



EXPECTATIVAS DE PREÇOS DO GAS NATURAL E OS PREÇOS DO PETRÓLEO NO MERCADO INTERNACIONAL: UM ESTUDO EMPÍRICO BASEADO EM UM MODELO DINÂMICO

André Assis de Salles

Universidade Federal do Rio de Janeiro
Centro de Tecnologia – Bloco F – sala F 101 – Ilha do Fundão – Rio – Brasil
e-mail: as@ufrj.br

Ana Beatriz Mendes Campanate

Universidade Federal do Rio de Janeiro
e-mail: anacampanate@poli.ufrj.br

RESUMO

Este trabalho tem como objetivo estudar a relação entre os preços do gás natural e do petróleo no mercado internacional, propondo investigar sua natureza e equilíbrio de longo prazo, através do desenvolvimento de modelos econométricos adequados para a determinação das expectativas futuras das principais referências de preço do gás natural ou de seus retornos. Com este propósito, as séries temporais dos preços de referência do petróleo bruto e do gás natural são utilizadas na feitura de testes estatísticos com o objetivo de verificar as hipóteses subjacentes aos modelos dinâmicos autoregressivos apropriados. A heteroscedasticidade condicional e a não-normalidade das séries dos retornos dos preços são consideradas na construção dos modelos econométricos utilizados. Para atingir o objetivo deste trabalho foram coletados os preços semanais de gás natural e de petróleo bruto negociados no mercado internacional.

PALAVRAS CHAVE. Gás Natural, Petróleo, Cointegração, Causalidade, Modelo ARDL.

PO na Área de Petróleo e Gás, Estatística, Gestão Financeira.

ABSTRACT

This work aims to study the relationship between natural gas and crude oil prices in the international market, proposing to investigate its nature and long term equilibrium, through the development of adequate econometric models for determining future expectations of major natural gas price benchmarks. In order to accomplish this, time series for both benchmark crude oil and natural gas prices are subjected to statistical tests with the purpose of verifying the underlying hypotheses behind the appropriate autoregressive dynamic models. The conditional heteroskedasticity and non-normality of the return series are considered when elaborating the econometric models used. To reach the purpose of this work weekly natural gas and crude oil prices benchmarks traded in the international market were collected.

KEYWORDS. Natural Gas Prices, Crude Oil Prices, Cointegration, Causality, ARDL Model.

OR in Oil and Gas, Statistics, Financial Management.



1. Introdução

Além da utilização expressiva na indústria, como fonte de calor e energia, e como um insumo para a produção, o gás natural tem uso relevante no sistema de aquecimento residencial e comercial. Essa ampla gama de usos para o gás natural torna este um recurso crítico para as economias nacionais. Uma implicação dessa versatilidade consiste no fato de que variações na demanda para certo uso do gás natural podem ter impacto significativo sobre os preços das demais aplicações. De fato, desde o início da última década, têm sido notória a volatilidade associada aos preços de gás. Após atingir níveis recordes com os picos de 2005 e 2008, os preços após a eclosão da crise financeira mundial decaíram de forma contínua. Parte dessa volatilidade está associada às dificuldades no transporte de gás onde a infraestrutura de gasoduto não está consolidada. Dadas essas limitações não há mercado global para o gás natural, e os preços locais podem depender fortemente da produção e disponibilidade regionais. Assim, como observado pela Union of Concerned Scientists (2014), a determinação de preços fica sujeita, até certo ponto, às forças de oferta e demanda locais. Entretanto, em relação à questão de formação de preços, um aspecto adicional importante a ser considerado são as relações de longo prazo no mercado de energia. Como fontes de energia podem ser substituídas para usos finais, é razoável supor que os preços da energia dessas diferentes fontes estejam de alguma maneira associados. Desse modo, muitas pesquisas sobre a cointegração entre os preços de diferentes commodities de energia foram desenvolvidas nos últimos anos. Dentre essas pesquisas muitas procuram explicar a relação entre os preços de gás natural e do petróleo bruto. A existência de uma relação estável de longo prazo entre os preços dessas duas *commodities* foi mostrada nos resultados de várias dessas pesquisas. Como destaca Hartley e Medlock (2014), os combustíveis competem entre si, embora algumas pesquisas tenham apresentado evidências contrárias a existência dessa relação. A existência ou não de uma relação estável entre preços de petróleo bruto e gás natural têm importantes consequências econômicas e sua verificação é fundamental para os formuladores de política econômica e pode permitir investimentos lucrativos para empresas, operações de arbitragem, estratégias especulativas e de *hedging*. As variações na relação de preços entre fontes de energia podem ser relevantes também para tomadores de decisões e formuladores de políticas econômicas. É possível que tais variações frustrem as tentativas de promover o uso de uma fonte de energia em detrimento de outra. A literatura empírica tratando da relação entre preços do gás natural e do petróleo bruto é extensa e abrangente. Nas décadas recentes muitos estudos têm sido desenvolvidos para verificar a existência de uma relação de longo prazo entre os preços de gás natural e do petróleo bruto, e a adequação de modelos de séries temporais para descrever essa relação. Dentre essas pode-se citar a pesquisa de Hartley e Medlock (2007), a pesquisa de Brown e Yücel (2007), o trabalho de Panagiotidis e Rutledge (2007), o trabalho de Leykam e Frauendorfer (2008), o trabalho mais recente de Brigida (2014), outra pesquisa recente de Nick e Thoenes (2014) e por último a taxonomia de Frey *et al.* (2009).

