \end{aligned} \quad (85)$$