Este trabalho tem por objetivo estudar da relação entre os preços de gás natural e do petróleo bruto negociados no mercado internacional, investigando a natureza e equilíbrio dessa relação, através do desenvolvimento de modelos econométricos adequados para determinação da expectativas de tendências dos preços, ou dos retornos dos preços, dos principais *benchmarks* dos preços do gás natural negociados no mercado internacional. Assim, as séries temporais de referências de preços do gás natural e de petróleo são analisadas por meio de testes estatísticos, com o intuito de verificar pressupostos necessários para elaboração de modelos dinâmicos autoregressivos apropriados. A heteroscedasticidade condicional e a não normalidade das séries de retornos são consideradas na construção desses modelos.

Além da presente introdução, este trabalho apresenta na Seção 2 a abordagem metodológica utilizada para atingir os objetivos da pesquisa, e a Seção 3 descreve a amostra utilizada. A Seção 4 mostra e comenta os resultados obtidos com os procedimentos econométricos utilizados no desenvolvimento da pesquisa que originou este trabalho. Por fim, a Seção 5 trata das considerações finais do estudo seguida das referências bibliográficas utilizadas.



2. Abordagem Metodológica Adotada

As pesquisas que envolvem séries temporais apresentam problemas que trazem vários desafios. Na sua maior parte as pesquisas empíricas com esse tipo de dado se assume que a série temporal subjacente é estacionária, o que significa, de forma ampla, que suas média, variância e autocovariância não variam sistematicamente ao longo do tempo. Como modelos envolvendo séries temporais são comumente usados para previsões, é importante verificar se o pressuposto da estacionariedade pode ser considerado válido ou se as inferências obtidas podem ser consideradas válidas se a série temporal subjacente não for estacionária. Um dos problemas mais comuns inerentes a séries não estacionárias se refere à regressão espúria, definida como uma regressão com coeficiente de determinação R^2 significativo entre duas variáveis que, *a priori*, não possuem relação alguma. Yule (1926) verificou que esse fenômeno permanece em séries não estacionárias, mesmo que a amostra utilizada seja muito grande. Sendo d a estatística de Durbin-Watson, de acordo com Granger e Newbold (1976), se $R^2 > d$ deve se suspeitar que a regressão estimada é espúria.

A forma mais direta de verificar a existência de estacionariedade seria, portanto, realizar o teste t ; entretanto, sob a hipótese nula, o valor t do coeficiente estimado de Y_{t-1} não segue a distribuição t , mesmo em grandes amostras. Uma alternativa foi elaborada por Dickey e Fuller (1979) que mostraram que, sob a hipótese nula de $\delta = 0$, a estatística t do coeficiente estimado de Y_{t-1} segue a estatística τ . Os valores críticos dessa estatística foram calculados por simulações de Monte Carlo, esse teste é conhecido como teste de Dickey-Fuller (DF). Uma variação desse teste é o teste de Dickey-Fuller Aumentado, com a adição de valores defasados da variável ΔY_t , que pode ser descrito através da equação (1), abaixo:

$$\Delta Y_t = \beta_1 + \beta_2 t + \delta Y_{t-1} + \sum_{i=1}^m \alpha_i \Delta Y_{t-i} + \varepsilon_t, \quad (1)$$

onde ε_t é um termo estocástico β_1 e β_2 são parâmetros. O número de *lags* da variável ΔY_t é determinado empiricamente, de forma que o termo de erro não possua correlação serial. O termo $\beta_2 t$ inclui uma tendência determinística, enquanto o termo β_1 se aplica para o caso em que há suspeita de que o processo estocástico seja um passeio aleatório com *drift*, em vez de passeio aleatório puro.

De forma geral, se uma série temporal precisa ser diferenciada n vezes para se tornar estacionária, é dito que tal série é integrada de ordem n . Nesse caso, a representação utilizada para a ordem de integração é $Y_t \sim I(n)$. Como observado em Wooldridge (2015), se a série em questão for estacionária, não sendo necessária a diferenciação, diz-se que a série é integrada de ordem zero, ou $Y_t \sim I(0)$. Sendo duas séries temporais X_t e Y_t ambas $I(1)$ para a regressão de Y_t sobre X_t tem-se as equações (2) e (3) abaixo.

$$Y_t = \beta_1 + \beta_2 X_t + u_t \quad (2)$$

$$u_t = Y_t - \beta_1 - \beta_2 X_t \quad (3)$$

Uma situação interessante ocorre quando u_t isolado na equação (3) é $I(0)$, ou seja, quando a combinação linear de duas séries temporais não estacionárias é estacionária. Nesse caso, a combinação linear cancela as tendências estocásticas das duas séries. Como resultado, a regressão de Y_t em X_t não é espúria e as duas variáveis são cointegradas. Enquanto a possibilidade de regressões espúrias estimula um certo cuidado em relação à utilização de variáveis $I(1)$, tornando sua diferenciação a ação mais segura, explorar a relação de cointegração torna o escopo das questões que podem ser respondidas mais abrangente. Pela perspectiva econômica, duas variáveis cointegradas têm uma relação de longo prazo de equilíbrio entre si, o que permite a



utilização da metodologia de regressão tradicional. Assim se o termo u_t da equação (2) é estacionário as séries Y_t e X_t são cointegradas. Para testar se esse é o caso, basta simplesmente aplicar o teste de Dickey-Fuller (DF) ou de Dickey-Fuller Aumentado (ADF) sobre os resíduos da equação de cointegração que nesse contexto, são chamados, respectivamente, de teste de Engle-Granger (EG) ou de Engle Granger Aumentado (AEG), como observado por Gujarati (2004). A distinção importante é que como os resíduos estimados da equação (2) são baseados no parâmetro de cointegração β_2 estimado, os valores críticos dos testes DF e ADF não são apropriados. Assim, Engle e Granger (1987) calcularam novos valores críticos para o teste de cointegração. Se Y_t e X_t não são cointegradas, a regressão da equação (2) é espúria e não tem sentido: não há relação de longo prazo entre Y e X . É possível ainda utilizar as variáveis na primeira diferença, ΔY_t e ΔX_t . Entretanto, nesse caso a interpretação se refere à explicação da diferença em Y em termos da diferença em X , e essa regressão nada diz necessariamente sobre a relação de tais variáveis no nível. Por outro lado, se as variáveis são cointegradas, os modelos dinâmicos mais gerais podem ser utilizados, como pode ser visto com mais detalhe em Wooldridge (2009).

É possível definir um modelo dinâmico de séries temporais denominado modelo autoregressivo de defasagens distribuídas (ARDL). Como observado por Pickup (2014), tal modelo é uma combinação dos modelos autoregressivo e de defasagens distribuídas já descritos, contendo defasagens das variáveis dependentes e independentes como mostrado na equação (4), a seguir, que representa um modelo $ARDL(p, m)$, com p defasagens da variável dependente e m defasagens da variável independente:

$$y_t = \alpha_0 + \sum_{j=1}^p \alpha_j y_{t-j} + \beta_1 x_t + \sum_{i=1}^m \beta_{i+1} x_{t-i} + \varepsilon_t \quad (4)$$

onde ε_t é um termo estocástico ou ruído branco, como se pode observar em Pickup (2014).

Os modelos definidos anteriormente têm como pressuposto a homocedasticidade dos erros. Entretanto, é possível se observar autocorrelação da variância do termo de erro no instante t em relação a seus valores passados nas análises de séries temporais de retornos dos preços de ativos financeiros e de *commodities*, ou uma heteroscedasticidade condicional autorregressiva, e que pode ser explicada ou representada pelo modelo ARCH apresentado no trabalho seminal de Engle (1982) e subsequentemente difundido na literatura econométrica para estimação da volatilidade de variáveis financeiras ou macroeconômicas originou diversos outros modelos utilizados na literatura econométrica formando o conjunto de modelos da família ARCH. Dentre esses tem-se o modelo heteroscedasticidade condicional autorregressiva generalizada ou GARCH apresentado por Bollerslev (1986).

As séries temporais financeiras e de mercados de energia apresentam alta volatilidade e são, em sua maioria, passeio aleatórios. Assim, seria natural modelar sua primeira diferença, que costuma ser estacionária. Entretanto, as primeiras diferenças de tais séries frequentemente apresentam grande volatilidade, sugerindo que a variância dessas séries temporais financeiras varia ao longo do tempo.

Considerando o seguinte modelo dinâmico representado pela equação (5):

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta_1 x_t + \varepsilon_t \quad (5)$$

Para modelar a dinâmica da variância de ε_t , é importante estabelecer a distinção entre as variâncias condicional e incondicional. É assumido que a variância incondicional de ε_t é constante e sem correlação serial, como mostrado na equação (12):

$$E(\varepsilon_t \varepsilon_{t-s}) = \begin{cases} \sigma^2, & \text{para } s = 0 \\ 0, & \text{caso contrário} \end{cases} \quad (6)$$



Embora o pressuposto de variância incondicional constante seja mantido, é permitido que a variância condicional sobre valores passados dos erros varie ao longo do tempo. Assim, os erros são agora modelados como um processo com variância condicional, ou seja, como uma função dos valores passados da variância. Uma forma comum de fazer isso é modelando o quadrado dos erros como de um processo autoregressivo de ordem m , que define o processo ARCH (m):

$$\varepsilon_t^2 = \zeta + \phi_1 \varepsilon_{t-1}^2 + \phi_2 \varepsilon_{t-2}^2 + \dots + \phi_m \varepsilon_{t-m}^2 + \omega_t \quad (7)$$

onde ω_t é um processo de ruído branco com média zero e variância incondicional constante. Modelos de séries de tempo que incluem tal processo na modelagem dos erros recebem o nome de modelos ARCH (ver Pickup (2014)). Segundo Engle (1982), a equação (7) pode ser estimada pelos métodos dos mínimos quadrados ordinários ou de máxima verossimilhança, gerando resultados consistentes e similares. Desse modo o teste ARCH é útil para verificar a existência de correlação na variância do erro, e consiste em estimar a equação (7) e empregar a estatística F para analisar a seguinte hipótese nula:

$$H_0: \phi_1 = \phi_2 = \dots = \phi_m = 0 \quad (8)$$

No caso da hipótese nula descrita pela equação (8) não ser rejeitada, tem-se $\varepsilon_t^2 = \zeta$, ou seja, a variância é constante e não há efeito ARCH. Uma das variações do modelo ARCH que se tornaram mais populares consiste no modelo GARCH, detalhado em Engle e Bollerslev (1986), que define o processo de geração da variância condicional como um processo do tipo ARMA. O modelo *GARCH*(p, q) pode ser descrito pela equação (9), a seguir.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 \quad (9)$$

Como observado por Gujarati (2004), a existência de uma relação entre duas variáveis, verificada pela análise de regressão não prova causalidade ou a direção da influência. Uma situação particular se materializa no caso de regressões envolvendo séries temporais: se o evento A precede o evento B , é possível que A cause B , mas não há possibilidade de o evento B causar o evento A , já que, evidentemente, o futuro não pode causar o passado. Granger (1969) partiu deste princípio na elaboração de um teste causalidade que se tornou popular na literatura econométrica. Dadas duas séries temporais X_t e Y_t a questão a ser considerada é se X_t “causa” Y_t , ou se Y_t “causa” X_t . Assumindo-se que a informação relevante para a predição dessas variáveis está contida inteiramente em suas séries de tempo. A condução do teste envolve a estimação das seguintes regressões:

$$X_t = \sum_{i=1}^n \alpha_i Y_{t-i} + \sum_{j=1}^n \beta_j X_{t-j} + u_{1t} \quad (10)$$

$$Y_t = \sum_{i=1}^n \lambda_i Y_{t-i} + \sum_{j=1}^n \delta_j X_{t-j} + u_{2t} \quad (11)$$

onde os termos residuais u_{1t} e u_{2t} devem ser não correlacionados entre si. A equação (10) propõe que o valor contemporâneo de X está relacionado aos seus próprios valores passados e aos de Y , enquanto a equação (11) prevê um comportamento similar para Y . Dessa forma, em uma regressão de Y em outras variáveis, incluindo seus próprios valores defasados, se a inclusão de valores defasados de X provoca uma melhora significativa na previsão de Y , é possível dizer que X causa Y no sentido Granger. No entanto, alguns pressupostos devem ser satisfeitos para garantir a validade das conclusões advindas do teste de causalidade de Granger. Dentre esses, as duas variáveis X e Y devem ser estacionárias e os termos estocásticos não devem ter correlação entre si. O número de lags a serem introduzidos no teste é uma questão empírica, a ser resolvida por meio de critérios de informação, como o critério de Akaike ou de Schwarz. É importante



notar, como observa Gujarati (2004), que a direção da causalidade pode depender fortemente do número de *lags* incluídos.

3. Amostra – Dados Utilizados

Os dois principais *benchmarks* dos preços do gás natural no mercado internacional são o Henry Hub, nos EUA, e o National Balancing Point (NBP), no Reino Unido. Em operação desde o final da década de 1990, o mercado de gás NBP é o mais antigo da Europa. Sendo o gás natural de qualquer local do país, desde que esteja contido no sistema nacional de transmissão, é tido como gás NBP. E assim, promove-se a reunião de compradores e vendedores simplificando o comércio. O preço do NBP é amplamente utilizado como indicador do mercado atacadista de gás natural da Europa. Enquanto o Henry Hub, localizado em Louisiana, nos EUA, é o ponto de comércio de gás natural mais conhecido da América do Norte e representa o ponto de entrega padrão de contratos futuros de gás natural negociados na New York Mercantile Exchange (NYMEX). No que se refere ao petróleo bruto as dois principais benchmarks do mercado internacional são o petróleo tipo Brent e o petróleo bruto do tipo WTI. O petróleo Brent designa todo o petróleo extraído no Mar do Norte e comercializado na Bolsa de Londres, enquanto o petróleo do tipo WTI, tem sua origem na principal região petrolífera dos EUA e é negociado na Bolsa de Nova York.

A amostra utilizada neste trabalho contém séries dos preços *spot* semanais do NBP e do Henry Hub (HH), coletadas no Bloomberg *web-site*, e dos preços semanais dos mercados *spot* do petróleo bruto do tipo Brent e do tipo WTI, coletados site da Agência Norte-Americana de Estatísticas de Energia (EIA). A amostra compreende o período de setembro de 2007 a janeiro de 2016, perfazendo quatro séries temporais de 437 observações. Os preços das cotações semanais do petróleo bruto do tipo Brent e do tipo WTI foram coletadas em dólares norte-americanos por barril, enquanto os preços das séries de preços do gás natural HH e NBP foram convertidos em dólares norte-americanos por milhões de BTU.

As séries de preços de petróleo e do gás natural apresentam comportamento similar. A estacionariedade destas séries é improvável, uma vez que o gráfico sugere grandes mudanças na média ao longo do período considerado, com uma tendência de decaimento para ambos os gases após meados de 2012. A variância também não parece ser constante ao longo do tempo. É possível identificar movimentos cíclicos na série; entretanto, não parecem se repetir em intervalos regulares ou possuir a mesma forma para justificar o emprego de *dummies* sazonais em modelagens. A suspeita de raiz unitária é mais uma vez levantada, e os movimentos semelhantes parecem sugerir correlação. Pode-se observar, ainda, que os séries temporais estudadas exibem comportamento clássico de séries de tempo financeiras, isto é, o padrão de passeio aleatório em nível. E como observa Gujarati (2004), a série obtida com as primeiras diferenças das observações dessas séries temporais costuma ser estacionária.

A Tabela 1 adiante apresenta uma descrição preliminar dessas séries temporais através de um resumo estatístico, com medidas resumo e os resultados dos testes da hipótese de normalidade e de estacionariedade dos dados. As médias de preços do período estudado diferem para os *benchmarks* do gás natural e do petróleo bruto, respectivamente. A observação das médias e dos desvios padrões permite se inferir que o gás natural HH apresenta uma volatilidade maior do que o NBP enquanto as cotações dos tipos de petróleo Brent e WTI apresentam comportamento semelhantes com uma diferença na cotação dos seus preços. No que se refere ao coeficiente de assimetria todos diferem do da distribuição normal, sendo o do HH o único positivo e o mais distante de zero ou de um conjunto de observações simétrica como a de uma distribuição normal, mesmo acontece com o coeficiente de curtose estimado dos preços do HH que difere das outras séries históricas de preços utilizadas neste trabalho. A série temporal dos preços do HH mostra uma leptocurtose, enquanto as demais séries de preços as estimativas dos coeficientes de curtose estão abaixo de 3 indicando serem essas platicúrticas. O teste de Jarque-Bera para normalidade confirma a hipótese de não normalidade, característica comum nas séries de preços de ativos financeiros e de *commodities* em geral. Como observado anteriormente,



as séries de preços selecionadas não apresentam estacionariedade e sua variação ou primeira diferença deve ser estacionária. Assim a partir das séries dos preços coletados foram calculados as variações desses preços ou séries de retornos dos preços.

	NBP	HH	Brent	WTI
<i>Média</i>	8,25	4,45	89,43	83,41
<i>Mediana</i>	8,60	3,89	97,13	88,50
<i>Máximo</i>	16,00	13,20	141,07	142,52
<i>Mínimo</i>	2,90	1,68	27,76	29,19
<i>Desvio Padrão</i>	2,35	2,06	25,58	22,26
<i>Assimetria</i>	-0,02	1,95	-0,49	-0,04
<i>Curtose</i>	2,47	7,13	2,11	2,75
<i>Jarque-Bera</i>	7,539	588,170	31,834	14,380
<i>(valor p)</i>	(0,023)	(0,000)	(0,000)	(0,001)
<i>ADF valor p</i>	0,389	0,199	0,382	0,339
<i>Nº Observações</i>	437	437	437	437

Tabela 1: Resumo Estatístico das Séries dos Preços Semanais

	NBP	HH	Brent	WTI
<i>Média</i>	-0,0012	-0,0023	0,0020	-0,0021
<i>Mediana</i>	0,0000	-0,0022	-0,0007	-0,0013
<i>Máximo</i>	0,3478	0,3010	0,2002	0,2512
<i>Mínimo</i>	-0,3455	-0,3093	-0,1646	-0,1910
<i>Desvio Padrão</i>	0,0857	0,0634	0,0424	0,0450
<i>Assimetria</i>	0,2092	0,0747	-0,0968	-0,0619
<i>Curtose</i>	6,0357	6,6423	5,7759	6,9453
<i>Jarque-Bera</i>	170,982	588,170	140,992	283,699
<i>(valor p)</i>	(0,000)	(0,000)	(0,000)	(0,000)
<i>ADF valor p</i>	0,000	0,000	0,000	0,000
<i>Nº Observações</i>	437	437	437	437

Tabela 2: Resumo Estatístico das Séries de Retornos dos Preços Semanais

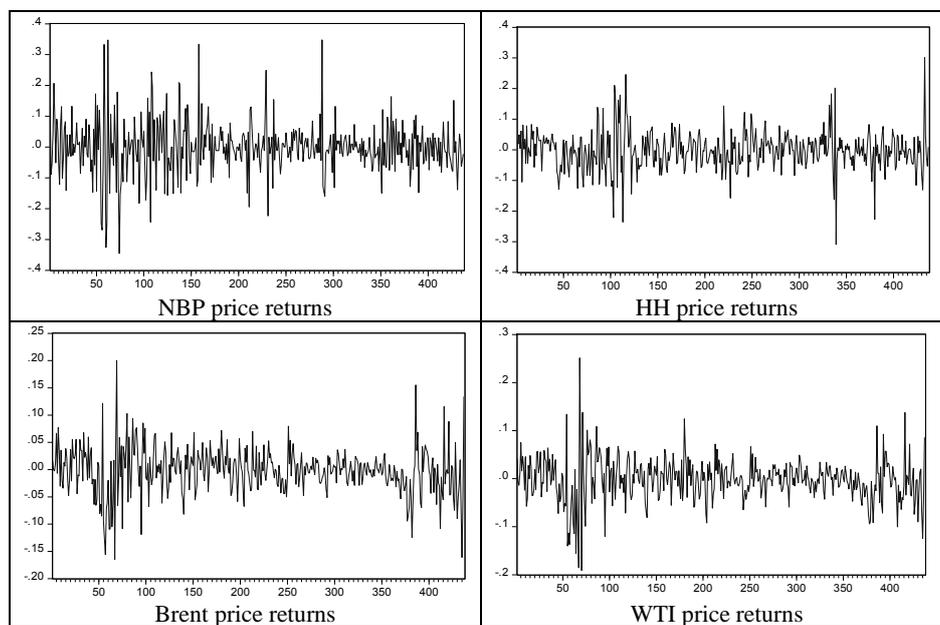


Figura 1: Retornos dos Preços Semanais do Petróleo (US\$ / Barril) e Gás Natural (US\$ / MMBtu)



Os retornos logarítmicos semanais foram calculados para cada uma das quatro séries de preços utilizadas neste trabalho por meio da fórmula a seguir:

$$R_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \quad , \quad (12)$$

onde R_t representa o retorno no instante t , e P_t , o preço no instante t . Como as séries de preço originais são integradas de ordem 1, ou $I(1)$, as novas séries advindas dessa transformação serão estacionárias, ou $I(0)$. A Tabela 2 mostra um resumo estatístico das séries dos retornos dos *benchmarks* dos preços do petróleo e do gás natural, respectivamente, assim como os resultados dos testes de normalidade e de estacionariedade dessas séries. As médias das séries apresentam valores similares, e a série que apresentou maior variabilidade, de acordo com o desvio padrão, foi a do gás natural NBP. Nota-se que todas as médias apresentaram valores negativos. Além disso, é possível rejeitar a hipótese de normalidade a um nível de significância menor que 1%. De acordo com o teste ADF, é possível rejeitar a hipótese de raiz unitária para as séries de retornos a um nível de significância menor que 1%. Assim, como se suspeitava pela análise dos gráficos e pelo comportamento esperado de séries financeiras, as séries de preços de petróleo e gás natural são integradas de ordem 1, ou $I(1)$, pois suas primeiras diferenças constituem séries estacionárias. A Figura 1 mostra as séries de retornos dos preços utilizadas para elaboração deste trabalho. Analisando os gráficos da Figura 1 é possível identificar um comportamento comum de séries de retornos, em que observações altamente positivas ou negativas aparecem em aglomerados, ou seja, é aparente o fenômeno de concentração de períodos de alta volatilidade seguido por outros de relativa calma. O que se traduz em indícios de que a volatilidade do período atual está relacionada à de períodos passados, o que pressupõe heteroscedasticidade condicionada autoregressiva. Tal proposição pode ser confirmada através do teste ARCH sobre as séries temporais de retorno dos preços.

4. Análise dos Resultados Obtidos

A relação entre os preços de gás natural e do petróleo bruto foi investigada para os mercados dos EUA e do Reino Unido, por meio da condução do teste de cointegração de Engle-Granger para as séries de preços. Considerando a posição geográfica do mercado norte-americano, os testes foram aplicados separadamente entre as referências de preços no mercado norte americano de gás natural e do petróleo bruto, respectivamente, HH e WTI. O modelo utilizado para o teste de raiz unitária de Dickey-Fuller tem como variável dependente a primeira diferença dos termos estocásticos, ou resíduos, da regressão dos preços HH sobre os preços do WTI, ou seja, a regressão linear com intercepto do primeira diferença dos resíduos sobre os resíduos com uma defasagem. Esse modelo, para explicar as primeiras diferenças ou variações dos resíduos, necessário para implementação do teste de Engle-Granger para a hipótese de cointegração foi o mais apropriado, de acordo com o critério de Akaike (AIC). O mesmo ocorre com os preços de referência do gás natural e do petróleo bruto do mercado do Reino Unido, respectivamente, NBP e Brent, onde o modelo mais apropriado foi o mesmo mas sem o intercepto. O modelo descrito acima pode ser representado pela seguinte equação para o teste de cointegração entre os preços do HH e do WTI:

$$\Delta u_t = \beta_1 + \delta u_{t-1} + \epsilon_t \quad (13)$$

E o modelo utilizado para testar a hipótese de cointegração entre os preços do NBP e o petróleo bruto do tipo Brent difere da equação acima por não ter o termo β_1 . A partir dos testes realizados os resultados da estatística de teste de Engle-Granger para as variáveis HH e WTI foi de -4,3992, com um valor p de 0,0074, e o para as variáveis NBP e Brent foi de -3,6557, com um valor p de 0,0018. Os valores críticos para o teste são -3,922, -3,350 e -3,054 para os níveis de significância de 1%, 5% e 10%, respectivamente. Assim os resultados dos testes não permitem se aceitar a



hipótese nula de não cointegração entre os preços do gás natural e do petróleo bruto tanto para o mercado norte-americano quanto para o mercado do Reino Unido, mostrando a existência de uma relação de longo prazo entre os preços do HH e o petróleo do tipo WTI assim como entre os preços do NPB e do petróleo do tipo Brent. Afastando a possibilidade de uma regressão espúria entre os preços do gás natural e do petróleo bruto para os preços de referência praticados nos mercados norte-americano e do Reino Unido.

Outra hipótese que deve ser testada é da causalidade entre as séries temporais dos retornos das variáveis aqui estudadas, uma vez que a existência ou não da relação de causalidade pode ser determinante na construção de modelos que tenham a finalidade de explicar os preços ou retornos do gás natural negociado no mercado internacional. Deve-se lembrar que teste de causalidade de Granger, apresentado na metodologia deste trabalho, tem como pressuposto a estacionariedade das variáveis envolvidas, hipótese que pode ser aceita para as séries temporais de retornos dos preços. Os resultados dos testes de causalidade realizados entre as séries de retornos os preços do HH e o petróleo do tipo WTI, referência do mercado norte-americano, e entre os retornos dos preços do NPB e do petróleo do tipo Brent, referência do mercado do Reino Unido, com defasagens de 1 a 12, estão listados a seguir na Tabela 3. Pode-se observar, através do valor p que para todas as defasagens consideradas a hipótese de causalidade dos retornos do WTI nos retornos do HH não pode ser rejeita, o que não acontece na direção oposta o que condiz com a estimação de um modelo estocástico que relacione essas duas variáveis e tenha como variável dependente os retornos do HH. Por outro lado, os resultados parecem sugerir uma relação de causalidade unilateral ou bilateral dos retornos do Brent para os do NBP, o que vai de encontro com a percepção comumente estabelecida de que os preços de petróleo são um fator na formação de preços de gás natural.

Lags	HH → WTI		WTI → HH		NBP → Brent		Brent → NBP	
	Estatística F	Valor p						
1	1,9461	0,1637	4,5946	0,0326	13,1024	0,0003	0,0941	0,7592
2	1,2045	0,3009	3,0446	0,0486	6,8624	0,0012	2,3708	0,0946
3	0,6755	0,5674	2,5863	0,0527	5,3943	0,0012	1,5485	0,2013
4	0,8822	0,4744	2,2729	0,0607	4,1854	0,0025	1,2799	0,2771
5	0,5932	0,7052	1,8291	0,1059	3,3652	0,0054	1,0560	0,3843
6	0,5245	0,7898	1,9232	0,0758	3,3183	0,0033	2,0818	0,0543
7	0,4701	0,8562	1,9863	0,0557	3,5623	0,0010	1,9493	0,0607
8	0,5591	0,8113	2,5955	0,0089	5,4900	0,0000	2,1218	0,0328
9	0,6236	0,7771	2,3104	0,0153	5,1628	0,0000	2,0720	0,0309
10	0,7164	0,7092	2,0425	0,0281	5,0606	0,0000	1,8944	0,0443
11	1,4361	0,1539	1,7859	0,0544	5,0482	0,0000	1,6453	0,0839
12	1,6879	0,0670	1,6574	0,0740	4,7386	0,0000	2,0151	0,0219

Tabela 3: Resultados do Teste de Causalidade de Granger para Séries de Retornos

Utilizando até 12 defasagens para as variáveis procedeu-se a estimação do modelo ARDL mais adequado para explicar os retornos das séries de gás natural. Após estimativas de muitos modelos para explicar os retornos dos preços do HH o modelo mais adequado, de acordo com o critério de Akaike, foi o **ARDL(3,3)**. A partir deste modelo preliminar, um novo modelo foi proposto no qual as variáveis precedidas de parâmetros não significativos foram retiradas. Dessa forma, o modelo ARDL estabelecido para explicar as variações de preços do HH foi dado pela expressão mostrada a seguir:

$$HH_t = \alpha_1 HH_{t-1} + \alpha_3 HH_{t-3} + \beta_1 WTI_t + e_t \quad (14)$$

Enquanto para explicar os retornos dos preços do gás natural NBP, com mesmos critérios estabelecidos para o modelo estimado para o HH o modelo selecionado foi um **ARDL(2,8)**, ou seja, após a exclusão de variáveis precedidas com parâmetros estatisticamente não significativos obteve-se o modelo que pode ser representado pela equação mostrada a seguir:



$$NBP_t = \alpha_1 NBP_{t-1} + \alpha_2 NBP_{t-2} + \beta_1 Brent_t + e_t \quad (15)$$

Os termos estocásticos dos modelos obtidos apresentados foram testados para verificar a condição de heteroscedasticidade sugerida pelos gráficos das séries temporais dos retornos, mostradas na Figura 1 apresentada na Seção 3. Assim, o teste ARCH foi implementado para série de termos estocásticos dos respectivos modelos selecionados para explicar os retornos dos preços do gás natural nos mercados norte-americanos e do Reino Unido, através do seguinte modelo: $e_t^2 = \zeta + \phi_1 e_{t-1}^2$. Os resultados do teste mostram a existência do efeito ARCH, uma vez que a estatística F para os modelos referentes aos retornos dos preços do HH a foi de 21,4334 e aos retornos dos preços do NBP foi de 30,3193, respectivamente, sendo os valores p bem próximos de zero, e desse modo é possível não se aceitar o efeito ARCH. Dessa forma, com o intuito de aprimorar os modelos, torna-se importante incluir os processos ARCH para tratar a heteroscedasticidade autoregressiva e possivelmente melhorar sua capacidade de previsão. Para tal fim, várias alternativas de processos **GARCH(p,q)** foram incluídos nos modelos de retornos do gás natural já apresentados, com **p** e **q** variando de zero até cinco. Os modelos mais apropriados, escolhidos pelo critério de Akaike, foram **GARCH(1,2)** e **GARCH(1,3)**, respectivamente, para os gases Henry Hub e NBP.

Outra consideração feita na ocasião do estudo das séries temporais utilizadas, foi a violação do pressuposto da normalidade das séries de retornos dos preços. Dessa forma, foi utilizada a distribuição *t* de Student na construção dos modelos para explicar os retornos dos preços do HH e do NBP.

Assim os modelos finais estimados para obter expectativas dos retornos dos preços e, por conseguinte, dos preços do gás natural no mercado internacional podem ser descritos da seguinte forma para os retornos do HH e para NBP, respectivamente, *RHH* e *RNBP*.

$$\begin{aligned} (RHH_t | I_{t-1}) &\sim Student(\mu_{HHt}; \sigma_{HHt}^2; \nu_{HH}) \\ \mu_{HHt} &= \alpha_1 RHH_{t-1} + \alpha_2 RHH_{t-2} + \alpha_3 RWTI_t + e_{HHt} \\ \sigma_{HHt}^2 &= \delta_1 + \delta_2 e_{t-1}^2 + \gamma_1 \sigma_{t-1}^2 + \gamma_2 \sigma_{t-2}^2 \end{aligned} \quad (16)$$

	<i>Coefficiente</i>	<i>Erro Padrão</i>	<i>Estatística t</i>	<i>Valor p</i>
α_1	0,1623	0,0462	3,5156	0,0004
α_2	-0,0923	0,0495	-1,8658	0,0621
α_3	0,1945	0,0556	3,4976	0,0005
δ_1	0,0002	0,0000	3,9469	0,0001
δ_2	0,0826	0,0634	3,4453	0,0006
γ_1	1,4729	0,0747	22,4583	0,0000
γ_2	-0,5969	6,6423	-12,3005	0,0000
ν	7,7660	2,5153	3,0875	0,0020
<i>RSS</i>	1,6509		<i>AIC</i>	-2,9372
<i>DW</i>	2,1201		<i>MSE</i>	0,0038

Tabela 4: Resultados das Estimativas do Modelo HH

$$\begin{aligned} (RNBP_t | I_{t-1}) &\sim Student(\mu_{NBPt}; \sigma_{NBPt}^2; \nu_{NBP}) \\ \mu_{NBPt} &= \alpha_1 RNBP_{t-1} + \alpha_2 RNBP_{t-2} + \alpha_3 RBrent_t + e_{NBPt} \\ \sigma_{NBPt}^2 &= \delta_1 + \delta_2 e_{NBPt-1}^2 + \gamma_1 \sigma_{NBPt-1}^2 + \gamma_2 \sigma_{NBPt-2}^2 + \gamma_3 \sigma_{NBPt-3}^2 \end{aligned} \quad (17)$$



	<i>Coefficiente</i>	<i>Erro Padrão</i>	<i>Estatística t</i>	<i>Valor p</i>
α_1	-0,1779	0,0489	-3,6399	0,0003
α_2	-0,0813	0,0487	-1,6680	0,0953
α_3	0,2624	0,0779	3,3681	0,0008
δ_1	0,0004	0,0002	2,2553	0,0241
δ_2	0,2581	0,0666	3,8771	0,0001
γ_1	1,0441	0,1378	7,5795	0,0000
γ_2	-0,8589	0,1703	-5,0442	0,0000
γ_3	0,5183	0,1085	4,7763	0,0000
ν	5,6189	1,7164	3,2738	0,0011
<i>RSS</i>	2,9815		<i>AIC</i>	-2,4454
<i>DW</i>	1,9845		<i>MSE</i>	0,0068

Tabela 5: Resultados das Estimativas do Modelo NBP

Assim, os modelos da média e da variância para o Henry Hub e NBP estão representadas pelas equações (16) e (17), respectivamente. Enquanto as Tabelas 4 e 5 exibem as estimativas obtidas para os parâmetros e os critérios de informação para os modelos dos retornos dos gases Henry Hub e NBP, respectivamente, que incluem o processo GARCH e se utilizam da distribuição t como alternativa à distribuição normal.

5. Considerações Finais

Este trabalho teve como objetivos a investigar da relação entre os preços de gás natural e de petróleo bruto, e verificar a relevância dos preços do petróleo na construção de um modelo adequado para determinação de expectativas de tendências dos retornos dos preços do gás natural negociado no mercado internacional. Da visão geral do trabalho pode-se observar que os objetivos foram atingidos. Através de procedimentos de inferência estatística clássica foi possível se verificar a existência de uma relação de equilíbrio de longo prazo, confirmada pelo teste de cointegração de Engle-Granger a nível de significância de 5%, tanto para o par Henry Hub e WTI, quanto para o par NBP e Brent. Desse mesmo modo, foi possível se testar a causalidade de Granger podendo-se inferir que os retornos passados do petróleo causam, de algum modo, as variações de preços do gás natural. Essa relação de causalidade se mostrou unilateral para as variações dos preços do WTI e do Henry Hub, enquanto para NBP e Brent que a hipótese de uma relação bilateral não pode ser rejeitada. O teste ARCH para os resíduos do modelo ARDL preliminar, desenvolvido entre as séries de retornos dos preços do gás natural e do petróleo, apontou para não rejeição da hipótese da existência de heteroscedastidade autoregressiva, sugerindo a que a inclusão processo GARCH nos modelos para precificação do gás natural deve ser obrigatória, para tratar a volatilidade dos preços. Assim como deve ser obrigatória a atenção da violação do pressuposto da normalidade da distribuição dos retornos, tratada aqui com a utilização da distribuição t de Student. Deve-se observar que a imposição de limites para a quantidade de *lags*, tanto para o modelo GARCH quanto para o ARDL, restringiu a gama de modelos. Como explicitado anteriormente, a estimação de possíveis modelos para esse trabalho foi restringida, no caso do ARDL (p, m) para ambos p e m com valor máximo de 12, e para o GARCH(p, q) variando de zero a cinco, com várias estimativas fora deste intervalo. Dadas as alternativas limitadas, é possível, com a mesma metodologia aqui adotada, que sejam obtidos modelos que apresentem melhor performance. Porém, deve-se ressaltar que os resultados obtidos apontam para um índice razoável de acerto nas expectativas de tendências dos preços ou na sinalização do mercado, ou seja, de indicações corretas de quando ao sinal, positivo ou negativo, dos retornos de cerca de 70% para os modelos escolhidos, a partir das séries dos termos



estocásticos obtidos na estimação dos modelos. Outro ponto que merece destaque é a não rejeição da hipótese da influência dos preços do petróleo bruto na determinação dos preços do gás natural.

Propostas de continuação deste trabalho incluem a ampliação das alternativas de modelos a serem testadas, por meio de limites de *lags* mais abrangentes ou variações dos modelos apresentados. O modelo GARCH, em especial, possui diversas variações, como o EGARCH ou TGARCH, que poderiam apresentar ajustes melhores. Os resultados das previsões poderiam, talvez, ser aprimorados por meio da inclusão de outras variáveis explicativas relevantes, uma vez que outros estudos mencionam que fatores como clima, condições de estocagem e sazonalidade têm papel relevante na relação entre preços de petróleo e de gás. Além disso, não deve ser descartada a utilização de outros enfoques metodológicos que possam permitir se obter os resultados mais robustos.

Referências

- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 31, 307-327.
- Brigida, M. (2014). The switching relationship between natural gas and crude oil prices. *Energy Economics*, 43, 48-55.
- Brown, S. P. A. & Yücel, M. K. (2008). What drives natural gas prices? *The Energy Journal*, 28(1), 45-60.
- Dickey, D. A. & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association*, 74, 427-431.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1007.
- Engle, R. F. & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, 5, 1-50.
- Engle, R. F. & Granger, C. W. (1987). Co-integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, 55, 251-276.
- Frey, G., Manera, M., Markandya, A. & Scarpa, E. (2009). Econometric models for oil price forecasting: A critical survey. *CESifo Forum*, 10(1), 29-44.
- Granger, C. (1969). Investigating Causal Relations by Econometric Models and Cross-Spectral Methods. *Econometrica*, 37(3), 424-438.
- Granger, C. & Newbold, P. (1976). R^2 and the Transformation of Regression Variables. *Journal of Econometrics*, 4, 205-210.
- Gujarati, D. (2004). *Basic Econometrics*. 4th Ed, The McGraw-Hill Companies, New York.
- Hartley, P. R. & Medlock III, K. B. (2014). The relationship between crude oil and natural gas prices: the role of the exchange rate. *The Energy Journal*, 35(2), 25-44.
- Leykam, K. & Frauendorfer, K. (2008). Co-integration and volatility in the European natural gas spot markets. Master Thesis, University of St. Gallen, St. Gallen, Switzerland.
- Nick, S. & Thoenes, S. (2014). What drives natural gas prices? - A structural VAR approach. *Energy Economics*, 45, 517-527.
- Panagiotidis, T. & Rutledge, E. (2007). Oil and gas markets in the UK: evidence from a cointegrating approach. *Energy Economics*, 29 (2), 329-347.
- Pickup, M. (2014). *Introduction to time series analysis*. Sage Publications, Washington DC.
- Union of Concerned Scientists. (2014). *Uses of Natural Gas*. Available at <http://www.ucsusa.org/clean-energy/coal-and-other-fossil-fuels/uses-of-natural-gas>.
- Wooldridge, J. M. (2009). *Introductory Econometrics: A modern approach*. 4th Ed., Mason, South Western Cengage Learning.
- Yule, G. U. (1926). Why do we sometimes get nonsense-correlations between Time-Series? - a study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, 89 (1), 1- 64.